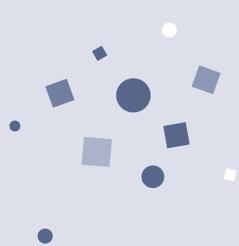


Série Notes méthodologiques

exPost
ExPost



Promouvoir les évaluations d'impact à
l'Agence Française de Développement :
renforcer l'appropriation et l'apprentissage institutionnels

Jean-David Naudet et Jocelyne Delarue
Agence Française de Développement



Département de la Recherche
Division Évaluation et capitalisation
Agence Française de Développement
5, rue Roland Barthes 75012 Paris < France
www.afd.fr

Avertissement

Les analyses et conclusions de ce document sont formulées sous la responsabilité de leurs auteurs. Elles ne reflètent pas nécessairement le point de vue officiel de l'Agence Française de Développement ou de ses institutions partenaires.

Directeur de la publication : Jean-Michel SEVERINO

Directeur de la rédaction : Jean-David NAUDET

ISSN : En cours

Dépôt légal : Décembre 2007

Mise en page : Solange Münzer

TABLE DES MATIÈRES

	Résumé	4
1.	La stratégie de l'AFD en matière d'évaluations d'impact	5
1.1.	La question de l'impact est liée à la production de connaissances et à la gestion axée sur les résultats	5
1.2.	Les évaluations d'impact, une partie importante de l'activité d'évaluation	6
2.	Une palette d'expériences	8
2.1.	Une phase préliminaire	8
2.2.	La deuxième évaluation d'impact financée par l'AFD	9
2.3.	Une évaluation d'impact expérimentale de la microfinance au Maroc	11
2.4.	Une évaluation d'impact expérimentale de la micro-assurance de santé au Cambodge	13
3.	Premiers enseignements	15
3.1.	L'apprentissage institutionnel	15
3.2.	L'évaluation d'impact pour un bailleur bilatéral	15
3.3.	Perspectives	16
	Références	18

Résumé

En 2006, la division Évaluation de l'Agence Française de Développement (AFD) a été intégrée au Département de la Recherche, dans le but d'introduire une dimension de recherche dans le travail d'évaluation. Cette nouvelle approche a conduit l'AFD à entreprendre des évaluations rigoureuses d'impact en partenariat avec des équipes universitaires. Deux d'entre elles sont finalisées, l'une sur Madagascar (microfinance) et l'autre sur la Guinée (développement agricole). Deux évaluations d'impact expérimentales ont été lancées récemment au Maroc (microfinance) et au Cambodge (micro-assurance de santé). En plus d'apporter une réponse aux exigences de redevabilité, il s'agit à travers ces exercices de produire des connaissances solides sur les questions de développement et d'appuyer l'élaboration des politiques. Par ailleurs, la forte implication de l'AFD et de ses partenaires locaux dans ces évaluations vise au renforcement des capacités internes. Il s'agira désormais de répondre à un nombre croissant de demandes internes pour réaliser des évaluations rigoureuses d'impact ce qui, vu les contraintes budgétaires et temporelles, impose de faire appel à toute une gamme de méthodes.

1. La stratégie de l'AFD en matière d'évaluations d'impact

1.1. La question de l'impact est liée à la production de connaissances et à la gestion axée sur les résultats

L'AFD a pour mission de contribuer au financement de projets de développement à vocation économique, sociale et/ou environnementale.

L'AFD fournit une assistance en matière de développement aux pouvoirs publics (État, entreprises publiques et collectivités locales), au secteur privé et aux réseaux associatifs locaux. Elle propose une large gamme d'instruments financiers pour la mise en œuvre de leurs projets de développement durable. L'Agence poursuit trois grands objectifs : (i) réduire la pauvreté et les inégalités d'ici 2015 (objectifs du Millénaire pour le développement – OMD) ; (ii) promouvoir la

croissance économique ; et (iii) protéger les biens communs de l'humanité (climat, biodiversité, santé).

Depuis le début de cette décennie et dans un processus plus large de réforme de la coopération française, l'AFD s'est engagée à renouveler ses orientations stratégiques.

