

# Annexe C

## Suivi et évaluation :

### Notes techniques et études de cas

Note technique C.1	Principaux types d'évaluation .....	1
Note technique C.2	Conceptions des évaluations de l'impact.....	3
Note technique C.3	Méthodes d'évaluation de l'impact des politiques et programmes à couverture totale.....	4
Note technique C.4	Types de sources de données pour l'évaluation de l'impact.....	4
Étude de cas C.1	Suivi des progrès du plan d'action d'élimination de la pauvreté en Ouganda.....	5
Étude de cas C.2	Proposition de plan de suivi de la stratégie de lutte contre la pauvreté en Tanzanie.....	13
Étude de cas C.3	Enquête de feedback des citoyens comme outil de participation de la société civile à l'évaluation des performances du secteur public : le cas de Bangalore en Inde .....	20
Étude de cas C.4	Évaluation des bénéficiaires pour les pauvres des programmes d'assistance-travail : le programme Trabajar en Argentine.....	21
Étude de cas C.5	Évaluation du projet de vulgarisation agricole du Kenya .....	24
Étude de cas C.6	Évaluation de la réforme scolaire au Nicaragua : une approche quantitative et qualitative .....	28
Étude de cas C.7	Résultats scolaires dans les écoles élémentaires philippines : évaluation de l'impact de quatre expériences .....	31
Notes	.....	34

#### **Tableaux**

C.1. Récapitulatif des études de cas pour les évaluations de l'impact .....	6
C.2. Liste revue des indicateurs de suivi, Ouganda.....	9
C.3. Proposition d'indicateurs de suivi du CSLP Tanzanie.....	15
C.4. Sources d'informations pour le suivi, Tanzanie .....	17

#### **Figures**

C.1. Cadre institutionnel pour le suivi du CSLP, Tanzanie.....	20
--	----

#### **Encadré**

C.1. Objectifs du plan d'action ougandais pour l'élimination de la pauvreté.....	7
--	---

### **Note technique C.1 Principaux types d'évaluation**

L'évaluation consiste à examiner de façon systématique la pertinence, l'application et les résultats de programmes et politiques, en les comparant à un ensemble de normes explicites ou implicites, de manière à améliorer les actions publiques. Les différentes questions liées à l'évaluation sont traitées par des types d'évaluation différents. On peut les classer dans trois grandes catégories :

- les questions de **processus** cherchent à comprendre si l'application du programme ou de l'une de ses composantes est conforme au concept original.
- Les questions de résultat cherchent à déterminer si la situation des personnes et des ménages sur le plan des résultats essentiels (savoir, comportement, bien-être, etc.) a changé, et dans quelle mesure les changements observés proviennent du programme. En effet, les résultats peuvent changer pour de nombreuses raisons, dont plusieurs sont indépendantes du programme. Les

questions d'attribution étudient si les changements observés sont causés par le programme ou s'ils seraient survenus de toute manière.

- Les questions relatives aux **raisons** cherchent à sonder les motifs qui sous-tendent le processus et les résultats observés ; elles demandent pourquoi on a obtenu ces résultats et comment.

On peut associer à ces questions trois grands types d'évaluation : évaluation du processus, évaluation des résultats et évaluation basée sur une théorie. Chaque type d'évaluation comporte à son tour plusieurs styles d'évaluation et méthodes de collecte possibles. Les styles d'évaluation rassemblent plusieurs techniques qu'il est possible d'utiliser selon différentes combinaisons de manière à répondre aux différentes questions d'évaluation. Ils précisent les unités d'analyse (ménages, individus, installations, communautés, etc.) et leur mode de sélection (par des stratégies opportunistes ou d'échantillonnage systématique) ; le type de comparaison à réaliser (sans comparaison, comparaison dans le temps et l'espace, comparaison entre groupes différents, etc.) ; et le calendrier de collecte des données (avant et après le programme, immédiatement après seulement, pendant sa mise en œuvre, etc.).

**Les évaluations de processus** s'intéressent à l'efficacité de la mise en œuvre d'une action publique ; elles s'axent sur différents aspects : qui participe, quelles sont les activités proposées, quelles ont été les actions prises, quelles sont les pratiques du personnel et les réactions des clients. On peut réaliser une évaluation du processus lorsque le système de suivi a révélé des problèmes (retards, surcoûts ou insatisfaction des bénéficiaires par exemple), ou régulièrement comme système d'avertissement précoce. Les évaluations de processus tendent à dépendre de styles d'évaluation et de modes d'enquête moins formels tels que l'auto-évaluation et le jugement d'experts.

**Les évaluations des résultats** déterminent ce qui est arrivé aux individus (ou à d'autres unités d'analyse) après la mise en œuvre de la politique ou du programme ; elles s'axent sur les résultats des interventions en cherchant à savoir par exemple si les individus sont en meilleure santé, plus instruits ou moins vulnérables aux chocs négatifs. Les styles d'évaluations des résultats varient selon un continuum de niveaux de complexité. On trouve à une extrémité du spectre les évaluations des résultats uniquement chargées de déterminer si les indicateurs clés du bien-être des participants au programme ont évolué : il s'agit en général d'exercices de suivi. Les styles d'évaluation et les méthodes de collecte et d'analyse des données dont on dispose à cette extrémité du spectre tendent à être relativement simples et à permettre d'obtenir rapidement des résultats ; ils laissent néanmoins une certaine marge pour des interprétations divergentes sur le degré auquel il est possible d'attribuer un changement à une intervention donnée. Ce type d'évaluation ne s'intéresse généralement qu'au groupe de participants au programme ; il ne réalise aucune comparaison avec les personnes et les communautés qui n'y ont pas participé et ne cherche pas à isoler les effets du programme ou de la politique d'autres événements survenant simultanément. L'évaluation peut étudier les résultats une fois que l'intervention est déjà lancée depuis un certain temps ou bien lorsqu'elle est achevée, ou encore avant et après. Les méthodes de collecte et d'analyse des données peuvent être quantitatives (enquête sur la prestation de services par exemple) ; qualitatives (entretiens avec des répondants clés ou des groupes types) ; ou participatives (méthodes d'estimation rapide).

À l'autre extrémité du spectre, les évaluations s'intéressent aux questions d'attribution en faisant intervenir des techniques spéciales (souvent complexes) afin de différencier les gains nets des interventions (voir notes techniques 2 et 3). Ce sont les évaluations de l'impact. Ces évaluations estiment le degré auquel des actions publiques ont produit les effets voulus et auquel il est possible d'attribuer un changement dans le bien-être des individus à un programme ou à une politique donnée. Elles estiment l'ampleur des effets d'un programme ou d'une politique et attribuent des liens de causalité. L'analyse des causes est essentielle si l'on veut comprendre le rôle relatif d'autres interventions programmées sur la réduction de la pauvreté et donc élaborer des stratégies appropriées de lutte contre la pauvreté.

**Les évaluations basées sur une théorie** étudient les liens entre les variables d'entrée, les activités et les résultats et cherchent à déterminer s'il y a une dégradation, et le cas échéant, où, pourquoi et comment. Elles donnent une théorie explicite ou implicite du comment et du pourquoi de l'efficacité d'une action publique en la présentant comme une série de petites étapes, et en font une analyse séquentielle afin de suivre le développement des hypothèses. En suivant ces étapes, ce type d'évaluation est capable de dire si le processus, des variables d'entrée jusqu'aux résultats du programme, a échoué, et le cas échéant, où.

## Note technique C.2 Conceptions des évaluations de l'impact

**Les styles expérimentaux ou aléatoires** consistent à réunir un ensemble d'individus (ou d'autres unités d'analyse) présentant les mêmes conditions d'éligibilité et acceptant de participer au programme, et à les répartir en deux groupes : ceux qui bénéficieront de l'intervention (groupe traité) et ceux qui n'en bénéficieront pas (groupe témoin). Par exemple dans le cas de certains fonds sociaux, on choisit au hasard parmi des projets économiques faisables soumis par les communautés les projets qui recevront les financements dans la première phase du projet (groupe traité), et les autres, qui recevront ce financement ultérieurement, servent de groupe témoin. Étant donné que les participants au programme sont choisis de manière aléatoire, toute différence par rapport aux participants ne bénéficiant pas du programme est due au hasard. C'est pourquoi l'on considère généralement que les évaluations expérimentales constituent la méthode la plus fiable, et qu'elles génèrent des résultats faciles à interpréter. En pratique cependant, ce style d'évaluation peut s'avérer malaisé à mettre en œuvre, notamment parce qu'il est difficile de ne pas accorder les mêmes bénéfices à des personnes tout aussi éligibles (voir étude de cas 7).

**Le style quasi-expérimental** est une autre option. Lorsqu'il n'est pas possible d'appliquer la méthode aléatoire, on peut constituer un groupe de comparaison. Pour cela il existe deux méthodes : la comparaison correspondante et la comparaison réflexive. La première consiste à sélectionner des non-participants présentant des caractéristiques essentielles comparables à celles des participants, d'après quelques caractéristiques seulement ou plusieurs d'entre elles, à l'aide de techniques statistiques. Par exemple, pour l'évaluation de Trabajar, programme argentin de grands travaux, on a élaboré un groupe de comparaison en appariant les participants au programme et les non-participants sur la base de plusieurs caractéristiques socioéconomiques, dont la scolarisation, le sexe, le logement, la perception subjective du bien-être et l'adhésion à des partis politiques (voir étude de cas 4). Il est souvent plus facile et moins onéreux de réaliser des évaluations selon cette méthode plutôt que par le style expérimental, mais les résultats sont moins fiables et parfois plus difficiles à interpréter.

La *comparaison réflexive* est un autre style d'évaluation quasi-expérimental. Dans une comparaison réflexive, le groupe fictif est élaboré d'après la situation que les participants au programme connaissent avant le programme. Ainsi, les participants sont comparés à eux-mêmes avant et après l'intervention et représentent à la fois le groupe traité et le groupe de comparaison. Ce style d'évaluation est particulièrement utile pour évaluer les interventions à couverture totale telles que les politiques et programmes nationaux auxquels toute la population participe et pour lesquels on ne peut constituer de groupe témoin (voir étude de cas 5). Cette méthode présente cependant un inconvénient majeur : la situation des participants avant et après l'intervention peut évoluer du fait de raisons indépendantes du programme. Par exemple il est possible que les perspectives d'emploi de participants à un programme de formation soient meilleures après le programme. Si cette amélioration peut provenir du programme, elle peut également être le fruit du redressement de l'économie après une crise et de la reprise de l'emploi. À moins d'être soigneusement réalisées, les comparaisons réflexives peuvent avoir des difficultés à différencier les effets du programme de ceux qui lui sont externes, ce qui compromet la fiabilité des résultats.

On peut utiliser le style non expérimental lorsqu'il est impossible de choisir un groupe témoin ou un groupe de comparaison. Les participants au programme peuvent être comparés aux non-participants à l'aide de méthodes statistiques destinées à expliquer les différences entre les deux groupes. Par le biais d'une analyse de régression, on peut « contrôler » les âge, revenu, sexe et autres caractéristiques des participants. Comme pour toute méthode quasi-expérimentale, ce style d'évaluation est relativement peu onéreux et facile à mettre en œuvre, mais puisque l'interprétation des résultats n'est pas évidente, les résultats risquent d'être moins fiables.

## Note technique C.3 Méthodes d'évaluation de l'impact des politiques et programmes à couverture totale

Lorsque des politiques ou des programmes concernent toute la population, il n'est en général pas possible d'identifier ou de constituer un groupe témoin ; dans ces conditions, il est extrêmement difficile de décider si ces interventions ont eu une influence sur les résultats. On peut alors utiliser plusieurs méthodes.

Les modèles d'équilibre général calculable (EGC) tentent de comparer des situations observées et fictives grâce à des simulations sur ordinateur. Ces modèles cherchent à suivre le fonctionnement de l'économie réelle et se basent généralement sur des matrices de comptabilité sociale détaillées fondées sur des données tirées de la comptabilité nationale, des enquêtes sur les dépenses des ménages et d'autres enquêtes. Les modèles EGC simulent la situation fictive, mais la solidité du modèle dépend entièrement de la qualité des données sous-jacentes et de la validité des hypothèses. Cela peut s'avérer problématique, dans la mesure où les bases de données sont souvent incomplètes et où bon nombre des paramètres requis ne peuvent pas être estimés via des méthodes économétriques claires. Les modèles EGC sont aussi longs à réaliser, peu faciles à manier et coûteux à développer.

Les comparaisons « avec et sans » comparent le comportement de variables clés dans un échantillon de pays ou de régions concernés par un programme par rapport au comportement de pays (ou de groupes de comparaison) non inclus dans ce programme. Ainsi, cette méthode fait appel aux expériences des pays non inclus dans le programme comme moyen de comparaison avec ce qui se serait autrement passé dans les pays concernés par le programme. Cette approche présente une restriction importante, à savoir qu'elle part de l'hypothèse que seule l'adoption d'une politique ou d'un programme donné différencie les pays ou régions concernées par le programme des zones non concernées et que soit les facteurs externes affectent également les deux groupes soit leur impact peut être isolé de celui de l'intervention.

Les méthodes de contrôle statistique sont des régressions qui contrôlent les différences dans les conditions initiales et dans les politiques engagées dans les pays ou régions qui bénéficient et ne bénéficient pas du programme. Cette approche identifie les différences entre les secteurs bénéficiant et ne bénéficiant pas du programme dans la période qui précède le programme puis réalise un contrôle statistique de ces différences afin d'identifier les effets isolés des programmes durant la période qui fait suite au programme.

Source : Baker, Judy. 2000. *Evaluating the Poverty Impact of Projects: A Handbook for Practitioners*. Directions in Development. Washington, D.C.: World Bank.

## Note technique C.4 Types de sources de données pour l'évaluation de l'impact

Les ensembles de données longitudinales ou de panel rassemblent des informations sur les mêmes individus (ou autres unités d'analyse) à au moins deux points différents dans le temps, un avant l'intervention (référence) et l'autre après. Les ensembles de données sont très précieux pour évaluer un programme, mais peuvent s'avérer coûteux et exiger une importante capacité institutionnelle (voir étude de cas 5 et chapitre 1, « Mesure et analyse de la pauvreté »).

Les données transversales réunissent des informations sur différentes personnes à différents moments dans le temps. Les évaluations utilisant des données transversales sont généralement moins chères que celles qui utilisent des informations collectées plus d'une fois mais, vu qu'il est souvent difficile de décider si les changements proviennent de l'intervention ou d'autres facteurs, les résultats tendent à être moins fiables, à l'exception des styles expérimentaux (voir étude de cas 4).

Les données de séries chronologiques collectent des informations sur les mesures de résultats clés à intervalles réguliers, à la fois avant et après le programme. Elles permettent d'étudier les changements de tendances avant et après le programme. Cependant, pour réaliser une analyse rigoureuse, il faut plusieurs points de collecte des données avant et après le programme. Les séries chronologiques sont principalement utilisées pour évaluer les politiques et programmes à couverture totale ou nationale.

## Études de cas

Les études de cas C.1 et C.2 donnent des exemples de systèmes nationaux de suivi de la pauvreté, alors que l'étude de cas C.3 illustre le recours à des enquêtes de feedback des citoyens comme outil permettant la participation de la société civile dans l'évaluation des performances du secteur public. Les études de cas C.4 à C.7 (adaptées de Judy Baker [2000], *Evaluating the Poverty Impact of Projects: A Handbook for Practitioners*, World Bank, Washington, D.C.) illustrent les évaluations de l'impact de projets et programmes dans différents secteurs. Elles décrivent un large éventail d'approches pour les styles d'évaluation, l'utilisation des données, la pertinence des résultats au niveau de la politique ainsi que l'impact sur l'évaluation de l'édification de la capacité (voir tableau C.1). Chaque étude est accompagnée d'une discussion sur les points forts et les points faibles de chaque type d'évaluation.

### Étude de cas C.1 Suivi des progrès du plan d'action d'élimination de la pauvreté en Ouganda<sup>1</sup>

#### C.1.1 Introduction

En 1995, le gouvernement d'Ouganda s'est lancé dans la formulation d'un plan d'action pour l'élimination de la pauvreté (PAEP) pour s'assurer que l'allègement de la pauvreté était au centre de sa stratégie globale de croissance et de développement. Ce plan a été élaboré par le biais d'un processus consultatif faisant intervenir des représentants du gouvernement et de la société civile de même que des organisations donatrices. L'objectif principal du PAEP est d'éliminer la pauvreté de masse, en réduisant la proportion de la population vivant dans des conditions de pauvreté extrême de 56 (1992) à 10 % et celle de personnes vivant dans des conditions de pauvreté relative de plus de 85 % à 30 % d'ici 2017.

Il a convenu d'autres objectifs dans quatre domaines (macroéconomie, gouvernance, génération de revenus et développement humain), qui ont été développés en un ensemble d'objectifs stratégiques (voir encadré C.1). Les soins élémentaires de santé, l'éducation primaire, la vulgarisation agricole et les routes de desserte rurale ont été identifiés comme secteurs de pauvreté prioritaires pour l'allocation des ressources. La définition d'objectifs et le choix d'objectifs stratégiques et de domaines prioritaires ont été des processus dynamiques, souvent revus à la lumière d'informations nouvelles, comme le projet ougandais d'évaluation participative de la pauvreté (POEPP) réalisé en 1998 et 1999, et du feedback apporté par le système de suivi de la pauvreté.

#### C.1.2. Système de suivi de la pauvreté

Le suivi continu de la pauvreté permet d'évaluer le chemin parcouru vers la satisfaction des objectifs. Activité ponctuelle au départ, elle a progressivement évolué vers un système de suivi participatif décentralisé délimitant plus clairement les rôles et les responsabilités, ainsi que les mécanismes de collaboration entre les institutions.

Ce système est composé de trois grandes institutions.

- Le bureau ougandais des statistiques (UBoS), qui réunit, analyse et publie les données provenant des enquêtes auprès des ménages.
- Les départements des statistiques des ministères d'exécution, qui réunissent et analysent les données sectorielles provenant des systèmes intégrés de gestion.
- L'unité de suivi de la pauvreté, dont la fonction principale consiste à relier les producteurs de données aux décideurs. Elle collecte des données sur la pauvreté auprès de diverses sources dont l'UBoS, les ministères d'exécution et d'autres organisations et institutions externes au gouvernement; elle analyse les données, diffuse les résultats et débat des tendances et des perspectives de la pauvreté avec les représentants et agences du gouvernement. L'unité va prochainement se développer et aura également pour mandat de réaliser l'analyse de la politique de réduction de la pauvreté. L'unité est située dans le ministère des finances, du plan et du développement économique, emplacement clé pour influencer la politique.

En outre, d'autres institutions telles que les organisations non gouvernementales (ONG), les institutions universitaires, les centres de recherche et les donateurs, jouent des rôles importants, mais non systématiques, dans la collecte et l'analyse de données supplémentaires. Les décideurs constituent également une partie essentielle du système en qualité d'utilisateurs principaux des résultats du suivi (notamment au niveau central, bien que l'on ait compris qu'il fallait également utiliser les statistiques sur les données collectées au niveau local dans la prise de décision locale).

Tableau C.1. Récapitulatif des études de cas pour les évaluations de l'impact

Programme/ projet	Pays	Type de base de données	Unité d'analyse	Mesures des résultats	Approche économétrique				Évaluation qualitative	Points forts
					Répartition aléatoire	Comparai- son correspon- dante	Comparai- son réflexive	Variables instru- mentales		
<b>Enseignement</b>										
Réforme de l'autonomie scolaire	Nicaragua	Enquête de panel et évaluations qualitatives	Élèves, parents, enseignants, directers	Résultats d'examens, degré de prise de décision locale	Non	Oui	Oui	Non	Oui	Mélange qualitatif / quantitatif
Intervention au niveau des abandons	Philippines	Enquête de référence et postérieure à l'intervention	Élèves, classes, enseignants	Résultats des examens et état des abandons	Oui	Non	Oui	Oui	Non	Analyse coût/béné- fice ; édification de la capacité
<b>Programmes sur le travail</b>										
Programme Trabajar	Argentine	Enquêtes auprès des ménages, recensement, documents administratifs, évaluations sociales	Travailleurs, ménages	Revenu, ciblage, coûts	Non	Oui	Non	Oui	Oui	Utilisation judicieuse des sources de données existantes, techniques d'analyse novatrices
<b>Agriculture</b>										
Projet national de vulgarisa- tion	Kenya	Données de panel, évaluations des bénéficiaires	Ménages, fermes	Productivité et efficience agricoles	Non	Non	Oui	Non	Non	Pertinence des résultats au niveau de la politique

Source: : Adapté de Baker, Judy. 2000. *Evaluating the Poverty Impact of Projects: A Handbook for Practitioners*. Directions in Development. Washington, D.C.: World Bank.

Ce système subit une révision majeure qui a pour but :

- d'accroître la participation, autrement dit de promouvoir une plus grande implication dans les activités de suivi au niveau local et une meilleure collaboration entre l'UBoS, l'unité de suivi de la pauvreté, les ONG et les ministères d'exécution pour la collecte, l'analyse et la diffusion des données. Les liens entre les agences de districts et centraux chargés d'étudier les statistiques sont également en cours de révision ;
- de développer la capacité, notamment pour le suivi au niveau local (district), et l'analyse des données et leur diffusion au niveau central, afin de réduire le délai entre la collecte des données et leur analyse et diffusion ;
- de définir les rôles institutionnels, autrement dit, d'explicitier les rôles et les responsabilités, y compris les mécanismes de collaboration ;
- d'harmoniser les rapports d'avancement, autrement dit, de définir un format commun pour la rédaction de rapports d'avancement des programmes sectoriels et de pauvreté.

#### **Encadré C.1. Objectifs du plan d'action ougandais pour l'élimination de la pauvreté**

**Objectif primordial : réduire la proportion de la population vivant dans des conditions de pauvreté extrême de 56 à 10 % et celle de personnes vivant dans des conditions de pauvreté relative de plus de 30 % d'ici 2017.**

**Objectif 1 : Mise en œuvre de politiques macroéconomiques permettant un environnement favorable à la réduction de la pauvreté**

- Préserver un taux de change stable, pour un secteur de l'exportation compétitif
- Préserver de faibles niveaux d'inflation pour faciliter la mobilisation de l'épargne et la planification à long terme
- Promouvoir des investissements du secteur privé dans les zones rurales
- Réduire les discriminations anti-exportations de la politique commerciale de manière à améliorer les perspectives d'exportation
- Promouvoir une croissance économique qui s'appuie sur une large base
- Réduire l'endettement externe à un niveau soutenable
- Réduire les écarts de pauvreté entre les districts
- Améliorer l'autonomisation économique et politique des femmes
- Élargir l'assiette fiscale
- Recentrer les dépenses publiques pour qu'elles soient directement liées à l'élimination de la pauvreté

**Objectif 2 : Création d'un cadre institutionnel favorable à la réduction de la pauvreté via une vaste participation, transparence et prise de responsabilité**

- Améliorer la fourniture réelle et efficace de services publics tout en encourageant la transparence et la responsabilisation
- Promouvoir la croissance du secteur privé en améliorant les investissements locaux et étrangers
- Renforcer les dispositifs de maintien de l'ordre, de respect de la loi et d'exercice de la justice tout en améliorant l'accès des pauvres aux services légaux
- Améliorer le respect des droits de l'homme, de la liberté et de la gouvernance démocratique
- Encourager la participation de la communauté à la planification et à la prestation des services

**Objectif 3 : Augmentation des opportunités de revenus pour les pauvres**

- Créer un réseau routier efficace
- Transformer et moderniser la production agricole
- Garantir la sécurité de la possession du sol, l'accessibilité adéquate à la terre et son usage efficient, tout en préservant l'environnement
- Soutenir le développement de marchés ruraux : infrastructure, informations commerciales et normes
- Proposer aux pauvres des services financiers via la promotion de la croissance d'institutions de microfinance et de banques de villages de campagne
- Améliorer la productivité du travail, en donnant la priorité à l'emploi des femmes, à la réduction de l'exploitation des enfants au travail et en préservant les droits des employés
- Créer un environnement favorable au développement de microentreprises et de petites entreprises

**Objectif 4 : Amélioration de la qualité de vie et du capital humain des pauvres**

- Satisfaire le droit constitutionnel de soins de santé élémentaires pour tous, en améliorant l'apport rentable de services de santé à toute la population
- Fournir à toute la population un accès facile à une source d'eau potable, tout en améliorant la rentabilité de l'approvisionnement en eau
- Parvenir à une éducation primaire pour tous et améliorer la qualité de l'éducation
- Promouvoir l'accès à une éducation de base pour les enfants vulnérables (les sans domicile fixe ou à la rue par exemple)
- Encourager pour tous les Ougandais l'acquisition d'un niveau d'éducation fonctionnel, son utilisation et sa préservation

L'une des options envisagées pour traiter certaines de ces questions consiste à établir une organisation de terrain dépendant de l'UBoS. Cette organisation sera chargée de contrôler le flux d'informations en provenance et en partance du siège ; de protéger le développement de statistiques à l'échelle du district ; de recruter, former et superviser le personnel de terrain ; de planifier le travail de terrain ; de collecter et de saisir les données ; et de réaliser toute autre fonction liée au travail de terrain. Six zones statistiques devront être créées. Chacune aura un bureau de zone dans lequel travaillera un petit nombre de personnel permanent (superviseur de zone, assistant statistique et opérateur de saisie de données) ainsi que des superviseurs et des enquêteurs de terrain qui seront recrutés sur une base temporaire.

Le ministère des finances, du plan et du développement économique a préparé en octobre 2001 une stratégie de suivi et d'évaluation de la pauvreté afin de débattre de certaines de ces questions et des solutions possibles.

### **Indicateurs**

Le choix des indicateurs a été le fruit d'un processus itératif. À l'origine, on choisissait les indicateurs d'après les travaux de groupes thématiques afin de suivre les progrès dans divers domaines : pauvreté monétaire, état de santé, éducation, environnement, infrastructure, gouvernance, emploi et accès aux informations, aux marchés et au crédit. La première liste d'indicateurs a été jugée trop longue, incomplète dans certains domaines et non concentrée sur les priorités. Travaillant au cadre de dépenses à moyen terme (CDMT), le groupe de travail pauvreté, composé de représentants du gouvernement, de la société civile et d'organisations donatrices, l'a peaufinée (voir tableau C.2). La liste sera encore affinée pour qu'elle reste cohérente avec le PAEP une fois amendé. Aujourd'hui, pratiquement tous les indicateurs sont suivis au niveau national ; un sous-ensemble est suivi au niveau du district et/ou de la région. Les données relatives à l'éducation sont les seules à être décomposées par sexe. Cela impose une contrainte majeure sur l'analyse exhaustive de la pauvreté. On s'attend à ce que ce problème soit résolu dans l'avenir.

En dépit d'un certain progrès vers l'alignement des indicateurs avec les objectifs du PAEP, il reste des points à améliorer. Plusieurs indicateurs sont définis en termes de nombre de cas. Si les quantités réelles sont importantes, les pourcentages et les ratios se révéleront souvent de meilleurs indicateurs. Par exemple, la proportion d'unités sanitaires disposant de médicaments essentiels constitue un meilleur indicateur que le seul nombre d'unités. Le fait que certains indicateurs ne constituent pas des mesures de progrès sans équivoque est un autre problème, car il n'est pas possible de déterminer, d'après ces indicateurs, si la situation s'est améliorée ou non. Par exemple, une hausse des dépenses des ménages consacrées à l'éducation n'est pas un indicateur d'amélioration sans équivoque. Peut-être est-elle due à une plus forte consommation des ménages ou au fait que comme les dépenses d'éducation sont supérieures, les ménages consomment moins. Enfin, la liste ne fait pas la distinction entre les indicateurs finaux et intermédiaires, distinction qui serait utile pour juger du progrès global.

Tableau C.2. Liste revue des indicateurs de suivi, Ouganda

<i>Indicateurs</i>	<i>Niveau de décomposition souhaité</i>
<b>PAUVRETÉ MONÉTAIRE</b>	
Proportion de la population vivant sous le seuil de pauvreté	National, régional, district
Nombre de personnes vivant dans des conditions de pauvreté extrême	National, régional
Part des dépenses alimentaires des ménages	National, régional
Proportion de la population vivant sous des toits de chaume	National, régional
Ratio de dépendance	National, régional, district
Coefficient de Gini	National, rural/urbain
Consommation par habitant pour les 20 % les plus pauvres	National, régional, district
PIB par habitant	National
Ratio épargne/PIB	National
Revenu par habitant et par district	District
Sécurité et vulnérabilité	
Proportion de ménages affectés par des vols ou des troubles civils	National, régional
<b>QUALITÉ DE VIE</b>	
Nombre de réfugiés dans le pays	National, régional
Nombre de décès de civils du fait d'une insurrection	National, régional
Nombre de crimes et délits signalés	National, régional
Proportion de ménages ayant connu un choc monétaire majeur l'an dernier	National, régional
Nombre de réfugiés en proportion de la population du district	District
Proportion de ménages en détresse économique vendant des actifs	National
Réseau routier	
Longueur de routes ouvertes	National
Longueur de route améliorées	National
Proportion de districts possédant plus de 50 % de routes en mauvais état	National, district
Proportion de zones non desservies par des routes	National, district
Terre	
Incidence de la pauvreté par possession de sol	National, district
Agriculture	
Taux d'adoption des méthodes d'exploitation modernes	National, district
Rendements	National, district
Pourcentage de fermiers cultivant des cultures de sécurité alimentaire	National, district
Marchés	
Disponibilité des marchés par type	National, district
Accessibilité des marchés	National, district
Volume de biens et services gérés dans un marché donné	National, district
Proportion de ménages dans lesquels le prix de vente du produit agricole principal est inférieur à 50 % du prix du marché urbain	National, district
Productivité de la main-d'œuvre et emploi	
Taux de chômage	National, district
Inscription à une formation professionnelle	National, district
Nombre moyen d'heures travaillées par jour	National, district
Crédit rural	
Développement du portefeuille de microfinance	National, district
Proportion de la population ayant accès au microcrédit	National, district
Croissance de l'épargne	National, district
Gestion du crédit (utilisation efficace)	National, district
Disponibilité des services de microfinance	National, rural/urbain

**Tableau C.2. Liste revue des indicateurs de suivi, Ouganda (suite)**

<i>Indicateurs</i>	<i>Niveau de décomposition souhaité</i>
<b>Indicateurs de santé</b>	
Incidence des maladies	National, district
Couverture vaccinale	National, district
Proportion de la population située dans un rayon de 5 km de l'unité de santé la plus proche	National, district
Dépenses de santé du ménage par habitant	National, district
Nombre d'unités de santé disposant de médicaments essentiels	National, district
Nombre de secteurs avec plus de 1 000 personnes par personnel de santé compétent	National, district
Couverture des soins anténataux	National, district
<b>Eau et assainissement</b>	
Proportion de la population située dans un rayon de 0,5 km de la source d'eau potable par région	National, district rural/urbain
Proportion de la population ayant des latrines en bon état et salubres	National, rural/urbain
Évacuation sûre des déchets	National
<b>Indicateurs d'éducation</b>	
Ratio net d'inscription primaire	National, district, sexe
Proportion d'élèves d'écoles primaires suivant plus de quatre années de scolarité	National, district, sexe
Ratio élèves-enseignants qualifiés	National, district, sexe
Distance par rapport aux écoles	National, district, sexe
Ratio élèves-classes	National, district, sexe
Ratio élèves-manuels	National, district, sexe
Dépenses d'éducation du ménage par habitant	National, district, sexe
<b>ENVIRONNEMENT</b>	
Niveau de conformité avec les normes environnementales	National pour tous
Actions correctives mises en œuvre par l'agence nationale pour la gestion de l'environnement	
Proportion de la population pratiquant des méthodes soutenables d'utilisation des sols	
Allocations budgétaires aux programmes environnementaux par les gouvernements locaux	
Proportion de terres classées par districts	
<b>GOVERNANCE ET RESPONSABILISATION</b>	
Niveau de connaissance qu'a la population de ses droits	National
Proportion de cas signalés et résolus	National
Nombre de personnes en arrêt de renvoi au delà de la période légale	National
Nombre de procès en retard	National
Cas de corruption soulevés à différents niveaux	National
Programmes d'allègement de la pauvreté réussis	National, district
Nombre de cas de corruption/détournement de fonds et d'abus de privilège aboutissant à une inculpation	National

## Collecte des données

Les enquêtes auprès des ménages, les systèmes intégrés de gestion et les études qualitatives constituent les principales sources de données de suivi.