Parmi les changements stratégiques opérés par l'AFD, deux méritent particulièrement d'être soulignés. Le premier a trait à la mise en œuvre d'une gestion axée sur les résultats, conformément aux engagements de la Déclaration de Paris. Dès 2002, le premier projet d'orientation stratégique de l'AFD (POS I) exprimait un réel souci d'évaluer l'impact de ses opérations. Il recommandait la mise en place d'un système de suivi axé sur les résultats, fondé en partie sur des indicateurs d'impact, ainsi que l'introduction d'un système de sélection des projets à financer en fonction des évaluations d'impact.

La direction de l'AFD accorde une place essentielle à la prise en compte des résultats et impacts de développement. À tous les niveaux, il s'agit d'accroître et de montrer l'efficacité de l'aide au développement.

L'approche de la « gestion axée sur les résultats » (GRD) est systématique tout au long du projet. Des indicateurs agrégés contrôlent les résultats attendus et obtenus en termes de développement. Les définitions sont normalisées et harmonisées avec celles des organisations internationales. Le suivi de la contribution aux OMD mesure à la fois l'engagement de la France et de l'AFD, en termes de moyens comme de résultats. L'analyse économique des projets de développement a pour but d'évaluer, au-delà de leur viabilité financière, leurs coûts et avantages économiques à l'échelle de la société tout entière, biens et services environnementaux compris. L'analyse des avantages que chaque partenaire retire d'un projet permettra de décider des mécanismes de transfert.

Le second changement oriente clairement les activités de l'Agence vers la production de connaissances, complément indispensable de l'activité financière. En accord avec le second projet d'orientation stratégique (AFD, 2007), cet objectif concernera particulièrement les grands sujets de l'APD – pour contribuer à l'élaboration des positions politiques françaises, pour rendre possible une participation au renforcement des capacités de ses partenaires et pour alimenter le débat international.

Ces nouvelles orientations stratégiques ont permis d'engager un débat institutionnel sur la question de l'impact, entraînant une prise de conscience généralisée du déficit d'information entourant l'impact des opérations de développement (le fameux « manque de connaissance » évoqué dans le rapport du groupe de travail sur le déficit d'évaluation du Center for Global Development [CGD, 2006]).

Bien que certains types de projets ou produits financiers – construction de routes en milieu rural ou microfinance, par exemple – aient mobilisé des financements substantiels, jamais jusqu'à une date récente l'AFD n'avait effectué de recherche rigoureuse sur l'efficacité de ses projets ni sur les résultats attribuables à ses programmes. Les évaluations réalisées jusqu'à présent n'ont fourni que très peu d'informations sur les effets à moyen ou long terme et aucune sur les impacts. Très logiquement, les gestionnaires privilégient, dans les premières phases d'un projet, l'élaboration et la mise en œuvre et remettent à plus tard la prise des décisions requises pour préparer des évaluations rigoureuses. Jusqu'ici, il n'y a guère eu d'incitations à concevoir un système d'information autour du projet, permettant d'en évaluer l'impact *ex post* ni de compensations prévues pour le coût et le temps passé à concevoir un tel système.

La nouvelle approche de l'AFD – reposant sur un positionnement stratégique et une gestion axée sur les résultats – oblige de plus en plus les services opérationnels, s'ils veulent obtenir des fonds, à démontrer l'impact de leurs interventions. Ceux-ci mettent donc davantage l'accent qu'auparavant sur la capitalisation et la mesure d'impact de leurs opérations. La généralisation de ces pratiques achoppe toutefois pour l'instant sur l'insuffisance des ressources humaines et budgétaires. Il semble d'ailleurs important que l'AFD vérifie d'abord si les évaluations d'impact rigoureuses permettent davantage que d'autres types d'évaluations d'améliorer la qualité du retour d'informations sur les opérations.

Parallèlement, le débat international sur l'évaluation d'impact met en évidence des difficultés conceptuelles, méthodologiques et pratiques. La question de l'attribution/contribution entre en ligne de compte avec l'application d'une gestion axée sur les résultats, notamment pour une institution comme l'AFD, qui est souvent partie prenante de cofinancements. La double acception du terme « impact » dans le discours sur le développement – couvrant à la fois les effets de long terme et les effets strictement attribuables au projet – demeure une source d'ambiguïté permanente. La subtile différence entre impact et additionnalité – concept plus fréquent (mais aussi plus équivoque) utilisé pour les mécanismes financiers et en

particulier à propos du changement climatique – constitue un sujet de préoccupation pour une institution qui finance directement des politiques publiques dans les pays pauvres et promeut, par des incitations financières, les biens publics mondiaux dans les pays émergents. Des questions méthodologiques sont directement liées à ces débats conceptuels : comment mesurer et attribuer les effets de long terme ? Comment élaborer un raisonnement contrefactuel ? Quel point de départ adopter ? Quels sont les indicateurs d'impact significatifs, etc.