**Les enquêtes auprès des ménages** sont planifiées au niveau central et mises en œuvre par l'UBoS avec une consultation ou une participation limitée au niveau du district. Le rôle des districts est actuellement à l'étude, dans le but d'élaborer une capacité locale et de promouvoir un accès rapide à des informations spécifiques au district que ceux-ci peuvent utiliser pour planifier, mettre en œuvre et suivre leurs programmes et politiques. L'UBoS et le ministère du gouvernement local travaillent à un système destiné à faire participer les unités de planification du district à la collecte des données de manière à ce que les statistiques et les informations qualitatives pertinentes soient utilisées pour suivre les performances au niveau du district. Un système de ce genre permettrait de compléter le système de collecte des données des ménages géré au niveau central. Les enquêtes auprès des ménages suivantes sont utilisées pour assurer le suivi de la pauvreté.

**Les enquêtes intégrées auprès des ménages (EIM)** rassemblent des données sur les caractéristiques des ménages et du logement ; le revenu et les dépenses du ménage ; les actifs, emprunts et épargne ; la production agricole ; et l'état sanitaire et nutritionnel des enfants. L'enquête réalisée en 1992 a donné des informations de référence sur 10 000 ménages dans tout le pays. Le questionnaire de l'enquête a été réexaminé compte tenu des données de l'évaluation participative de la pauvreté en Ouganda et intègre maintenant des questions sur des sujets tels que la sécurité des ménages. L'enquête revue (Enquête nationale ougandaise auprès des ménages) a été réalisée en 2001/2002 et une nouvelle session est prévue en 2003/2004.

**Les enquêtes de suivi** rassemblent des informations semblables à celles collectées par l'EIM, mais portent sur un plus petit échantillon de 5000 ménages et utilisent un questionnaire plus court (sans module de consommation). Elles ont été réalisées avec une périodicité annuelle de 1992-1993 à 1997.

**Les enquêtes démographiques et de santé (EDS)** rassemblent des informations sur la santé des mères et des enfants, la vaccination, l'accès aux soins de santé, l'incidence des grandes maladies etc. Les données de référence ont été obtenues lors de l'enquête de 1995 et une enquête de suivi a été réalisée en 2001.

En outre, le recensement national de 2002 apportera des informations actualisées sur la structure démographique de la population : âge, statut marital, ethnicité, religion, taille du ménage, taux de dépendance, etc. D'autres enquêtes, comme l'enquête de suivi des dépenses publiques (voir chapitre 6, « Dépenses publiques ») et l'enquête nationale sur la prestation de services, ont généré des informations utiles mais n'ont pas encore été intégrées au système de suivi régulier. L'enquête nationale sur la prestation de services rassemble des informations sur l'utilisation des services publics et le degré de satisfaction que l'on en tire. Cette enquête a été fait l'objet d'un essai pilote en 1996 avant d'être réalisée en 1999 à l'échelle nationale (aujourd'hui par le ministère des services publics mais dans l'avenir par le bureau des statistiques). Une enquête sur les installations sanitaires a été récemment réalisée.<sup>2</sup>

**Les systèmes intégrés de gestion (SIG)** collectent des informations sectorielles sur les résultats, l'accès aux services, et, dans une certaine mesure, la qualité des services. Par exemple, dans le secteur de la santé, le SIG collecte des informations sur le nombre d'installations sanitaires par types (gestion publique ou privée de la capacité en lits) ; les installations proposant des services vitaux ; les effectifs et les grandes causes de morbidité. Dans le domaine de l'éducation, le ministère de l'Éducation et des sports réalise un recensement annuel sur l'éducation en collectant des informations au niveau du district sur le taux de scolarisation, le nombre d'enseignants, le matériel didactique, les installations et les finances.

On a identifié un certain nombre de problèmes que posent les données provenant des SIG. Tout d'abord, les informations sont incomplètes. En 1996, par exemple, le recensement de l'éducation a obtenu un taux de réponse de 60 % de la part d'institutions soutenues par le gouvernement et de 30 % des écoles privées. Deuxièmement, les données ne sont pas fiables. Dans le secteur de l'éducation, dépendre des directeurs d'école pour communiquer les données relatives à l'école est problématique, le nombre d'élèves étant souvent gonflé de manière à obtenir plus de subventions, pratique révélée lors de vérifications aléatoires. Dans le secteur de la santé, les outils d'aide au diagnostic, la capacité en termes d'effectif et l'infrastructure de communication sont limités dans de nombreuses régions, notamment dans les zones rurales isolées, de sorte que l'incidence des maladies est très rarement signalée. Troisièmement, il y a le problème de l'opportunité ; les données arrivent lentement des districts, et leur analyse, leur compilation et la rédaction de rapports au niveau du centre sont retardées. Les résumés statistiques sur l'éducation

par exemple, sont généralement réalisés un an et demi après la collecte des données. De sorte que les données ne sont pas utilisées pour la prestation de services et leur planification.

Les performances médiocres des SIG sont en partie dues à leur fort degré de centralisation. On demande aux districts de collecter des informations sans les faire participer en tant qu'acteur au processus de suivi. En conséquence, ils ont peu de raisons de veiller à la collecte opportune de données fiables. Les efforts réalisés pour corriger cette situation font intervenir des activités au niveau du district et du secteur avec les ministères d'exécution centraux. Au niveau du district, il s'agit de mettre en œuvre le système d'information sur les ressources du district (DRIS, District Resource Information System). Le DRIS est la deuxième phase d'une tentative antérieure qui visait à collecter des données sur les services sociaux et les infrastructures pertinentes auprès de tous les districts (District Resource Endowment Profile Study, ou DREPS). Il crée un lien direct entre les districts et l'UBoS et se concentre sur un grand nombre de variables, administration, prestation de services et infrastructure compris.

Quant aux **études qualitatives**, le projet ougandais d'évaluation participative de la pauvreté (POEPP) est la source principale de données qualitatives pour le suivi du PAEP. Il s'agit d'un projet sur trois ans destiné à intégrer les impressions des pauvres dans le dialogue local et national sur la réduction de la pauvreté et à mieux comprendre les tendances révélées par les données quantitatives. Le travail de terrain de la dernière évaluation participative sur la pauvreté a été réalisé fin 2001. Le POEPP est un partenariat entre le gouvernement, les donateurs, les neuf autorités des districts dans lesquels le projet est appliqué, et Oxfam, l'agence chargée de sa mise en œuvre.<sup>3</sup> Au sein du gouvernement, le POEPP est installé au ministère des finances, du plan et du développement économique dans le cadre de l'unité de suivi et d'analyse de la pauvreté.

L'agriculture et la gouvernance sont deux domaines qui requièrent plus d'efforts. En dépit de l'importante proportion de pauvres engagés dans des activités agricoles, les données pour l'agriculture sont relativement difficiles d'accès. Le programme de modernisation de l'agriculture, élément central d'un troisième pilier du PAEP, ne définit aucun indicateur, et le système d'information sur les ressources du district n'intègre aucune donnée sur l'agriculture. De même, il existe peu d'informations utiles sur les questions de gouvernance à l'exception de l'enquête sur l'intégrité nationale et des rapports de la Commission des droits de l'homme. Une quantité énorme de données a été collectée par l'inspectorat général sur la gouvernance, mais cette organisation manquant de personnel, il n'existe aucune statistique récapitulative. Les indicateurs du PAEP dans ce domaine n'ont pas été définis.

### **Analyse, diffusion des données et feedback**

L'analyse des données est réalisée principalement au niveau central par l'UBoS. Elle est limitée au niveau du district. Quelques districts ont lancé leurs propres systèmes de suivi dans le cadre du projet de développement du gouvernement local, dont l'objectif majeur est de renforcer la planification participative et le développement de systèmes de budgétisation et de suivi au niveau du district ; le projet a été récemment élargi des 5 puis 9 districts de l'origine à 30 de plus.

L'unité de suivi de la pauvreté rédige tous les deux ans un rapport sur l'état de la pauvreté afin d'évaluer les progrès et d'identifier les difficultés dans la mise en œuvre du PAEP. Ce rapport présente les progrès vers la satisfaction des objectifs du PAEP et indique également où en sont les actions d'élimination de la pauvreté, allocations budgétaires comprises. Ces informations servent ensuite à identifier les lacunes, les défis clés et les domaines prioritaires. Le prochain rapport est prévu en 2003.

Les rapports sont diffusés aux niveaux national et du district et utilisés pour réexaminer le PAEP, le CDMT et d'autres bilans sectoriels. L'unité de suivi de la pauvreté et le groupe de travail sur la pauvreté (GTP) veillent à ce que les résultats du suivi soient utilisés dans les processus relatifs aux décisions et aux allocations budgétaires. Plus spécifiquement, le GTP, auquel participent des agents du gouvernement et des représentants d'ONG et d'universités, veille à ce que les données collectées via le système de suivi de la pauvreté soient prises en compte et suivies par le groupe de travail sectoriel pertinent dans les processus de CDMT et de budgétisation. Le GTP fait également des recommandations sur l'allocation budgétaire globale des ressources pour la réduction de la pauvreté de même que sur d'autres politiques budgétaires affectant les pauvres. Malgré ces efforts, les performances des dépenses publiques restent mesurées en grande partie en termes d'entrées et d'activités plutôt que de contributions à la réduction de la pauvreté. Les progrès réalisés pour atteindre l'objectif continuent à jouer un rôle restreint dans le processus sectoriel d'allocation du budget. Il est nécessaire d'élaborer un système d'incitation, reliant l'allocation des ressources et l'évaluation des performances aux contributions aux résultats du PAEP.

### C.1.3 Édification de la capacité statistique

Les activités du système de suivi de la pauvreté sont étayées par un programme majeur destiné à actualiser les systèmes statistiques d'Ouganda. L'objectif principal du programme est d'édifier une capacité nationale permettant de collecter, de traiter, de stocker et de diffuser des informations statistiques pour le suivi et l'évaluation de la pauvreté à la fois au niveau national et du district. Le programme vise essentiellement à renforcer la capacité de l'UBoS à créer un programme statistique central permettant un suivi régulier et opportun des objectifs de développement nationaux. Il crée une nouvelle infrastructure des technologies de l'information pour le système intégré de gestion. Ce système a pour objet de veiller à ce que toutes les données directement collectées par l'UBoS ou que celui-ci reçoit d'autres sources comme données secondaires soient stockées au niveau central selon un format commun qui facilite l'accès libre aux données par les utilisateurs, que les données soient sur support papier ou informatique. Le dépôt de données central regroupe toutes les données nettoyées et prêtes à l'emploi, avec la garantie que tous les tableaux et analyses proviennent de la même source de données. Le système intègre également des données macroéconomiques et des tableaux de résultats qui servent ensuite de sources de données pour la rédaction de rapports et de notes d'information ou pour une diffusion électronique.

Il est également essentiel de mettre à jour les capacités de l'UBoS en matière d'enquête auprès des ménages. Le programme de stratégie intégrée d'enquête auprès des ménages sur trois ans et la création d'un noyau d'équipes mobiles de terrain constituent les activités principales. Ces équipes auront pour mission de réaliser les enquêtes de l'UBoS et également d'être une source de support technique pour les districts planifiant leurs propres enquêtes. L'UBoS conduira une étude pilote sur une simple enquête de suivi d'indicateurs ; elle pourrait être réalisée par les autorités de district afin de répondre à leurs besoins en matière d'informations.

Ce programme permet également de renouveler la réalisation d'une enquête nationale annuelle sur la prestation de services. L'enquête se sert d'un petit questionnaire auprès d'un vaste échantillon (environ 20 000 ménages) de manière à pouvoir décomposer les résultats au niveau du district. Elle intègre un nombre de caractéristiques du questionnaire unifié des indicateurs de développement, y compris l'emploi d'une lecture optique de marques pour accélérer le processus de saisie des données. Cette enquête sera progressivement généralisée et prise en charge par l'UBoS. Des sessions futures de l'enquête seront renforcées par des entretiens de groupes types.

#### Sources :

- Hauge, Arild. 2001. "Strengthening Capacity for Monitoring and Evaluation in Uganda: A Results Based Management Perspective." ECD Working Paper Series, No. 8. World Bank, Operations Evaluations Department, Washington, D.C.
- Republic of Uganda. 1997. "Poverty Eradication Action Plan: A National Challenge for Uganda." Ministry of Finance, Planning, and Economic Development, Kampala.
- . 1999. "Uganda Poverty Status Report, 1999." Ministry of Finance, Planning and Economic Development, Kampala.
- . 1999. "Five-Year Strategy for Poverty Monitoring and Policy Analysis." Planning and Poverty Eradication Section, Ministry of Finance, Planning, and Economic Development, Kampala
- World Bank. 1999. "Uganda's Integrated Information Management System: A New Approach in Statistical Capacity-Building." Africa Region Findings, No. 142, Washington, D.C.<sup>4</sup>

## Étude de cas C.2 Proposition de plan de suivi de la stratégie de lutte contre la pauvreté en Tanzanie

### C.2.1 Introduction

La SLP tanzanienne s'appuie sur des stratégies précédentes destinées à traiter le problème de la pauvreté et à améliorer le développement humain. Elle renforce des stratégies antérieures à moyen et long termes telles que Vision 2025, la stratégie nationale d'élimination de la pauvreté de 1997 et la stratégie d'aide à la Tanzanie, et définit un plan axé sur trois grands objectifs :

- réduire la pauvreté monétaire ;
- améliorer la qualité de vie et le bien-être social ; et
- parvenir à créer un environnement favorable au développement, et à le préserver.

La préparation de la SLP s'est caractérisée par la forte participation des parties prenantes. Tout au long du processus, via des ateliers réalisés par zones, on a pu recueillir les opinions des acteurs issus du peuple (autorités locales, communautés locales et société civile compris). Projets de cibles, priorités et actions ont également été débattues lors d'un atelier national, auquel participaient des représentants du gouvernement central et régional, des organisations du secteur privé, de la communauté des donateurs et des médias.

## C.2.2 Suivi de la stratégie de lutte contre la pauvreté

Le cadre stratégique de lutte contre la pauvreté (CSLP) propose un plan provisoire pour assurer le suivi et l'évaluation de la stratégie. Ce plan a été perfectionné depuis le lancement de la SLP et continuera à évoluer à mesure que des leçons sont tirées pendant sa mise en œuvre. Cette étude de cas souligne trois aspects de l'expérience tanzanienne : choix des indicateurs, sources de données et cadre institutionnel prévu pour le suivi et l'évaluation de la SLP. D'autres activités, telles que les études participatives, la signalisation des résultats et le travail de participation ne sont pas présentées ici.

### **Choix des indicateurs**

Le système de suivi et d'évaluation (S&E) englobe un ensemble d'indicateurs finaux et intermédiaires. Les indicateurs finaux ont été choisis parmi une liste plus vaste d'indicateurs de pauvreté et de bien-être née d'un processus consultatif ; ils serviront à suivre les progrès vers la satisfaction des principaux objectifs de la stratégie. Les indicateurs intermédiaires serviront à suivre la mise en œuvre de la stratégie sur le plan des ressources allouées et des biens et services générés par le biais des actions clés. Sachant qu'il est difficile de mesurer certains indicateurs finaux à intervalles rapprochés, le système de suivi comprend également un ensemble d'indicateurs de substitution susceptibles d'être suivis chaque année. Par exemple l'un des objectifs de la SLP est de réduire la pauvreté monétaire. Il est donc important de suivre régulièrement la proportion de la population vivant en dessous du seuil de pauvreté. Cependant, dans le cas de la Tanzanie, comme dans bon nombre de pays, puisqu'il est impossible de collecter souvent des données sur les revenus et les dépenses, il a été décidé d'inclure des indicateurs de possession d'actifs ménagers et de matériaux de construction de logements (que l'on peut suivre sur une base annuelle) en tant qu'indicateurs de substitution de la pauvreté monétaire.

Comme le montre le tableau C.3, on peut classer les indicateurs proposés dans quatre catégories globalement alignées avec les objectifs de la SLP : pauvreté monétaire, qualité de vie et bien-être social (santé, éducation, vulnérabilité et bien-être social), stabilité macroéconomique et gouvernance. Selon les catégories, l'adéquation des indicateurs en termes de pertinence, de clarté, de fiabilité, de disponibilité opportune et d'équilibre entre les indicateurs finaux et intermédiaires connaît d'importantes variations.

La SLP choisit un indicateur final approprié spécifique au pays pour la pauvreté monétaire, incidence de la pauvreté monétaire mesurée en fonction du seuil de pauvreté national, et une cible plus ambitieuse que celles des objectifs de développement pour le millénaire : réduire de moitié l'incidence de la pauvreté d'ici 2010 au lieu de 2015. L'incidence de la pauvreté sera décomposée par zones rurales et urbaines. Bien qu'elle reste un phénomène largement rural, la pauvreté monétaire devient de plus en plus un problème urbain. Comme nous l'avons dit, la pauvreté monétaire n'étant pas mesurée chaque année, des indicateurs de substitutions ont été identifiés ; un module d'enquête a été élaboré afin de collecter ces informations selon une périodicité annuelle. Les indicateurs intermédiaires choisis sont pertinents dans le contexte de la Tanzanie et devraient également être accessibles chaque année pour l'étude de la SLP. Le mélange d'indicateurs intermédiaires et finaux est bien équilibré.

Les indicateurs de santé, de survie et de nutrition traduisent l'objectif primordial visant à porter l'espérance de vie à 52 ans d'ici 2010. Il serait toutefois possible de mieux définir certains indicateurs pour qu'ils génèrent davantage d'informations. Ainsi par exemple, les mises en œuvre des programmes de contrôle du paludisme et de prise en charge intégrée des maladies de l'enfance (PCIME) sont mal définies, à moins qu'elles ne se rapportent à une liste d'indicateurs spécifiques contenus dans d'autres documents, tels que les plans de suivi par secteur ou programme. Si tel n'est pas le cas, elles doivent être mieux définies ; on pourrait par exemple reformuler l'indicateur de mise en œuvre du programme de contrôle du paludisme comme la proportion d'installations de soins de santé primaires et secondaires recevant un approvisionnement régulier d'antipaludiques de première et deuxième intentions. De même, le pourcentage d'installations de soins de santé primaires et secondaires disposant d'un personnel formé au PCIME constitue un meilleur indicateur que de savoir si le programme de prise en charge a été mis en œuvre.

**Tableau C.3. Proposition d'indicateurs de suivi du CSLP Tanzanie**

<b>Objectifs</b>	<b>Indicateurs finaux</b>	<b>Indicateurs intermédiaires</b>
<b>1. Réduction de la pauvreté monétaire</b>	<ul style="list-style-type: none"> <li>• Incidence de la pauvreté</li> <li>• Possession d'actifs ménagers (indicateur de substitution)</li> <li>• Type de matériaux de construction des logements ; p.e. planchers, murs et toiture (indicateur de substitution)</li> </ul>	<ul style="list-style-type: none"> <li>• Croissance du PIB réel</li> <li>• Investissement (physique et humain)</li> <li>• Productivité des investissements</li> <li>• Croissance de la valeur ajoutée de l'agriculture</li> <li>• Développement de la stratégie du secteur privé</li> <li>• Production saisonnière d'aliments et de cultures marchandes clés</li> <li>• Kilomètres de routes de campagne réhabilitées</li> <li>• Allocation réelle et budgétisée destinée aux routes de campagne</li> <li>• Allocation réelle et budgétisée destinée à la vulgarisation agricole</li> </ul>
<b>2. Amélioration de la qualité de vie et du bien-être social</b>		
<b>A. Santé, survie et nutrition</b>	<ul style="list-style-type: none"> <li>• Taux de mortalité des enfants de moins de un an et de cinq ans</li> <li>• Pourcentage d'enfants de moins de deux ans vaccinés contre rougeole et DCT</li> <li>• Taux de séropositivité des femmes enceintes</li> <li>• Mortalité maternelle</li> <li>• Espérance de vie</li> <li>• Taux de mortalité des enfants de moins de cinq ans lié au paludisme</li> <li>• Charge de la maladie/morbidité</li> <li>• Proportion de ménages ayant accès à une source sûre d'eau potable</li> <li>• Prévalence du nanisme</li> <li>• Prévalence de maladies cachectisantes</li> </ul>	<ul style="list-style-type: none"> <li>• Proportion de districts disposant de campagnes d'information actives sur le VIH/SIDA</li> <li>• Pourcentage de naissances assistées par un personnel compétent</li> <li>• Pratiques d'allaitement</li> <li>• Mise en œuvre de programmes de contrôle du paludisme</li> <li>• Mise en œuvre de programmes de prise en charge intégrée des maladies de l'enfance</li> <li>• Allocation réelle et budgétisée destinée aux soins de santé primaires</li> <li>• Allocation réelle et budgétisée destinée à la lutte contre le VIH/SIDA</li> <li>• Allocation réelle et budgétisée pour l'eau et l'assainissement</li> </ul>
<b>B. Éducation</b>	<ul style="list-style-type: none"> <li>• Taux d'illettrisme</li> <li>• Égalité des sexes dans l'enseignement primaire et secondaire</li> <li>• Proportion d'enfants d'âge scolaire qui réussissent à achever le cycle primaire</li> <li>• Taux net d'inscription scolaire primaire</li> <li>• Taux brut d'inscription</li> <li>• Taux d'abandons</li> <li>• Taux de passage du primaire au secondaire</li> <li>• Proportion d'élèves de 5<sup>e</sup> passant avec une note donnée à un examen standard</li> </ul>	<ul style="list-style-type: none"> <li>• Allocation réelle et budgétisée destinée à l'éducation de base</li> </ul>
<b>C. Vulnérabilité</b>	<ul style="list-style-type: none"> <li>• Capacité édifée dans toutes les communautés requérant des programmes de filet de sécurité</li> </ul>	<ul style="list-style-type: none"> <li>• Base de données pour les groupes vulnérables identifiés</li> <li>• Encouragement à produire des cultures résistantes à la sécheresse dans toutes les zones menacées de sécheresse</li> <li>• Promotion de programmes d'irrigation gérés par la communauté dans toutes les zones d'irrigation potentielles</li> </ul>

**Tableau C.3. Proposition d'indicateurs de suivi du CSLP Tanzanie (suite)**

<i>Objectifs</i>	<i>Indicateurs finaux</i>	<i>Indicateurs intermédiaires</i>
D. Bien-être social	<ul style="list-style-type: none"> <li>• Pleine réalisation de la SLP</li> </ul>	<ul style="list-style-type: none"> <li>• Pleine réalisation du programme de réforme du gouvernement local</li> <li>• Ratio de jurisprudence par rapport aux cas classés</li> <li>• Délai moyen de résolution des litiges commerciaux</li> <li>• Ratio de sessions effectives de la Cour d'appel par rapport aux sessions prévues</li> <li>• Nombre d'ateliers de SLP, participation et composition des comités</li> <li>• Diffusion des rapports</li> </ul>
<b>3. Création et préservation d'un environnement favorable au développement</b>		
A. Stabilité macroéconomique	<ul style="list-style-type: none"> <li>• Taux d'inflation</li> </ul>	<ul style="list-style-type: none"> <li>• Équilibre budgétaire</li> <li>• Réserves internationales brutes officielles</li> <li>• Taux de change</li> <li>• Solde de compte actuel</li> </ul>
B. Gouvernance	<ul style="list-style-type: none"> <li>• Nombre de votes du budget gérés via des systèmes financiers intégrés</li> <li>• Promesses de dépenses et arriérés enregistrés via les systèmes financiers intégrés</li> <li>• Ampleur de la corruption</li> <li>• Intégrité et transparence du système comptable</li> <li>• Niveau de décentralisation efficace réel du système de gouvernance</li> <li>• Meilleures efficacité professionnelle et rentabilité du système de services publics</li> <li>• Meilleures capacité, motivation et performances du service public</li> <li>• Meilleure gestion du budget au niveau central et également à d'autres niveaux inférieurs</li> </ul>	<ul style="list-style-type: none"> <li>• Application du système financier intégré à tous les ministères et sous-trésoreries</li> <li>• Plans d'action anti-corruption spécifiques élaborés et agréés par les ministères de l'agriculture et des coopératives, de l'éducation et de la culture ; de la santé ; et de l'eau ; et centre pour l'étude de la démocratie basé sur la stratégie nationale de lutte contre la corruption</li> <li>• Modules d'amélioration des performances élaborés et agréés pour les secteurs prioritaires</li> <li>• Préparation opportune de budgets à tous les niveaux</li> <li>• Pluralisme institutionnel dans la prestation de services publics</li> </ul>

Pour certains des indicateurs finaux et intermédiaires, l'opportunité peut poser problème (par exemple, pratiques en matière d'allaitement), étant donné que la principale source de données pour ces indicateurs sera l'enquête démographique et de santé qui devrait être réalisée durant la mise en œuvre de la SLP.

Les indicateurs et cibles d'éducation conviennent à l'objectif de suppression de l'illettrisme d'ici 2010 : porter le taux brut d'inscription à 85 %, le taux de passage du primaire au secondaire de 15 à 21 %, le taux d'abandon au primaire de 6,6 à 3 % et le taux net d'inscription primaire de 57 à 70 %. La plupart des indicateurs finaux sont clairement définis, à l'exception de l'égalité des sexes qui pourra être mesurée par rapport aux taux bruts d'inscription, taux nets d'inscription ou à tout autre indicateur. La liste des indicateurs intermédiaires semble incomplète : des indicateurs de résultats provenant des dépenses publiques, tels que le ratio élèves-enseignants, la disponibilité de manuels, le pourcentage de classes réhabilitées et le temps de trajet moyen jusqu'à l'école, pourraient apporter des informations utiles à la compréhension des tendances révélées par les indicateurs finaux. Les indicateurs d'éducation sont généralement disponibles à intervalles rapprochés puisqu'ils sont obtenus des systèmes routiniers de collecte de données mis en place par le ministère de l'Éducation. Les données relatives aux inscriptions seront validées avec les informations provenant du recensement 2002.

Les indicateurs de vulnérabilité sont moins bien élaborés. S'ils reflètent bien les actions qui seront appliquées en la matière, ils s'intéressent au suivi des activités et non aux résultats ; ils ne comprennent aucun « indicateur final ». Il faudrait également élaborer des indicateurs intermédiaires plus spécifiques. Par exemple, le pourcentage de fermiers de zones menacées de sécheresse ayant adopté la culture de produits résistant à la sécheresse peut se révéler un meilleur indicateur que de savoir si la production de cultures résistant à la sécheresse a été encouragée. De même, mesurer l'utilisation de la base de données

sur les groupes vulnérables pourrait être un meilleur indicateur que de savoir si un tel type de base de données a été élaboré.

Les indicateurs de bien-être social reflètent également la difficulté de préciser des indicateurs mesurables. Des sujets très divers sont traités sous ce titre. Ces indicateurs essaient de saisir l'évolution du transfert des responsabilités pour les services clés aux autorités locales ; de l'accès à la justice, de l'efficacité et de la transparence du système administratif ; et du niveau de participation de tous les acteurs au processus d'élaboration d'une SLP, mais ils ne donnent qu'une vision très partielle des progrès réalisés. C'est un domaine dans lequel les objectifs, les indicateurs et les cibles devront être développés davantage.

Les indicateurs macroéconomiques et de gouvernance ont pour objet de mesurer le degré de réussite dans la création d'un environnement favorable au développement. Notamment, sur le plan macroéconomique, la SLP cherche à atteindre un taux d'inflation plus ou moins cohérent avec les prévisions d'inflation des principaux partenaires commerciaux de la Tanzanie. Cet objectif complète celui qui consiste à atteindre une croissance annuelle de 6% du produit intérieur brut (PIB) sur les trois prochaines années ; cela constituerait le socle qui permettrait d'atteindre les objectifs de réduction de la pauvreté à moyen et long termes. Les indicateurs intermédiaires et finaux proposés sur le plan macroéconomique apporteront des informations pertinentes et fréquentes pour le suivi de l'objectif de stabilité.

Côté gouvernance, les principaux objectifs cherchent à améliorer les performances du secteur public, dont la prestation de services publics ; à réduire les fuites de ressources et à promouvoir la responsabilisation. Les indicateurs proposés pour suivre ces objectifs ne sont généralement pas au point. À la différence d'autres domaines, les points cités dans la catégorie « indicateurs finaux » ne sont pas tous des indicateurs ; par exemple, « un système de gouvernance décentralisé efficace » constitue un objectif et non un indicateur de décentralisation. Plusieurs autres points ne constituent pas des indicateurs mesurables. Par exemple, « ampleur de la corruption » n'identifie pas ce qu'il faut faire pour mesurer la corruption. Il est nécessaire d'identifier des indicateurs appropriés à la Tanzanie. Suivre les questions d'égalité entre les hommes et les femmes fait partie intégrante du système de suivi de la SLP. Les indicateurs de santé et d'éducation tels que la mortalité des enfants de moins de un an et de moins de cinq ans, les taux de vaccination, d'inscription et de passage du primaire au secondaire seront décomposés par sexes. En outre, le système de suivi intègre des indicateurs spécifiques aux hommes et aux femmes tels que le taux de séropositivité chez les femmes enceintes, la mortalité maternelle et le pourcentage de naissances assistées par du personnel compétent.

Les indicateurs de santé et d'éducation seront également décomposés par zones urbaines et rurales et par régions administratives. Cela revêt une importance cruciale vu les fortes variations des conditions sociales entre les différentes régions du pays. Par exemple, la mortalité des enfants de moins de un an et de moins de cinq ans est trois fois supérieure dans la région la plus pauvre.

L'un des défis majeurs consiste à choisir un nombre gérable d'indicateurs capables de donner des informations pertinentes et suffisantes pour évaluer l'évolution de la SLP. Dans le cas de la Tanzanie, cela a été un processus itératif. Le système de suivi a débuté avec un total de quelque 111 indicateurs au niveau national ; on en compte aujourd'hui environ 70. Le perfectionnement de la liste des indicateurs se poursuivra à mesure que les représentants du gouvernement et leurs homologues sauront identifier les indicateurs les plus utiles et ceux qui manquent à la liste.

**Tableau C.4. Sources d'informations pour le suivi, Tanzanie**

<i>Type d'indicateur</i>	<i>Source de référence</i>	<i>Fréquence du suivi et source de données</i>
Pauvreté en nombre d'habitants	Estimations préliminaires : Enquête 1991-1992 sur le budget des ménages (EBM) Mise à jour : EBM 2000-2001	Aucune EBM supplémentaire n'a été prévue au cours de la période de mise en œuvre du CSLP
Indicateurs de substitution pour la pauvreté monétaire	EBM 2000-2001 Recensement 2002	Les enquêtes annuelles de suivi de la pauvreté mesureront les indicateurs de substitution pour le revenu
Indicateurs macroéconomiques	Comptabilité nationale et enquête économique réalisée par le bureau national des statistiques et la Commission de planification	Mises à jour annuelles auprès des mêmes sources

**Tableau C.4. Sources d'informations pour le suivi, Tanzanie (suite)**

<i>Type d'indicateur</i>	<i>Source de référence</i>	<i>Fréquence du suivi et source de données</i>
Infrastructure rurale	Rapports sur le secteur routier préparés par le ministère des travaux et le ministère du gouvernement régional et local	Même source, fréquence non précisée
Santé	Enquête 1999 sur la santé génésique et juvénile en Tanzanie (EDS intermédiaire)	Recensement 2002 EDS (à maintenir durant la mise en œuvre de ce CSLP) Système d'information sanitaire (pour des mises à jour annuelles de la couverture vaccinale)
Proportion de districts disposant de campagnes d'information actives sur le VIH/SIDA	Programme national de contrôle du SIDA	Même référence, fréquence non précisée
État nutritionnel des enfants	Enquête 1999 sur la santé génésique et juvénile en Tanzanie (nouvelles estimations à partir de l'EDS suivant)	Suivi communautaire et suivi systématique dans les dispensaires Les enquêtes annuelles de suivi de la pauvreté peuvent également intégrer un module anthropométrique
Indicateurs d'éducation	Système de collecte de données systématique du ministère de l'Éducation Cartographie des écoles	Suivi annuel utilisant des données administratives. Le recensement 2002 permettra de vérifier les données administratives relatives aux inscriptions
Allocation des ressources	RDP, CDMT et processus de budgétisation annuels	Même source ; revues trimestrielles

### **Sources de données**

Le calcul de références fiables pour les indicateurs retenus était particulièrement difficile dans certains cas du fait de l'absence de données récentes. Les données les plus récentes sur la consommation destinées à évaluer l'incidence de la pauvreté proviennent de l'enquête sur le budget des ménages de 1991-1992 (EBM). L'estimation de référence pour 2000 et les cibles provisoires ont été déterminées en extrapolant les résultats de l'enquête de 1991-1992 d'après les estimations de la population tirées des listes réalisées pour l'EBM 2000-2001, mais des problèmes de méthodologie sont associés à ces estimations qui seront reprises compte tenu des résultats de l'EBM 2000-2001.