Cette phase des débats est très enrichissante pour l'AFD. Elle contribue à persuader l'institution qu'une approche standard (un outil de type boîte noire de mesure d'impact applicable à de multiples situations) ne peut pas constituer une réponse satisfaisante à l'insuffisance des connaissances sur les impacts. L'AFD a donc décidé de poursuivre et de diversifier son investissement dans des approches méthodologiques et dans la production de connaissances en matière d'évaluation d'impact, notamment à travers des opérations pilotes et des partenariats spécialisés. Elle s'est également engagée dans le débat international sur cette question.

1.2. Les évaluations d'impact, une part importante de l'activité d'évaluation

En accord avec ces nouvelles orientations, l'AFD a engagé depuis 2006 une réforme de sa fonction d'évaluation. Le processus repose sur deux piliers : la décentralisation (vers les agences locales) de l'évaluation externe systématique de chaque financement ; et la réorientation de la division Évaluation vers une exigence de qualité, d'évaluation stratégique et de production de connaissances, y compris en termes de mesures d'impact.

Jusqu'à-là, la fonction d'évaluation était essentiellement interne, conduite et mise en œuvre au sein de la Direction de la stratégie par la division Évaluation. Environ 15 % des projets financés par l'AFD étaient soumis à une évaluation finale. Les résultats étaient présentés et discutés en Comité d'évaluation, sous l'égide de la Direction générale. Ce processus pêchait à

plusieurs niveaux : l'AFD n'utilisait pas au mieux son système d'évaluation ; la division Évaluation était isolée du reste de l'institution ; le retour d'informations était faible ; et le Comité d'évaluation était difficile à convoquer.

Début 2006, la division Évaluation a donc été intégrée au Département de la recherche (l'équivalent du Département des connaissances dans certaines institutions), lequel est rattaché à la Direction de la stratégie. Cette structure originale exprime clairement la volonté d'établir un lien entre production de connaissances et évaluation. Bien plus, les principes d'évaluation stipulent qu'une partie des travaux doivent inclure de la recherche en sciences sociales. Cette revitalisation de l'évaluation s'appuie sur quatre principes :

1. partage des évaluations : l'évaluation ne doit pas être une activité restreinte, confiée à un service dédié et n'ayant pour seule vocation que d'informer la direction. Les rapports d'évaluation doivent impliquer tous les acteurs du développement et, en particulier, les départements opérationnels et les partenaires locaux ;

2. synergie entre évaluation et recherche : une partie du travail d'évaluation doit relever d'une recherche appliquée à l'analyse des processus et des résultats des projets et des programmes ;

3. priorité à la capitalisation des expériences et à l'apprentissage institutionnel : les évaluations destinées à nourrir ce processus sont formatives plutôt que sommatives, centrées sur la pratique professionnelle plutôt que sur les seules politiques et les stratégies et elles portent plutôt sur des opérations en cours ;

4. rapprochement graduel (et parfois fusion) des études externes et internes : tandis que l'évaluation externe doit devenir la règle, l'évaluation interne demeure nécessaire, non seulement pour rendre le processus d'apprentissage institutionnel effectif, mais également pour mettre à jour le savoir-faire de la division Évaluation et lui conférer une attractivité professionnelle.

L'AFD vient d'enrichir sa palette d'outils avec les évaluations décentralisées, les évaluations stratégiques, la capitalisation thématique et les évaluations rigoureuses d'impact. Les évaluations décentralisées seront confiées à des experts extérieurs (de préférence locaux), sous le pilotage des départe-

ments géographiques et les agences locales et partagées avec les partenaires locaux. Les évaluations stratégiques continueront d'être confiées à la division Évaluation et pilotées par elle, à l'initiative de la direction et des ministères de tutelle. La division développe également la capitalisation thématique, sur la base d'analyses comparées de groupes d'opérations de développement, finies ou en cours, et financées par l'AFD seule ou en partenariat. Les évaluations rigoureuses d'impact sont réalisées en collaboration avec des équipes universitaires de haut niveau et les partenaires locaux intéressés.