Les autres sources importantes de données de référence sont l'enquête de 1999 sur la santé génésique et juvénile en Tanzanie pour les indicateurs de santé et de nutrition ; les données administratives pour les indicateurs d'éducation ; et la comptabilité nationale et l'enquête économique pour les indicateurs macroéconomiques (voir tableau C.4). Les données provenant du recensement 2002 et de l'EBM 2000-2001 aideront à valider la fiabilité des données administratives concernant les inscriptions scolaires et à mettre à jour les chiffres de la mortalité à partir de l'enquête sur la santé génésique et juvénile de 1999. Cette vérification est importante, car l'enquête sur la santé génésique et juvénile constitue un EDS intermédiaire sur un échantillon relativement plus réduit que la véritable enquête. La plupart des indicateurs seront suivis selon une périodicité annuelle, à l'exception de la pauvreté en nombre d'habitants et de certains indicateurs de santé et de nutrition. Comme nous l'avons déjà dit, la pauvreté en nombre d'habitants sera remplacée par un ensemble d'indicateurs de revenu de substitution dont les données de référence seront calculées d'après l'EBM 2000-2001 et le recensement 2002, et qui pourra être suivie par le biais d'une enquête annuelle de suivi de la pauvreté. Il a été décidé de ne pas réaliser chaque année une enquête spécifique de suivi de la pauvreté, mais plutôt d'élaborer un module spécial contenant des indicateurs du revenu de substitution, et de l'intégrer à toute enquête auprès des ménages prévue une année donnée. Les indicateurs de santé et de nutrition seront suivis au moins une fois durant les trois ans en fonction du moment où sont communiqués les résultats de l'EDS suivant.

### **Accords institutionnels prévus pour le suivi et l'évaluation**

Le cadre institutionnel proposé pour suivre la SLP résulte de vastes consultations entre les différents acteurs. Tout d'abord, les différentes propositions ont été présentées lors d'une réunion des acteurs à laquelle participaient des représentants du gouvernement, de la société civile, des ONG, du secteur privé, des universités et des centres de recherche. Ces propositions ont été débattues lors d'une réunion suivante à laquelle participaient des représentants d'organisations multilatérales et bilatérales, du ministère des finances, de la commission de planification et du bureau national des statistiques (BNS). La réunion a été organisée par le bureau du vice-président (BVP) en sa qualité de coordinateur de la préparation du CSLP. Un plan directeur du suivi de la pauvreté a été agréé en novembre 2001.

Le point fort du système de suivi et d'évaluation est le comité consultatif national de suivi de la pauvreté (NPMSC, National Poverty Monitoring Steering Committee). Son rôle consiste à apporter des conseils en général sur le suivi de la SLP et à veiller à ce que le feedback du système de suivi soit intégré au processus décisionnel national. Le comité est formé de représentants du gouvernement, du secteur privé, d'ONG et de la société civile. La division « élimination de la pauvreté » du bureau du vice-président remplit le rôle de secrétariat.

Comme l'illustre la figure C.1, le comité consultatif est aidé dans sa tâche par quatre groupes de travail coordonnés par la division « élimination de la pauvreté ».

Le groupe de travail chargé des enquêtes et des recensements, sous la direction du bureau national des statistiques, réalise de vastes enquêtes auprès des ménages ainsi que le recensement, et coordonne également les activités de stockage des données par le biais de l'initiative de base de données socioéconomique.

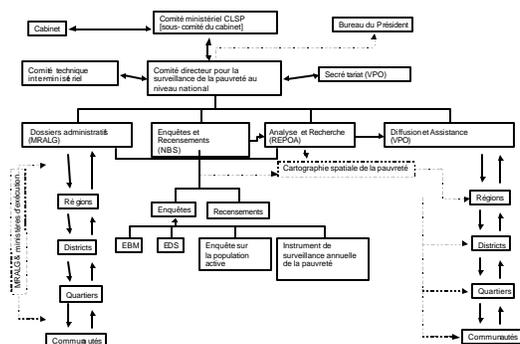
Le groupe de travail chargé des données systématiques, sous la direction du ministère de l'administration régionale et du gouvernement local (MRALG, Ministry of Regional Administration and Local Government), coordonne et gère la collecte des données sectorielles auprès des ministères d'exécution ainsi que les données collectées via les systèmes administratifs des unités gouvernementales décentralisées.

Le groupe de travail chargé des recherches et analyses, sous la direction de la commission de planification du président et du groupe de recherches sur l'allègement de la pauvreté (REPOA, Research on Poverty Alleviation), coordonne les études et les initiatives spéciales telles que la cartographie spatiale de la pauvreté.

Le groupe de diffusion, de sensibilisation et d'assistance, sous la direction du bureau du vice-président, est chargé de coordonner les activités de diffusion à tous les niveaux et de veiller à ce que les opinions des gouvernements locaux soient reflétées dans le système de suivi.

Le comité consultatif national sur le suivi de la pauvreté devrait former un lien essentiel entre les décideurs et le système de suivi, travaillant en liaison avec le comité ministériel pour le CSLP via le bureau du vice-président.

**Figure C.1. Cadre institutionnel pour le suivi du CSLP, Tanzanie**



Ce comité, auquel participent plusieurs ministres ainsi que le gouverneur de la banque de Tanzanie, a été formé dans le but de guider le processus d'élaboration et de mise en œuvre d'un CSLP. Il reçoit l'appui du comité technique interministériel, coordonné par le ministère des finances et formé de représentants du bureau du vice-président, du premier ministre, de la commission de planification, de la banque de Tanzanie et de plusieurs autres ministères d'exécution.

Le cadre institutionnel proposé pour suivre et évaluer la SLP pourrait constituer un lien adéquat entre les producteurs et les utilisateurs de données. Il contient cependant de nombreux défis. Premièrement, le MRALG et le bureau national des statistiques doivent posséder de solides capacités institutionnelles pour réussir dans leur rôle de coordination. Deuxièmement, le rôle du ministère des finances (Mdf) en tant que coordinateur du suivi des dépenses publiques n'est pas clairement défini dans le cadre actuel. Il serait important de créer des mécanismes de coordination entre le Mdf et le MRALG qui est chargé de coordonner toutes les données administratives du suivi. Troisièmement, la mise en œuvre complète de la réforme du gouvernement en cours est nécessaire pour qu'un flux adéquat de données administratives proviennent de différents niveaux du gouvernement. Cette réforme devrait clarifier la répartition des responsabilités de gestion des systèmes d'information entre le MRALG et les ministères d'exécution comme celui de l'éducation et de la santé, répartition assez floue à l'heure actuelle. Si l'on repoussait encore la mise en œuvre de la réforme, certaines activités seraient redondantes ou bien des informations manqueraient. Enfin, le cadre institutionnel proposé présente une structure de suivi et d'évaluation qui, bien qu'assez claire au niveau national, l'est moins à l'échelle de la région et du district. Dans la mesure où les efforts de décentralisation délèguent le pouvoir décisionnel aux gouvernements locaux, il est essentiel que soit mise en place une structure locale de suivi et d'évaluation. Dans l'ensemble, il est important que les incitations à collecter, analyser et synthétiser des informations exactes et opportunes soient homogènes à tous les niveaux du système de suivi.

Sources :

United Republic of Tanzania, The. 2000. "Poverty Reduction Strategy Paper." Dar Es Salaam  
 \_\_\_\_\_. 2001. "Poverty Monitoring Master Plan." Dar Es Salaam

### Étude de cas C.3 Enquête de feedback des citoyens comme outil de participation de la société civile à l'évaluation des performances du secteur public : le cas de Bangalore en Inde

À Bangalore, en Inde, une ONG a réalisé des enquêtes de feedback auprès des citoyens axées sur les services fournis par la municipalité, tels que l'eau, l'électricité, le ramassage des déchets et les hôpitaux. Les citoyens ont été interrogés sur leur degré de satisfaction vis-à-vis de ces services publics, en indiquant les aspects les plus ou les moins satisfaisants, sur l'efficacité des fonctionnaires et sur le versement de pots-de-vin aux fonctionnaires pour l'obtention de ces services.

Les objectifs de l'enquête étaient :

- de générer un feedback des citoyens sur les services publics et d'attribuer à chaque service municipal une note globale sur ses performances ;

- d'identifier les services spécifiques où la prestation est bonne ou mauvaise ;
- d'identifier l'étendue et l'importance de la corruption ;
- de pousser les citoyens à plus de proactivité ;
- d'apporter un outil de diagnostic aux départements de la municipalité de sorte que leurs dirigeants soient mieux à même de comprendre les performances de leurs services et d'identifier ce qu'il est possible de faire pour les améliorer ; et
- d'encourager les organes publics à s'intéresser davantage aux clients et à être plus transparents.

En fonction des résultats des enquêtes de Bangalore, les organes municipaux ont tous été classés selon le niveau de satisfaction ressenti par les citoyens pour leurs services. Hôpitaux et banques étaient en tête alors que le service de l'aménagement urbain (qui connaissait les plus hauts niveaux de corruption avérée) était le plus mal noté.

Les résultats des enquêtes ont été largement diffusés et ont bénéficié d'une forte couverture médiatique. Des ateliers ont été mis sur pied pour communiquer les résultats aux groupes de citoyens et à d'autres ONG. Même si ces résultats ne les étonnaient pas, ils leur ont fourni des preuves solides et ont permis de souligner des points problématiques spécifiques. Ils ont également stimulé la participation des citoyens et la formation de groupes de résidents.

L'ONG qui a réalisé ces enquêtes a transmis des rapports détaillés aux responsables de tous les services publics. La plupart des responsables des organes et des hauts dirigeants s'en sont méfiés, mais certains ont bien réagi, comme le directeur du service d'aménagement urbain qui, suite à cela, a testé une approche de partenariat avec les groupes de citoyens et les ONG. Cette initiative a abouti à des innovations dans la prestation de services et à la création d'un nouveau système permettant aux clients d'exprimer leurs critiques. Avec l'aide de l'ONG, des programmes de formation destinés aux fonctionnaires ainsi qu'un groupe partenaire ayant pour objet de diffuser les informations et de servir d'organe de surveillance ont été mis sur pied.

Des enquêtes similaires ont été réalisées depuis dans d'autres villes d'Inde, dont Madras, Mumbai, Calcutta et Pune. On a ainsi pu publier des comparaisons pour plusieurs villes.

Source : Adapté de MacKay, Keith et Sulley Gariba (eds.). 2000. *The Role of Civil Society in Assessing Public Sector Performance in Ghana: Proceedings of a Workshop*. Evaluation Capacity Development, Operations Evaluation Department, World Bank, Washington, D.C.

## **Étude de cas C.4 Évaluation des bénéficiaires pour les pauvres des programmes d'assistance-travail : le programme Trabajar en Argentine**

### **C.4.1 Introduction**

Le programme argentin Trabajar vise à réduire la pauvreté en créant des opportunités d'emploi pour les personnes démunies et en améliorant simultanément l'infrastructure sociale des communautés pauvres. Les rémunérations proposées par le programme sont relativement faibles afin de n'attirer que des chômeurs pauvres : c'est une forme « d'auto-sélection ». (Pour plus d'information sur ce programme et d'autres programmes publics de grands travaux, voir chapitre 18, « Protection sociale ».) Les administrations locales et les ONG proposent les projets d'infrastructure sur lesquels travaillent les participants et prennent en charge les coûts non salariaux du projet. Les projets sont approuvés au niveau régional selon les directives du gouvernement central.

Le programme a évolué dans le temps en intégrant divers enseignements. Trabajar I, un programme pilote, a été lancé en 1996 en réponse à une crise économique et à des taux de chômage supérieurs à 17 %. Trabajar II a débuté en 1997 : c'était une version étendue et amendée du programme pilote ; puis cela a été le tour de Trabajar III en 1998. Trabajar II comprenait plusieurs réformes destinées à améliorer le ciblage du projet : dans ses critères d'allocation budgétaire, le gouvernement central a accordé plus de poids aux indicateurs de pauvreté provinciale et de chômage et aux projets proposés par les régions pauvres ; au niveau local, des efforts ont été réalisés en vue de renforcer la capacité des agences provinciales de manière à aider les régions pauvres à préparer des projets et à accroître la qualité des infrastructures construites dans le cadre de ce programme.

## C.4.2 Style d'évaluation

L'effort d'évaluation a débuté durant la préparation de Trabajar II. L'évaluation avait pour but de vérifier si le programme atteignait ses objectifs et d'identifier les domaines dans lesquels il serait possible d'accroître son efficacité par des réformes. Elle regroupait plusieurs composantes isolées destinées à établir : a) l'augmentation des gains de salaire net revenant aux participants au programme ; b) l'allocation des ressources du programme entre les régions (ciblage) ; c) la qualité des projets d'infrastructure financés. En outre, l'étude s'est intéressée au rôle de la communauté et des ONG dans les résultats du projet.

Deux des composantes de l'évaluation de Trabajar illustrent les meilleures techniques. La première, l'évaluation des gains de salaire net, démontre un progrès par rapport aux évaluations classiques des programmes d'assistance-travail (qui mesurent en général les gains de salaire des participants simplement en termes de salaire *brut*), en évaluant justement les gains de salaire *net*. En se fondant sur les données tirées d'enquêtes existantes et nouvelles réalisées auprès des ménages et en utilisant les progrès récents des techniques de comparaisons correspondantes, l'étude a pris en compte le salaire sacrifié (salaire auquel les participants renoncent en rejoignant le programme Trabajar), et a abouti à une estimation plus précise, plus faible, des gains de salaire net pour les participants.

La deuxième composante, l'étude des résultats du ciblage, a introduit une nouvelle technique d'évaluation du ciblage (l'allocation des financements du programme) quand l'incidence des dépenses publiques au niveau local n'est pas observée.

Le style d'évaluation global a également associé les meilleures composantes et techniques de recherche (depuis l'analyse quantitative jusqu'aux visites des sites industriels en passant par l'évaluation sociale), procurant par là un flux intégré de résultats dans les temps requis.

## C.4.3 Techniques de collecte et d'analyse des données

**L'évaluation des gains de salaire net** qui reviennent aux participants au programme s'est appuyée sur deux sources de données, une enquête nationale sur le niveau de vie (Encuesta de Desarrollo Social, ou EDSO) et une enquête auprès des participants à Trabajar réalisée spécialement dans le but de l'évaluation.<sup>5</sup> Ces enquêtes ont été effectuées en août et septembre 1997 par le bureau national des statistiques, par les mêmes équipes d'entretien qui utilisaient le même questionnaire. L'échantillon de l'enquête EDSO a couvert 85 % de la population nationale, abstraction faite de quelques zones rurales et communautés très petites. L'échantillon de l'enquête sur les participants à Trabajar a été tiré d'un échantillon aléatoire de projets de Trabajar II intégrés à l'échantillon de l'EDSO ; il a généré des données pour 2 802 participants actuels au programme (entre mai 1997 et janvier 1998, le nombre total de participants à Trabajar II était de 65 321).

Afin d'établir un groupe témoin correspondant à partir de l'enquête EDSO, l'étude s'est servie d'une technique dite de notation à la propension. Pour une correspondance idéale, il faudrait deux individus, le premier dans l'échantillon des participants et le deuxième dans le groupe témoin, pour lesquels toutes les variables (x) prédisant la participation au programme seraient identiques. Le problème que l'on rencontre fréquemment ici, est que cette méthode est peu pratique au regard du grand nombre de variables contenues dans (x). Cependant, il est possible de calculer les correspondances d'après la notation à la propension de chaque individu, qui correspondent simplement à la probabilité de participation sous condition de (x).<sup>6</sup> On a pu estimer le salaire sacrifié des participants à Trabajar II à partir des données relatives aux salaires du groupe témoin correspondant composé de non-participants. Le salaire net provenant de la participation au programme a ensuite été calculé comme la somme des salaires du programme moins le salaire sacrifié.

**L'analyse de ciblage** est intéressante en ce sens qu'elle n'a exigé aucune collecte de données spéciale. Elle s'est inspirée des données provenant du bureau des projets du ministère sur les allocations de financements par département géographique. Elle s'est également inspirée de l'indice de pauvreté de chaque région, calculé à partir du recensement de 1991 comme la proportion de ménages ayant des « besoins élémentaires non satisfaits ».<sup>7</sup> Pour analyser l'incidence du ciblage, les données sur les dépenses publiques par secteur géographique, ici par département, ont été régressées en se fondant sur les taux de pauvreté géographiques correspondants. Le coefficient obtenu a constamment estimé un « différentiel de ciblage » induit par la différence d'allocation moyenne du programme entre les pauvres et les non-pauvres. Le différentiel de ciblage national pouvait alors être décomposé en composantes dues au mécanisme de ciblage du gouvernement central (allocations de financements entre les départements) et au ciblage réalisé au niveau provincial.