Depuis son intégration au Département de la recherche, la division Évaluation s'est vue confier la responsabilité de développer les évaluations d'impact. Le lien entre production de connaissances et évaluations est en effet décisif pour encourager l'investissement dans ces dernières. La distinction formelle qui sépare souvent les activités de recherche des activités d'évaluation mais aussi l'idée fautive selon laquelle la recherche en sciences sociales et l'évaluation sont deux disciplines indépendantes pourraient bien représenter un sérieux obstacle à l'élaboration d'évaluations d'impact au sein des unités d'évaluation des institutions de développement.

Le programme de l'AFD en matière d'évaluation d'impact sera développé progressivement, ce qui implique un certain investissement intellectuel de ses agents sur ces questions. Globalement, le programme se fixe plusieurs objectifs : le premier est stratégique et vise à produire des connaissances solides sur les facteurs de réussite et d'échec des politiques de développement à l'intention des partenaires du Sud, des responsables des politiques sectorielles de l'AFD et, plus généralement, de la communauté du développement. Le deuxième est méthodologique, l'AFD voulant renforcer les compétences internes pour utiliser les différents instruments de mesure d'impact. Le dernier cherche à encourager les partenariats, au travers d'initiatives conjointes de production de connaissances, avec divers partenaires du Sud et en collaboration active avec des équipes scientifiques.

Enfin, cette expérience progressive de l'AFD sera partagée et enrichie grâce à la participation active à des réseaux internationaux d'évaluation d'impact et à la mise en commun des résultats et des méthodologies.

2. Une palette d'expériences

2.1. Une phase préliminaire

La première évaluation rigoureuse d'impact à l'AFD remonte à 2003, avec le financement d'une recherche sur l'impact des projets de développement agricole financés par différents bailleurs sur les revenus des producteurs guinéens (Delarue, 2007). Confiée à l'AgroParisTech¹, cette évaluation avait pour but de proposer une méthodologie adaptée à la quantification d'impacts, fondée sur une approche qualitative. Menée entièrement *ex post* sur la base d'enquêtes par échantillonnage raisonné, l'évaluation s'est basée sur la collecte de données sur la situation avant-projet et sur des entretiens approfondis avec 100 producteurs. Afin de pouvoir quantifier l'effet net induit par l'un des projets sur le revenu des producteurs, le chercheur a identifié un ensemble de producteurs directement ou indirectement impliqués dans le projet ainsi qu'un groupe témoin fiable. Deux types de projets ont été évalués : le développement de la culture de riz irrigué dans les bas-fonds et la production agroindustrielle de caoutchouc et d'huile de palme (projet public).

La situation avant-projet a été reconstituée en réalisant une première étude en profondeur d'un nombre limité d'unités de production. Celle-ci a permis d'identifier une typologie des systèmes de production en place avant le projet. Dans le but d'établir un scénario avec projet et un scénario contrefactuel, un échantillonnage raisonné a été constitué en choisissant des unités de production appartenant au même type de système de production initial et qui s'étaient développées avec ou sans le projet.

¹ L'AgroParisTech est un établissement de tout premier plan pour les sciences du vivant et l'ingénierie, membre de l'Institut des sciences et technologies de Paris (ParisTech), qui regroupe dix des plus prestigieuses écoles d'ingénieurs françaises.

L'évaluateur, qui avait une bonne connaissance des facteurs endogènes et exogènes ayant influencé l'évolution et les trajectoires potentielles des systèmes de production, a pu identifier précisément les individus aux trajectoires comparables, « avec » et « sans » projet. Son implication directe dans la collecte de données a été essentielle – d'où l'importance des échantillons de petite taille. Un sondage de grande envergure en milieu rural n'aurait pas permis de réunir des données aussi fiables sur les rendements, l'évolution des structures de production au fil du temps et les stratégies des producteurs.

Grâce à cette compréhension fine des projets et des trajectoires des producteurs (qu'ils aient ou non été exposés au projet), un modèle quantitatif fondé sur la méthode de Gittinger d'analyse économique des projets de développement a pu être élaboré (Gittinger, 1982). La diversité des unités de production ayant été clairement identifiée avant l'échantillonnage, un modèle d'évolution a été construit pour chaque type de système de production en place avant le projet. Cette modélisation économique et cette comparaison des trajectoires de développement pour chaque système de production, qu'il ait été exposé ou non au projet, a ainsi permis d'évaluer l'impact du projet sur les revenus des producteurs.

Les différences objectives entre chaque unité de production étudiée auraient pu laisser une certaine place à la subjectivité du chercheur dans la construction de la typologie, de l'échantillon et du modèle. Toutefois, la logique qui sous-tend le concept de systèmes de production a permis de dépasser le risque d'arbitraire. En effet, le saut méthodologique opéré entre un petit nombre d'entretiens et un modèle n'en est plus un quand on sait que, dans la réalité, il existe un nombre fini de types de systèmes de production.

Le premier intérêt de cette nouvelle méthode tient à ce qu'elle a permis de construire une évaluation crédible d'impact entièrement *ex post*. Ensuite, elle a donné une estimation de cet impact sur différents types de systèmes de production, expliquant, ce faisant, la répartition inégale des bénéfices liés aux projets. Enfin, elle a permis de comprendre de façon nuancée pourquoi les impacts attendus se matérialisaient ou non.

Quatre ans de travail de terrain et de traitement des données ont été nécessaires pour obtenir les résultats de cette première évaluation d'impact. Ils ont été présentés aux autorités guinéennes et aux représentants locaux des principaux bailleurs opérant en milieu rural. Sur le terrain, les résultats ont été communiqués aux communautés locales interviewées et aux organisations paysannes. Le ministre de l'Agriculture a affirmé vouloir favoriser davantage d'évaluations d'impact sur des projets de développement agricole. Malheureusement, faute de cadre institutionnalisé pour mener des discussions entre les différents partenaires, il n'y a guère d'espoir que les conclusions de cette recherche modifient la politique nationale relative à ce type de projets.

2.2. La deuxième évaluation d'impact financée par l'AFD

Cette évaluation concernait l'AdéFI, une institution de micro-finance (IMF) au service des micro-entrepreneurs d'Antananarivo (Madagascar). Lancée à la demande de la direction de l'AdéFI et de l'AFD, elle devait produire des données fiables avant de les analyser à l'aide d'une méthode scientifique robuste. L'évaluation d'impact s'est déroulée entre 2003 et 2005, sous l'égide de chercheurs de l'IRD-DIAL, un centre de recherche français (Gubert et Roubaud, 2005).

Initialement, la méthodologie impliquait de comparer la situation d'un échantillon représentatif des micro-entreprises clientes de l'AdéFI avec celle d'un groupe témoin constitué au moyen de méthodes d'appariement sur le score de propension (*propensity score matching* – PSM). Cette première éva-

luation quantitative d'impact était fondée sur un schéma d'évaluation appelé « projet *post-test* et groupes témoins » (*post-test project and comparison groups* ; voir Bamberger et al., 2006). Elle reposait sur 255 interviews datant de 2001 et était alors complétée par une analyse qualitative sur la base d'entretiens ouverts avec un nombre limité de clients de l'AdéFI.

Cette analyse a produit des résultats encourageants concernant la composition du groupe cible (conformément à la théorie du projet) et en termes d'impact (sur la production ou le chiffre d'affaires des entreprises clientes). Toutefois, cette première esquisse n'étant pas suffisamment robuste, une seconde phase a été programmée qui intégrait une méthode de « double différence ». Cette phase, qui s'est conclue par deux enquêtes successives réalisées auprès du même panel en 2003 et en 2004, a permis de recueillir des informations sur les dynamiques à l'œuvre. Elle a également amélioré la qualité des résultats, en permettant l'intégration de nouvelles variables dans le processus d'appariement, rendant ainsi les clients et les non-clients encore plus semblables que dans la première phase.

Malheureusement, il n'a pas toujours été possible de retrouver les entreprises des groupes traitement et des groupes témoins interviewés en 2001 : en 2003, le taux de défection dans ces deux groupes était respectivement de 22 et 23 %. En 2004, 55 % seulement des entreprises du panel d'origine étaient encore en activité. Cette faiblesse du taux de survie traduit la grande fragilité des petites entreprises. Contrairement à ce qui était supposé, les clients de l'AdéFI ont été davantage affectés que les non-clients. Tandis que 255 entreprises avaient été interviewées en 2001, seuls 107 entretiens ont pu être utilisés pour l'analyse du panel en 2004.