**L'analyse de la qualité de l'infrastructure** étudie en deux temps le rapport coûts-bénéfices des projets d'infrastructure de Trabajar. Dans un premier temps, une note de qualité globale a été décernée à un échantillon de 50 projets achevés de Trabajar I d'après des indicateurs classés dans six catégories : technique, institutionnelle, environnementale, socioéconomique, supervision et exploitation et maintenance, et des analyses coûts-bénéfices ont été réalisées à chaque fois que c'était utile (mais pas pour les écoles ou les dispensaires). Dans un deuxième temps, une étude de suivi portant sur 120 projets de Trabajar II a été réalisée un an plus tard, pour suivre l'impact des réformes sur la qualité des infrastructures.

**Des évaluations sociales** ont été réalisées pendant la préparation du projet pour Trabajar I comme Trabajar II. Elles ont apporté un feedback sur les questions de mise en œuvre du projet telles que le rôle des ONG, la disponibilité de l'assistance technique dans la préparation et la construction du projet et le choix des bénéficiaires. Elles ont été faites par des sociologues au moyen de groupes types et d'entretiens.

#### C.4.4 Résultats

**Impact du programme.** Il est important de tenir compte du salaire sacrifié de manière à obtenir une image précise des avantages du programme d'assistance-travail. Les participants au programme ne pouvaient se permettre d'être au chômage en l'absence de programme ; c'est pourquoi une certaine partie du salaire est sacrifiée pour participer au programme. On estime le revenu sacrifié en observant les revenus des non-participants « correspondants » à ceux des participants au programme. Les estimations par la méthode de correspondance montrent qu'ignorer le salaire sacrifié exagère substantiellement les gains moyens obtenus grâce au programme, que l'on estimait à environ la moitié des salaires reçus via le programme. L'évaluation a également révélé que la répartition des gains était résolument favorable aux pauvres, avec 80 % des participants au programme appartenant aux 20 % les plus pauvres de la population, même après avoir réduit les gains salariaux du montant du salaire sacrifié. La participation des femmes est faible dans le programme (15 %), mais les gains en salaire net sont pratiquement identiques pour les hommes comme pour les femmes qui participent à Trabajar.

**Ciblage des performances.** Les performances se sont nettement améliorées du fait des réformes de Trabajar II. On a vu une multiplication par sept de l'allocation implicite des ressources aux ménages pauvres entre Trabajar I et Trabajar II. On doit cette hausse à un meilleur ciblage au niveau central pour un tiers et au niveau provincial pour deux tiers. On a en revanche noté des différences significatives entre les provinces dans les résultats du ciblage. Un département dont 40 % de la population est considérée comme pauvre peut s'attendre à recevoir, en fonction de la province dont il dépend, de zéro à cinq fois l'allocation départementale moyenne. De plus, le ciblage semblait être moins performant dans les provinces les plus pauvres.

**Qualité du projet d'infrastructure.** La qualité s'est révélée correcte, mais les réformes de Trabajar II n'ont malheureusement pas généré d'amélioration notable. C'était en partie due à la très forte expansion du programme, qui a accru la difficulté de respecter certaines des normes opérationnelles spécifiées auparavant. Toutefois, les projets d'infrastructure de Trabajar II ont mieux réussi à satisfaire les besoins prioritaires de la communauté. L'évaluation sociale a révélé la nécessité d'obtenir de la part des ONG et des municipalités rurales une meilleure assistance technique, ainsi qu'une plus grande diffusion publique et transparence des informations concernant le programme Trabajar.

#### C.4.5 Implications sur la politique générale

Les résultats de l'évaluation ont démontré clairement que les participants au programme Trabajar étaient en majorité des pauvres. L'auto-sélection obtenue en proposant des salaires faibles est une stratégie qui fait ses preuves en Argentine, et les participants voient leurs revenus augmenter du fait de leur participation (même si les gains nets sont inférieurs au salaire brut, du fait du salaire sacrifié). Le programme ne semble pas faire de discrimination entre les hommes et les femmes. Les réformes de Trabajar II ont réussi à améliorer les résultats du ciblage géographique ; le programme arrive mieux aujourd'hui à diriger les fonds vers les régions pauvres. Mais les performances varient et sont immanquablement faibles dans quelques provinces qui demandent plus d'attention. Enfin, les résultats décevants de la qualité du projet d'infrastructure ont poussé l'équipe projet à perfectionner les procédures opérationnelles : insister davantage sur les visites des sites pour l'évaluation et la supervision, pénaliser les organismes peu performants lors de l'achèvement du projet et approfondir le manuel d'évaluation.

## C.4.6 Coûts d'évaluation et administration

**Coûts.** On estime à quelque 350 000 dollars le coût de réalisation de l'enquête Trabajar pour l'étude des gains de salaire net et le traitement des données. Les deux évaluations coûts-bénéfices de la qualité des sous-projets ont chacune coûté environ 10 000 dollars, de même que les évaluations sociales, ce qui porte les dépenses totales de l'évaluation à environ 390 000 dollars.

**Administration.** L'évaluation a été mise en œuvre conjointement par les services de la Banque mondiale et l'équipe de projet argentine. Tout au long des différentes étapes, l'effort d'évaluation a requis une coordination avec plusieurs agences locales du gouvernement, y compris l'agence statistique, le ministère du travail (bureaux locaux compris), et la division d'analyse politique du ministère du développement social.

## C.4.7 Enseignements tirés et travaux futurs

**Importance de tenir compte du salaire sacrifié dans l'évaluation des gains apportés par le programme d'assistance-travail.** Le salaire sacrifié représente environ la moitié des salaires bruts touchés par les participants au programme d'assistance-travail en Argentine. Ce résultat montre que les méthodes d'évaluation classiques (à l'aide des seuls salaires bruts) exagèrent largement les gains de salaire, et donc également la situation que connaîtraient les pauvres si ce programme n'existait pas.

**Utilité des méthodes de correspondance des notations à la propension.** Les notations à la propension permettent de déduire des correspondances fiables entre les échantillons de participants et de non-participants (groupe témoin).

**Emploi judicieux des sources de données nationales existantes.** Les sources de données disponibles, telles que le recensement national ou les enquêtes auprès des ménages, apportent souvent des données précieuses pour les activités d'évaluation. Travailler à partir de sources existantes réduit la nécessité de pratiquer des collectes de données coûteuses dans un but spécial. Des techniques d'évaluation novatrices peuvent compenser les données manquantes, comme l'a très justement démontré l'évaluation des résultats du ciblage géographique de Trabajar.

**Large éventail de composantes d'évaluation.** Le style d'évaluation de Trabajar illustre un mélange efficace d'outils et de techniques d'évaluation. Analyse des données des enquêtes, visites de sites et évaluations sociales ont toutes été utilisées pour générer un large éventail de résultats qui ont offert une vision précieuse de l'efficacité du projet et souligné les domaines à réformer.

**Opportunité des résultats.** Bon nombre des composantes de l'évaluation ont été clairement conçues dans l'optique du cycle du projet et réglées pour générer des résultats durant les étapes de préparation du projet de sorte que ces résultats servent aux conceptions suivantes. Plusieurs composantes fournissent à présent des données régulières dans un processus continu de suivi du projet.

**Travaux futurs.** Trois études sont prévues : la technique de comparaison correspondante sera à nouveau appliquée pour évaluer l'impact de la participation au programme sur l'activité du marché du travail ; la qualité des projets d'infrastructure sera réévaluée ; et une composante de recherche qualitative se penchera sur les opérations et les procédures du programme.

*Sources et autres lectures :*

Jalan, Jyotsna et Martin Ravallion. 1999. "Income Gains to the Poor from Workfare: Estimates for Argentina's Trabajar Program." Policy Research Working Paper 2149. World Bank, Development Economics Research Group, Washington, D.C.

Ravallion, Martin. 1999. "Monitoring Targeting Performance when Decentralized Allocations to the Poor Are Unobserved." Policy Research Working Paper 2080. World Bank, Development Economics Research Group, Washington, D.C.

## Étude de cas C.5 Évaluation du projet de vulgarisation agricole du Kenya

### C.5.1 Introduction

**Description du projet.** Le premier projet national de vulgarisation (NEP I) du Kenya a introduit en 1983 le système de gestion « Formation et visite » (F&V) dans les services de vulgarisation agricole. Le projet

avait pour double objectif de développer les institutions et de proposer des services de vulgarisation aux fermiers, dans le but d'accroître la productivité agricole. NEP II, qui a suivi en 1991, avait pour but de renforcer les bénéfices réalisés grâce à NEP I en augmentant les contacts directs avec les fermiers, en améliorant la pertinence des informations et des technologies liées à la vulgarisation, en développant les compétences du personnel et des fermiers et en renforçant le développement institutionnel.

**Évaluation de l'impact.** Les performances controversées du système de vulgarisation kenyan font partie d'un débat plus vaste sur la rentabilité de l'approche F&V appliquée à la vulgarisation. En dépit de l'intensité de ce débat et de la très grande quantité de sommes investies, les tentatives de mesure de l'impact de l'approche F&V selon des méthodes rigoureuses ont été rares. Dans le cas du Kenya, les estimations très élevées de retours sur investissement des F&V selon une étude antérieure et le manque de résultats visibles et convaincants (notamment les mauvaises performances de l'agriculture kenyane ces dernières années) ont alimenté un débat particulièrement vif.

Le département d'évaluation rétrospective des opérations (OED) de la Banque mondiale a réalisé une évaluation du système de vulgarisation kenyan entre 1997 et 1999. Dans un cadre de gestion basé sur les résultats, l'évaluation s'est intéressée à l'impact des services de projets sur la productivité et l'efficacité des exploitations. Elle a également développé des mesures de variables de sortie du programme (par exemple, fréquence et qualité des contacts) et des résultats (autrement dit, la connaissance des fermiers sur le plan des nouvelles techniques et leur aptitude à les adopter).

### C.5.2 Style d'évaluation

La stratégie d'évaluation illustre les meilleures techniques en utilisant un large éventail de méthodes d'évaluation et en exploitant les données disponibles.<sup>8</sup> Elle s'est appuyée sur des méthodes quantitatives et qualitatives et a ainsi pu compléter les conclusions sur l'impact du programme, tirées selon des méthodes empiriques rigoureuses, par des évaluations des bénéficiaires et des entretiens avec le personnel qui mettent en exergue les questions pratiques liées à la mise en œuvre. L'étude a également appliqué la méthode d'évaluation conditionnelle afin de déterminer si les fermiers étaient prêts à payer pour bénéficier de services de vulgarisation.<sup>9</sup> La mise en œuvre nationale du système de F&V a compliqué l'évaluation quantitative, empêchant de ce fait toute comparaison entre un groupe «avec programme» et un groupe «sans programme» (groupe témoin). La méthodologie d'évaluation a donc cherché à exploiter les données sur la production agricole des ménages disponibles avant le projet afin de réaliser, à l'aide de méthodes de données de panel, des comparaisons entre la situation antérieure et postérieure au projet. Pour cela, les données des ménages disponibles ont été complétées par les données issues d'une enquête récente de manière à former un panel. Vu qu'il n'a pas été possible de mener à bien les évaluations des bénéficiaires conçues pour cette étude, l'évaluation s'est inspirée des résultats pertinents provenant de deux évaluations de bénéficiaires effectuées récemment au Kenya.

### C.5.3 Techniques de collecte et d'analyse des données

La méthode d'évaluation s'est basée sur plusieurs sources de données qualitatives et quantitatives existantes. L'évaluation quantitative s'est basée en très grande partie sur une enquête de 1998 réalisée auprès des ménages par l'OED. Les données de panel de cette enquête ont été obtenues en retrouvant, et en revoyant, le plus grand nombre de ménages possible parmi ceux qui avaient été contactés lors d'une enquête réalisée par le département technique pour l'Afrique de la Banque mondiale, laquelle s'était inspirée d'un sous-échantillon de l'enquête de 1982 sur le budget des ménages ruraux.<sup>10</sup>

L'étude a évalué le développement institutionnel en s'appuyant sur une enquête concernant le personnel chargé de la vulgarisation, plusieurs études sur le service de vulgarisation menées ou sollicitées récemment par le ministère de l'Agriculture, ainsi que sur des entretiens individuels et en groupes types avec les prestataires de services de vulgarisation.

L'étude s'est également fondée sur deux évaluations de bénéficiaires récentes, une étude de 1997 réalisée par Actionaid, Kenya, qui a recueilli les opinions des utilisateurs réels et potentiels des services de vulgarisation kenyans, et une évaluation participative de la pauvreté réalisée en 1994 conjointement par le gouvernement kenyan, l'African Medical and Research Foundation, l'administration britannique du développement outre-mer, l'UNICEF et la Banque mondiale, qui s'est intéressée aux services publics, y compris aux services de vulgarisation. La qualité et la quantité des services fournis ont été évaluées en associant les résultats des évaluations participatives (des bénéficiaires) et des enquêtes auprès du personnel, ainsi que par le biais de mesures de la portée, de la nature et de la fréquence des contacts entre

les agents de vulgarisation et les fermiers tirées de l'enquête réalisée par l'OED en 1998. Les méthodes d'évaluation conditionnelle ont été utilisées pour inciter les fermiers à payer des services de vulgarisation.

Les données de l'enquête ont également été utilisées pour évaluer les résultats du programme en termes de connaissance des recommandations relatives à la vulgarisation par les fermiers et de leur aptitude à les adopter. Les résultats du programme, ses effets réels sur la production agricole au Kenya, ont été évalués en comparant la prestation de services de vulgarisation aux changements de productivité et d'efficacité au niveau des exploitations. En se basant sur les données de panel des ménages, on a estimé ces impacts à l'aide d'une analyse de l'enveloppement des données (technique non paramétrique) afin de mesurer l'évolution de l'efficacité et de la productivité des fermiers, et d'une analyse économétrique qui a mesuré l'impact de la prestation de services de vulgarisation sur la production de l'exploitation.

#### **C.5.4 Résultats**

Le développement institutionnel des projets NEP I et NEP II a été limité. Quinze ans plus tard, l'efficacité des services de vulgarisation a peu progressé. En dépit d'une rénovation récente des approches de la vulgarisation qui a été utile, le programme de vulgarisation a manqué dans l'ensemble de vision stratégique pour un développement futur. La gestion du système continue à être faible et les systèmes d'information sont pratiquement inexistantes. La qualité et la quantité des prestations de services sont médiocres. Les bénéficiaires et le personnel des services de vulgarisation sont unanimes : les visites sont rares et inefficaces. Alors que les demandes de services techniques utiles restent insatisfaites, le service public de vulgarisation a continué à mettre l'accent sur des messages agronomiques simples et élémentaires. Pourtant, l'approche retenue, des contacts très fréquemment renouvelés avec un nombre de fermiers restreint, est plutôt bien adaptée à la diffusion des informations techniques. Il en a résulté un système de prestation de services coûteux et inefficace. L'analyse a montré que les activités de vulgarisation avaient peu d'influence sur la connaissance et l'adoption des recommandations, d'où un potentiel d'impact limité. En termes d'impact réel sur la production et les performances agricoles, les données ont montré un faible impact positif des services de vulgarisation sur la capacité des fermiers à produire le plus possible à partir des ressources disponibles (efficacité technique). En revanche, aucun effet n'a été révélé sur l'efficacité d'allocation (utilisation des ressources compte tenu des prix du marché) ou sur l'efficacité économique globale (somme de l'efficacité technique et de l'efficacité d'allocation). En outre, aucun impact significatif de la prestation de services de vulgarisation sur la productivité au niveau de l'exploitation n'a pu être établi. Par contre, les données ont montré que l'impact était plutôt supérieur dans les régions auparavant moins productives, dans lesquelles les lacunes techniques étaient certainement les plus importantes. Ces résultats étaient cohérents avec les résultats de l'évaluation conditionnelle : une grande majorité de fermiers, à la fois parmi les bénéficiaires et les non-bénéficiaires actuels du programme, acceptait de payer pour recevoir des conseils, ce qui est le signe de besoins non satisfaits. Mais la valeur perçue du service, en termes de montant proposé, était bien en deçà des dépenses du gouvernement pour ce service.

#### **C.5.5 Implications sur la politique générale**

L'évaluation du service de vulgarisation kenyan se démarque par l'éventail de conclusions pratiques qu'il est possible de déduire des résultats, dont bon nombre sont pertinentes pour la conception de projets futurs de vulgarisation agricole. Tout d'abord, l'évaluation a révélé la nécessité d'améliorer le ciblage des services de vulgarisation, qui devra se concentrer sur les régions et les groupes où la différence entre les pratiques moyennes et supérieures est la plus forte et où l'impact devrait donc être maximum. Ensuite, les conseils doivent être soigneusement élaborés pour répondre aux besoins des fermiers, en tenant compte des variations locales au niveau des conditions technologiques et économiques. Parvenir à un tel niveau de personnalisation des services requiert des flux réguliers et opportuns d'informations appropriées et fiables, ainsi qu'un système de suivi et d'évaluation capable d'apporter un feedback régulier des bénéficiaires sur le contenu des services.

Pour accroître l'efficacité du programme, il faudrait sans doute réduire la présence des agents de vulgarisation tout en élargissant la couverture. Les innovations techniques ne sont pas suffisantes pour justifier une fréquence de visites élevée mais il y a une demande qui n'est pas satisfaite de la part des fermiers qui ne bénéficient pas de ces services. L'approche globale du programme vis-à-vis de la prestation de services, qui dépend surtout d'une méthodologie unique (visites des exploitations) axée sur la transmission de messages standard simples, limite également l'efficacité du programme. Aujourd'hui,

les programmes radiodiffusés sont populaires, les jeunes fermiers sont plus instruits et d'autres prestataires (organisations non gouvernementales) commencent à intervenir dans les campagnes kenyanes. Une approche pluraliste et souple de la prestation de services, qui utilise notamment des moyens de communication à coûts réduits, devrait permettre d'améliorer la rentabilité du programme.

Enfin, les principaux résultats ont révélé la nécessité de pratiquer une réforme institutionnelle. L'axe central de l'institution doit être le fermier. La décentralisation de la conception du programme, dont les mécanismes participatifs qui permettent aux fermiers de s'exprimer (comme le partage des coûts et les organisations d'exploitants), devrait faire partie intégrante du mécanisme de prestation. La pérennité financière est essentielle. L'ampleur et la densité du service devraient se fonder sur les déficits de technologie et de connaissance qui existent et sur le rythme de progression des nouvelles technologies. Le recouvrement des coûts, même s'il n'est que partiel, offre plusieurs avantages : il apporte des incitations appropriées, aborde les problèmes de responsabilisation et de contrôle de la qualité, oriente les services vers la demande et les rend plus réactifs et procure un certain répit budgétaire. Une telle décentralisation des institutions reste ignorée au Kenya et dans de nombreux programmes de vulgarisation en Afrique et ailleurs.

### C.5.6 Coûts d'évaluation et administration

**Coûts.** On estime à 350 000 dollars le coût total de l'évaluation, couvrant la collecte des données par des enquêtes auprès des ménages et leur traitement (65 000 dollars, bien que ce montant soit très certainement sous-évalué par rapport aux coûts réels) ; l'enquête auprès du personnel chargé de la vulgarisation, les données et les rapports des consultants (12 500 dollars) ; et d'autres coûts de collecte de données (12 500 dollars) ; auxquels il faudrait ajouter, pour refléter vraiment le coût total de l'étude, environ 160 000 dollars pour les heures de travail et les déplacements du personnel de la Banque mondiale liés au traitement des données, à leur analyse et à la rédaction de rapports.

**Administration.** Pour plus d'objectivité et pour dissocier le travail d'enquête du service de vulgarisation du gouvernement et de celui de la Banque mondiale, l'enquête auprès des ménages a été mise en œuvre par le Tegemeo Institute of Egerton University, institut de recherche indépendant au Kenya. L'analyse a été réalisée par les services de la Banque mondiale.

### C.5.7 Enseignements tirés

Associer une évaluation basée sur une théorie à un cadre fondé sur des résultats peut former une base solide pour évaluer l'impact des interventions de projets, notamment lorsque de nombreux facteurs risquent d'affecter les résultats escomptés. La conception de cette évaluation a permis de mesurer les indicateurs clés à des stades critiques du cycle du projet, en reliant les entrées du projet aux résultats escomptés de manière à rassembler suffisamment de preuves de l'impact.

Une évaluation empirique nécessite une supervision constante et rigoureuse. Un système de suivi et d'évaluation bien huilé et d'excellente qualité, qui s'appuie notamment sur une bonne base de données de référence facilitera grandement l'évaluation. On dispose rarement de ressources adéquates pour ces activités. Cette évaluation a également bénéficié d'un accès, même limité, à des données lors de la phase d'avant-projet et également à des sources de données indépendantes à des fins comparatives.

Pour rassembler un faisceau de preuves crédible, il est important de procéder à une vérification des conclusions via différentes approches analytiques et sources de données. Des données imparfaites et des problèmes de mise en œuvre restreignent le degré de confiance accordé aux réponses apportées aux questions d'évaluation clés par le biais de méthodes individuelles. Les évaluations qualitatives et quantitatives sont parfaitement complémentaires. L'expérience tirée de cette évaluation montre que même en l'absence d'évaluation participative des bénéficiaires, on peut intégrer des questions soigneusement élaborées dans une enquête afin de collecter des informations qualitatives autant que quantitatives. Ces informations peuvent apporter des données précieuses en complément des évaluations quantitatives.

Correctement appliquée, l'évaluation conditionnelle peut s'avérer un outil utile, surtout pour évaluer la valeur d'un service public en place. Les résultats de l'application de cette évaluation sont encourageants, et les réponses semblent être rationnelles et raisonnables.

*Source :* Gautam, Madhur. 1999. "World Bank Agricultural Extension Projects in Kenya: An Impact Evaluation." Report No. 19523. World Bank, Operations Evaluation Department, Washington, D.C.

## Étude de cas C.6 Évaluation de la réforme scolaire au Nicaragua : une approche quantitative et qualitative

### C.6.1 Introduction

**Description du projet.** En 1991, le gouvernement nicaraguayen a engagé une réforme radicale de son système d'éducation public : il a décentralisé la gestion des écoles (décisions relatives au personnel, aux budgets, aux programmes et à la pédagogie) et transféré les responsabilités de financement au niveau local.

Les réformes ont été introduites progressivement, en commençant en 1991 par un décret établissant des conseils parentaux et communautaires dans toutes les écoles publiques. Puis, en 1993, un programme pilote réalisé dans 20 écoles secondaires triées sur le volet a transformé ces conseils en conseils de gestion des écoles dotés de plus grandes responsabilités sur les plans du personnel, des budgets, des programmes et de la pédagogie. En 1995, les conseils de gestion des écoles étaient opérationnels dans 100 établissements secondaires et plus de 300 écoles primaires, qui ont rejoint le programme par le biais d'un processus d'auto-sélection basé sur une pétition rédigée par les enseignants et les directeurs d'école.

Il était prévu qu'à mesure que la gestion des écoles évoluait vers un processus démocratique et participatif et que les revenus générés localement augmentaient, les modes de dépenses devaient soutenir davantage les activités ayant une action directe sur la pédagogie et encourageant la réussite scolaire.

**Évaluation de l'impact.** L'évaluation des réformes concernant l'autonomie scolaire du Nicaragua représente l'une des premières actions systématiques d'évaluation de l'impact de la décentralisation des écoles sur les résultats des élèves. Ce style est novateur en ce qu'il associe des méthodes d'évaluation qualitatives et quantitatives. La composante quantitative est unique en son genre car elle intègre un module d'évaluation des processus décisionnels scolaires isolé. L'évaluation illustre également les meilleures techniques en l'absence de données de référence et quand une application sélective (non aléatoire) des réformes empêche toute évaluation expérimentale.

Le but de la composante qualitative de l'évaluation était de déterminer si les réformes de gestion et de financement prévues étaient effectivement appliquées dans les écoles et également de comprendre ce que les différentes parties prenantes pensaient du processus de réforme. La composante quantitative cherchait à répondre à la question suivante : Les changements dans la gestion et le financement des écoles permettaient-ils d'obtenir de meilleurs résultats du point de vue de l'apprentissage des enfants ? Les résultats qualitatifs ont montré que la réussite de la mise en œuvre des réformes dépendait en grande partie de la situation et de l'environnement propres à l'école (c'est-à-dire le niveau de pauvreté de la communauté), alors que les résultats quantitatifs laissaient penser qu'une plus forte prise de décision par les écoles était en fait significativement liée à de meilleures performances scolaires.

### C.6.2 Style d'évaluation

L'évaluation de la réforme concernant l'autonomie des écoles au Nicaragua se base sur la comparaison de données provenant d'un échantillon représentatif d'écoles participant au programme à des données issues d'écoles n'y participant pas. On choisit l'échantillon d'écoles non-participantes de telle sorte qu'il corresponde le plus possible aux caractéristiques des écoles participantes ; il constitue la situation fictive. On a retenu ce style parce que l'absence de données de référence rendait impossible toute évaluation avant et après le programme et que les réformes n'étaient pas appliquées de façon aléatoire aux écoles, proscrivant toute évaluation expérimentale.

### C.6.3 Techniques de collecte et d'analyse des données

L'étude qualitative s'appuie sur les données provenant d'un échantillon de 12 écoles, 9 bénéficiant de la réforme, et 3 n'en bénéficiant pas et constituant le groupe témoin.<sup>11</sup> L'échantillon de 12 écoles a été sélectionné pour représenter des écoles primaires et secondaires, rurales et urbaines, ainsi que des écoles bénéficiant de divers degrés d'autonomie réelle dans la prise de décision. Au total, 82 entretiens et réunions de groupe ont été réalisés et ont cherché à évaluer la compréhension que les directeurs d'école, membres du conseil, parents et enseignants avaient du processus de décentralisation et ce qu'ils en pensaient. Tous les entretiens ont été menés par des Nicaraguayens formés à poser des questions dirigées sans donner de pistes pour la réponse, par le biais de simulations d'entretiens et d'essais pilotes. Les entretiens étaient enregistrés, transcrits puis condensés dans un rapport de deux à quatre pages qui était

ensuite analysé afin d'identifier des ensembles discrets de preuves et de thèmes fondamentaux émergeant entre les écoles et les acteurs, et entre les écoles réformées et le groupe témoin.

La collecte des données quantitatives était constituée de deux éléments : une enquête de panel des écoles, réalisée en deux phases (de novembre à décembre 1995, et d'avril à août 1997), et des tests de connaissances proposés aux élèves de ces écoles, réalisés en novembre 1996. L'enquête auprès des écoles a recueilli des données sur le taux d'inscription scolaire ; les taux de redoublement et d'abandon ; les ressources physiques et humaines ; le processus décisionnel suivi dans l'établissement ; et les caractéristiques des directeurs d'école, des enseignants, des élèves et de leur famille. Le module décisionnel appliqué aux écoles est unique en son genre : il propose une série de 25 questions qui permettent de mesurer, le cas échéant, l'effet de la réforme sur l'augmentation du pouvoir décisionnel des écoles. L'enquête a concerné 116 écoles secondaires (73 bénéficiant de la réforme et 43 n'en bénéficiant pas et formant le groupe témoin), et 126 écoles primaires (80 avec réforme et 46 sans). Ici également, les groupes témoins ont été choisis pour correspondre aux caractéristiques des écoles réformées. L'enquête a également rassemblé des données sur 400 enseignants, 182 membres de conseil et 3 000 élèves et leurs parents, de 10 à 15 élèves étant choisis au hasard dans chaque école.

L'analyse quantitative des données se base sur des techniques de régression afin d'évaluer une fonction de production de l'éducation. Elle a étudié l'impact du degré d'autonomie des écoles sur les niveaux de réussite scolaire, avec contrôle des variables d'entrée des écoles et des caractéristiques des ménages et des élèves. L'analyse a mesuré l'effet de l'autonomie de droit et de fait. L'autonomie de droit indique simplement si l'école a rejoint la réforme sur le plan juridique, alors que l'autonomie de fait mesure le degré d'autonomie réelle dont bénéficie l'école (en pourcentage) à l'aune de 25 décisions clés prises directement par l'école. On anticipait une variation de l'autonomie de fait entre les écoles puisque les réformes ont été introduites progressivement (de sorte que les écoles de l'échantillon se situaient à des stades différents du processus de réforme), et puisque la capacité à réussir la mise en œuvre des réformes variait selon la situation de l'école, résultat identifié dans l'étude qualitative).

#### **C.6.4 Résultats**

L'étude qualitative a souligné que les changements d'orientation de la politique au niveau central ne se traduisaient pas toujours par des changements au niveau local. En général, les réformes ont été associées à une plus forte participation des parents et à des améliorations de la gestion et du leadership. Mais le degré de réussite dans la mise en œuvre des réformes a varié en fonction de la situation de l'école. Le degré d'appauvrissement de la communauté avoisinante (dans les communautés pauvres, augmenter le financement des écoles locales a été difficile), et le degré de cohésion entre les membres du personnel de l'école (lorsque des acteurs clés comme les enseignants ne se sentaient pas intégrés au processus de réforme, la décentralisation était limitée) revêtaient une importance particulière. Les décideurs ignorent souvent l'extrême variabilité des situations locales où sont introduits les nouveaux programmes. Les résultats du Nicaragua ont montré que l'objectif d'accroissement du financement local des écoles avait peu de chances d'être atteint dans les communautés pauvres.

L'étude quantitative a renforcé les conclusions que les écoles réformées prenaient effectivement plus de décisions, notamment sur les sujets relatifs à la pédagogie et au personnel. L'autonomie de droit (acquise lorsque l'école a signé le contrat de réforme) ne s'est pas forcément traduite par un niveau de prise de décision de l'école plus fort pas plus qu'elle n'a affecté en conséquence le mode de fonctionnement de l'école. Le degré d'autonomie atteint dépendait du niveau de pauvreté de la communauté et de la période de participation de l'école au processus de réforme. Les résultats de l'analyse de régression ont montré que l'autonomie de droit avait peu d'influence sur les résultats en matière de réussite scolaire, alors que l'autonomie de fait améliorerait substantiellement ces résultats. En outre, des simulations montrent qu'une plus grande autonomie de l'école a plus d'influence sur la réussite scolaire que toute autre action classique, telle que l'augmentation du nombre de manuels, l'accroissement de la formation des enseignants et la réduction du nombre d'élèves par classe.

#### **C.6.5 Implications sur la politique générale**

Les résultats de l'évaluation démontrent de manière concrète que la réforme scolaire au Nicaragua a produit des résultats tangibles. Les écoles réformées prenaient plus de décisions localement et cette plus grande prise de décision locale a amélioré la réussite scolaire.

Les résultats ont également souligné les domaines dans lesquels il était possible d'améliorer la politique de réforme, et le ministère de l'Éducation a introduit plusieurs modifications dans le programme de réforme scolaire. Le programme place davantage l'accent sur la formation des enseignants et la promotion des aspects pédagogiques de la réforme. De plus, en réponse aux problèmes de financement rencontrés par les communautés pauvres, le ministère a élaboré un projet de subventions basé sur une carte de la pauvreté. Enfin, les bénéfices tangibles de cette évaluation ont poussé le ministère à intégrer au programme de réforme une composante d'évaluation continue.

### C.6.6 Coûts d'évaluation et administration

**Coûts.** On estime à quelques 495 000 dollars le coût total de l'évaluation, soit moins de 1,5 % du crédit accordé par la Banque mondiale pour le soutien des réformes.<sup>12</sup> De cette somme totale, 39 % ont été dépensés au support technique fourni par des consultants externes, 35 % à la collecte des données, 18 % en heures de travail des services de la Banque mondiale et 8 % aux déplacements.

**Administration.** L'évaluation a été réalisée conjointement par le ministère de l'Éducation du Nicaragua, la Banque mondiale et des chercheurs de Harvard School of Education.

### C.6.7 Enseignements tirés

**Valeur de l'approche des méthodes mixtes.** En utilisant à la fois des techniques de recherche qualitatives et quantitatives, il a été possible d'obtenir un ensemble précieux de résultats utiles et pertinents par rapport à la politique générale de réforme. L'étude quantitative a offert un large panorama statistique de la situation et des résultats des écoles, renforcé par l'étude qualitative dont les données ont expliqué pourquoi certaines écoles obtenaient les résultats escomptés quand d'autres ne les obtenaient pas, permettant ainsi d'effectuer des ajustements d'orientation. En outre, étant plus intuitive, l'étude qualitative était plus accessible et a donc intéressé le personnel du ministère, qui en retour a facilité une édification rapide de la capacité et accru la crédibilité du processus d'évaluation au sein du ministère.

**Importance de l'édification de la capacité locale.** L'édification de la capacité locale s'est révélée coûteuse et a nécessité de fréquents contacts et une étroite coordination avec les fonctionnaires nicaraguayens, leurs homologues de la Banque mondiale et des consultants externes. Cependant, ceci a permis de développer rapidement la prise en compte de manière responsable du processus d'évaluation par les acteurs locaux, et d'obtenir un fort degré d'acceptation des résultats de l'évaluation, positifs comme négatifs. Les résultats de cette évaluation ont permis d'intervenir directement dans la réforme à mesure qu'elle évoluait. L'impact de l'évaluation au niveau de la politique a été également renforcé par la cohésion de l'équipe locale dans laquelle évaluateurs et décideurs travaillaient de concert et par le soutien du ministère de l'Éducation.

*Sources et autres lectures :*

- Fuller, Bruce, and Magdalena Rivarola. 1998. "Nicaragua's Experiment to Decentralize Schools: Views of Parents, Teachers, and Directors." Working Paper Series on Impact Evaluation of Education Reforms 5. World Bank, Development Economics Research Group, Washington, D.C.
- King, Elizabeth, and Berk Ozler. 1998. "What's Decentralization Got to Do with Learning? The Case of Nicaragua's School Autonomy Reform." Working Paper Series on Impact Evaluation of Education Reforms 9. World Bank, Development Economics Research Group, Washington, D.C.
- King, Elizabeth, Berk Ozler et Laura Rawlings. 1999. "Nicaragua's School Autonomy Reform: Fact or Fiction?" Working Paper Series on Impact Evaluation of Education Reforms 19. World Bank, Development Economics Research Group, Washington, D.C.
- Nicaragua Reform Evaluation Team. 1996. "Nicaragua's School Autonomy Reform: A First Look." Working Paper Series on Impact Evaluation of Education Reforms 1. World Bank, Poverty and Human Resources Division, Policy Research Department, Washington, D.C.
- . 1996. "1995 and 1997 Questionnaires, Nicaragua School Autonomy Reform." Working Paper Series on Impact Evaluation of Education Reforms 7. World Bank, Development Economics Research Group, Washington, D.C.
- Rawlings, Laura B. 2000. "Evaluating Nicaragua's School-Based Management Reform." In Michael Bamberger, ed., *Integrating Quantitative and Qualitative Methods in Development Research*. Washington, D.C.: World Bank.

## Étude de cas C.7 Résultats scolaires dans les écoles élémentaires philippines : évaluation de l'impact de quatre expériences

### C.7.1 Introduction

**Description du projet.** Dans la plupart des pays en voie de développement, les décideurs sont préoccupés par les forts taux d'abandon et le mauvais apprentissage des élèves durant le cycle primaire. C'était certainement le cas aux Philippines : près d'un quart des enfants philippins abandonnait l'école avant d'avoir atteint la sixième, et ce en ayant souvent acquis moins de la moitié des enseignements. Entre 1990 et 1992, le gouvernement s'est engagé dans un programme d'action contre l'abandon scolaire (DIP, ou Dropout Intervention Program) pour faire face à ces problèmes. Quatre expériences ont été menées : fourniture de supports didactiques à niveaux multiples (MDNM) et de repas scolaires (RS), ces deux expériences étant chacune associée à un partenariat parents-enseignants (PPE). La fourniture de supports didactiques à niveaux multiples permet aux enseignants d'ajuster leur enseignement aux divers besoins des élèves et la fourniture de repas est bien moins onéreuse que les programmes d'alimentation scolaire. Les partenariats parents-enseignants ne coûtent pratiquement rien et pourtant ils peuvent aider l'élève à travailler ses leçons chez lui comme à l'école.

**Évaluation de l'impact.** Cette évaluation est remarquable en ce qu'elle vise explicitement à édifier la capacité du pays hôte de manière à ce que l'évaluation devienne une partie intégrante des futures initiatives ; par ailleurs, les besoins en matière de données sont étudiés avant et non après la mise en œuvre du projet. Une autre contribution majeure de l'évaluation a été de vérifier la fiabilité des résultats par différentes approches économétriques. Enfin, l'analyse coûts-bénéfices réalisée à la fin de l'évaluation était importante car elle reconnaissait explicitement le fait que des résultats significatifs ne suffisent pas à justifier une intervention : les interventions peu coûteuses restent parfois préférables à d'autres plus coûteuses.

### C.7.2 Style d'évaluation

L'objectif clé de l'étude était d'évaluer l'impact de quatre actions différentes sur les taux d'abandon et les résultats scolaires. Le style d'évaluation était conditionné par les exigences tant pragmatiques que programmatiques. L'équipe du DIP a suivi un processus de sélection des écoles en trois phases :

- deux districts à faibles revenus ont été identifiés dans chacune des cinq régions du pays ;
- dans chaque district, l'équipe a retenu trois écoles qui a) offraient le cycle d'éducation primaire complet avec une classe par niveau ; b) présentaient un fort taux d'abandon ; et c) n'avaient aucun programme d'alimentation scolaire en place ;
- dans un district, les options disponibles étaient MDMN, MDMN-PPE, ou rien ; dans l'autre, RS, RS-PPE, ou rien.

Les trois écoles de chaque district étaient affectées au groupe témoin ou à l'une des deux actions selon une sélection aléatoire.

Des tests dans trois matières (mathématiques, philippin et anglais) ont été réalisés avant et après l'intervention, au début et à la fin des années scolaires 1991 et 1992 dans toutes les classes des 30 écoles.

### C.7.3 Techniques de collecte et d'analyse des données

La collecte des données de référence a débuté en 1990-1991, et les interventions ont été mise en œuvre en 1991-1992. Des informations approfondies ont été réunies pour 29 écoles, quelque 180 enseignants et environ 4 000 élèves pour chacune des deux années. Alors que ces questionnaires étaient très détaillés, il s'est avéré qu'une grande partie des questions étaient inutiles : seule une petite portion des informations a été effectivement utilisée, indiquant par là qu'il était possible de réduire le poids du processus d'évaluation.

Les données étaient structurées de manière à suivre à la fois les élèves et les écoles sur les deux années ; malheureusement il est apparu que les identifiants des élèves n'étaient pas uniques. Il convient de remarquer que l'on ne le savait pas a priori, et que cela n'est devenu évident que lorsque six mois de travail ont révélé des incohérences internes. Il a été impossible de retrouver les identifiants originaux auprès du département de l'éducation des Philippines. Il a heureusement été possible de récupérer les données des CP et on a ainsi pu réaliser une analyse longitudinale.

La structure de la procédure d'échantillonnage a soulevé d'intéressants problèmes économétriques pour les taux d'abandon et les résultats des tests scolaires. Dans chaque cas, il y a deux ensembles de témoins évidents : le premier est le groupe témoin d'écoles ; le deuxième, l'enquête de référence réalisée dans l'année qui a précédé l'intervention. Les auteurs ont traité différemment ces outils témoins.

Dans l'analyse des taux d'abandons, on utilise habituellement une approche de différence en différences pour comparer la modification du taux d'abandon moyen de chaque classe bénéficiant de l'intervention entre les deux années avec celle des classes témoins. Mais deux problèmes se sont immédiatement posés. Tout d'abord les résultats, bien qu'importants en nombre, n'étaient significatifs que pour l'intervention MDNM, ce qui est probablement dû à la petitesse des échantillons. Ce phénomène n'est pas rare dans ce type de procédure. Il est dû en partie au manque de financement pour des expériences à large échelle dans les pays en voie de développement. Ensuite, une rapide vérification destinée à savoir si les caractéristiques et les résultats des élèves étaient en fait semblables dans toutes les écoles l'année précédant les interventions a suggéré que ces caractéristiques présentaient certaines différences significatives. Ces deux facteurs ont mené les auteurs à vérifier la robustesse des résultats via des techniques de régression logistique qui vérifiaient les caractéristiques personnelles et la situation familiale. Le résultat principal n'a pas bougé. Mais la technique de régression a révélé une importante cause essentielle indirecte des abandons, à savoir de mauvaises performances scolaires. Cela a naturellement mené à un deuxième ensemble d'analyses concentré sur la réussite.

Plusieurs problèmes économétriques se sont posés lors de l'évaluation de l'impact de l'intervention sur les performances scolaires d'un individu dans une école donnée et à un moment donné : tenir compte de la corrélation groupée d'erreurs susceptible d'exister pour des élèves d'une même classe et d'une même école ; capturer toute hétérogénéité non observée ; et aborder les problèmes de distorsions dans la sélection (si les élèves ayant de moins bonnes performances scolaires ont plus de risques d'abandonner l'école, alors les estimations des effets du programme peuvent être faussées à la hausse).

Pour résoudre le premier problème, on a appliqué une correction de Huber-White aux erreurs types. Pour le deuxième, on pouvait en principe y remédier au niveau individuel en utilisant la différence dans les résultats des tests comme variable indépendante. Mais les auteurs ont avancé que cette méthode ne convenait pas parce qu'elle présupposait que la valeur du coefficient de performance scolaire était de 1, ce qui n'a pas été validé par les tests. Ils ont donc utilisé les performances scolaires de la première période comme variable indépendante, mais cela a alors posé le problème de la corrélation avec le terme d'erreur, ou distorsion endogène de l'estimateur. Ce problème est résolu en utilisant comme variable instrumentale des résultats préalables au test dans chaque matière, les résultats préalables au test dans les autres matières. Les auteurs ont remarqué cependant que la réduction de la distorsion se faisait au prix d'une réduction de l'efficacité, et ils ont donc indiqué les résultats des moindres carrés et ceux des variables instrumentales. Ils ont utilisé les effets fixes des écoles et enseignants pour déceler toute hétérogénéité non observée dans l'environnement d'apprentissage et les conditions de travail dans les classes.

Le troisième problème, celui de la distorsion de la sélection, est un problème qui revient souvent dans la littérature, et pour lequel il n'existe aucune solution faisant l'unanimité. Il est évident que les performances scolaires individuelles dépendent de la décision de ne pas abandonner l'école. Bien que ce problème ait souvent été traité à l'aide de la procédure en deux phases de Heckman (Heckman 1976, 1979), le résultat reste insuffisant pour trois raisons : sa sensibilité à l'hypothèse d'une répartition normale, le choix et la pertinence des variables appropriées utilisées lors de la première étape, et son fréquent recours à l'identification par la non-linéarité de la première étape. Mais il n'existe toujours aucun consensus sur une autre solution appropriée. Dans la solution proposée par Krueger (1997), on attribue des résultats scolaires aux élèves qui quittent l'échantillon en leur attribuant leur percentile de test le plus récent pour les années qu'ils n'ont pas faites puisqu'ils ne faisaient pas partie de l'échantillon. Les auteurs signalent trois séries de résultats : la simple régression des résultats par rapport à l'intervention, la méthode Krueger, la procédure de Heckman et les résultats des variables instrumentales obtenus avec la procédure de Krueger.

#### **C.7.4 Résultats**

L'étude a démontré que l'effet positif des matériels didactiques à niveaux multiples, notamment associés à un partenariat parents-enseignants, sur les taux d'abandons et les performances scolaires était robuste par rapport aux différentes spécifications. L'effet des repas scolaires était en général faible. L'analyse coûts-bénéfices s'est révélée une composante intéressante de l'étude, et a démontré un point majeur, à savoir qu'il ne suffisait pas d'avoir des résultats significatifs. Un simple calcul des coûts directs et

indirects (opportunité) des programmes a mené à la conclusion que l'approche MDNM était à la fois efficace et rentable.

Il se peut néanmoins que le manque d'efficacité du programme d'alimentation scolaire soit surestimé : une approche plus ciblée de ce type de programmes, cibler par exemple seulement les enfants souffrant de malnutrition ou économiquement faibles, pourrait être plus rentable. De plus, étant donné le court délai entre la mise en œuvre du programme et son évaluation, il est possible que cette évaluation n'aborde pas l'impact à long terme des interventions.

### **C.7.5 Administration**

La collecte des données a été réalisée par le bureau de l'éducation élémentaire du département de l'éducation, de la culture et des sports philippin. L'analyse a été exécutée par un collaborateur de la Banque mondiale et deux chercheurs universitaires.

### **C.7.6 Enseignements tirés**

Plusieurs leçons ont été tirées au cours de cette évaluation. Une des plus importantes est qu'il est possible de perdre une quantité considérable d'informations longitudinales vitales si les identifiants des observations (ici, les élèves) ne sont pas uniques dans le temps. Deuxièmement, seule une quantité infime des informations collectées par le biais des enquêtes approfondies a été utilisée ; ainsi la charge sur les personnes interrogées pourrait être très largement réduite. Troisièmement, l'étude a souligné la valeur de différentes approches économétriques et les avantages qu'il y a à trouver des résultats cohérents entre les différentes techniques. Quatrièmement, l'étude était exemplaire à la fois pour identifier et évaluer les coûts des différentes interventions. Enfin, bien que des erreurs aient été clairement commises lors de cette étude, les auteurs ont noté que l'un des objectifs majeurs était d'édifier une capacité d'évaluation aux Philippines.

*Sources et autres lectures :*

- Heckman, James J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Economic and Social Measurement* 5:475-492.
- Heckman, James J. 1979. "Sample Selection as a Specification Error." *Econometrica* 47:153-161.
- Krueger, Alan. 1997. "Experimental Estimates of Educational Production Function," NBER Working Paper 6051, Cambridge, MA
- Tan, J. P., J. Lane, and G. Lassibille. 1999. "Schooling Outcomes in Philippine Elementary Schools: Evaluation of the Impact of Four Experiments." *World Bank Economic Review* 13(3):493-508.

## Notes

1. Une version précédente de cette étude de cas a été élaborée par Margaret Kakande, Directrice de l'unité de suivi de la pauvreté, Ministère des finances, du plan et du développement économique, gouvernement d'Ouganda et Kimberley McClean, Directrice des projets internationaux, Aus Health International.
2. Pour plus d'informations sur l'enquête nationale sur la prestation de services, aller sur le site du département de la recherche de la Banque mondiale sur la prestation de services publics à l'adresse [http://www.worldbank.org/research/projects/publicspending/tools/tools.htm#Quantitative Service Delivery](http://www.worldbank.org/research/projects/publicspending/tools/tools.htm#Quantitative%20Service%20Delivery).
3. Voir <http://www.uppap.or.ug/>.
4. Disponible à l'adresse <http://www.worldbank.org/afr/findings/english/find142.htm>.
5. La Encuesta de Desarrollo Social a été financée dans le cadre d'un projet autre que Trabajar, soutenu par la Banque mondiale. Son but est d'améliorer la qualité des informations sur le bien-être des ménages en Argentine, notamment dans le domaine de l'accès aux services sociaux et des programmes sociaux mis en place par le gouvernement.
6. Le questionnaire EDSO est très complet. Il réunit des données détaillées sur les caractéristiques des ménages qui permettent de prévoir la participation au programme et de simplifier l'utilisation de la technique de notation à la propension. On calcule la notation à la propension pour chaque observation dans le groupe des participants et dans le groupe témoin à l'aide de modèles de répartition standard.
7. Il s'agit d'un indice composite qui comprend le surpeuplement résidentiel, les installations d'assainissement, la qualité du logement, le niveau d'instruction des adultes, l'inscription des enfants à l'école, l'emploi et la dépendance (ratio de membres de la famille qui travaillent par rapport à ceux qui ne travaillent pas). Bien que l'indice datait quelque peu, son avantage était que la mesure de pauvreté du département n'était pas influencée par les interventions du programme Trabajar.
8. Aucune étude n'a été faite pour étudier l'impact sur le bien-être des ménages, qui devrait être affecté par plusieurs facteurs qui vont bien au-delà de la portée des activités de F&V.
9. La méthode d'évaluation conditionnelle détermine les valeurs d'utilisation et de non-utilisation des individus pour une variété de biens et de services publics et privés. On demande aux personnes interrogées d'indiquer si elles seraient prêtes à payer afin d'éviter un changement hypothétique dans la disponibilité des biens et des services, soit le résultat « conditionnel ». Dans le cas présent, on a demandé aux fermiers quel serait le montant qu'ils seraient prêts à verser pour assurer la continuité des services de vulgarisation agricole, si le gouvernement devait les supprimer.
10. Ces trois enquêtes ont généré un ensemble de données de panel pour quelque 300 ménages. Les enquêtes couvrent les données démographiques des ménages, les caractéristiques des fermes et des données entrées-sorties sur la production agricole ; les enquêtes de 1990 et 1998 ont également rassemblé des informations sur les contacts avec les services de vulgarisation, dont la connaissance et l'adoption des informations de vulgarisation.
11. En fait, des données ont été recueillies auprès de 18 écoles, mais seules 12 de ces écoles ont été incluses dans l'analyse qualitative en raison des retards dans la préparation des transcriptions, et il a donc été décidé de concentrer la plus grosse partie de l'analyse sur les écoles réformées, qui ont fourni plus de matériel utile pour l'analyse.
12. Ce total ne tient pas compte du coût des équipes locales homologues au ministère de l'Éducation du Nicaragua.