D'autres aspects méthodologiques ont été améliorés dans la seconde phase. Les indicateurs sélectionnés pour le calcul du score de propension (probit) comprenaient le sexe des micro-entrepreneurs, leur âge, leur niveau d'éducation, le type d'études, le secteur économique de l'entreprise, le type de locaux pour exercer l'activité, la date de création de l'entrepre-

se, le nombre initial de travailleurs, la valeur initiale du capital, etc. En 2001, ces informations sur la situation initiale avaient été recueillies pour l'année de création de l'entreprise, qui n'était pas la même pour toutes, ce qui constituait un biais. En conséquence, pour les enquêtes de 2003 et 2004, l'année de référence pour ces indicateurs a été fixée pour toutes les entreprises à 1997, pour pouvoir tenir compte des différences dans les caractéristiques des entreprises au moment où l'ADéFI a démarré ses activités.

Plusieurs variables intéressantes en termes de performance économique des micro-entreprises ont pu ainsi être étudiées : le chiffre d'affaires, la production, la valeur ajoutée, le personnel, le capital et, pour finir, la productivité du travail et du capital. L'impact du microcrédit sur ces variables semble positif et statistiquement significatif en 2001 et en 2004. Globalement néanmoins, l'impact mesuré en 2004 semble moindre que celui constaté à l'aide des données de 2001. Cette différence s'explique en partie par la méthodologie utilisée : l'appariement des données a été plus rigoureux en 2004 et les résultats de 2001 sont probablement biaisés.

En effet, du fait que l'appariement de 2001 était fondé sur les caractéristiques des entreprises l'année de leur création, le score de propension ne reposait pas sur des variables de production (le chiffre d'affaires ou la production). En 2004, au contraire, le chiffre d'affaires de 1997 faisait partie des variables du modèle pour prévoir la participation, ce qui a permis un appariement de meilleure qualité. Une simple comparaison entre les deux ensembles de variables utilisés pour calculer le score de propension sur la base des données de 2004 fait apparaître un écart significatif dans les impacts mesurés, même si cet impact reste positif pour toutes les variables. Ce test démontre en même temps l'importance et la difficulté de constituer un groupe témoin adéquat, particulièrement lorsqu'aucune étude de référence n'a été conduite au préalable.

Enfin, le recours à la méthode de double différence entre 2001 et 2004 a donné des résultats très différents de ceux que nous venons d'évoquer, obtenus par des mesures portant sur

une seule période. Dans ce cas, le projet avait un impact positif sur la productivité et sur les différents résultats. En revanche, en appliquant la méthode de double différence, aucun des impacts mesurés n'était significatif. Cela signifie que la trajectoire économique moyenne des clients et des non-clients est identique et que le projet n'a pas réussi à créer une dynamique de croissance pour ses clients.

Cette évaluation d'impact a démontré combien il est difficile de collecter des données de panel sur les clients d'un projet de microfinance. Dans ce cas précis, cela tenait au fort taux d'attrition, étroitement lié à la vulnérabilité des micro-entreprises et à leur propension à changer de lieu, obligeant ainsi les enquêteurs à les suivre à la trace, souvent sans succès. Cette évaluation a également révélé la nature extrêmement sensible de la relation entre les mesures d'impact et la qualité de l'appariement des données.

Cette étude est l'une des rares évaluations d'impact portant sur une institution de microfinance et sur des micro-entrepreneurs. Achevée en 2005, elle est également la première expérience de l'AFD pour la conduite de ce type d'évaluations. La rigueur et la transparence avec lesquelles l'équipe de chercheurs a mené à bien ce travail scientifique a contribué à l'apprentissage institutionnel de l'AFD dans ce domaine. Elle a également encouragé l'AFD à lancer de nouvelles études, pour dépasser les limites méthodologiques de cette première expérience – entraînant ce faisant une augmentation des allocations budgétaires aux évaluations d'impact. Les évaluations suivantes ont été programmées bien avant le démarrage du projet, avec des enquêtes de référence approfondies pour garantir la collecte d'informations utiles sur la situation de départ.

2.3. Une évaluation d'impact expérimentale de la microfinance au Maroc

La première évaluation d'impact expérimentale financée par l'AFD concerne Al Amana, une institution de microfinance qui intervient en zone rurale au Maroc.

Les programmes de microfinance pâtissent encore d'un « déficit d'évaluation » : « *Les opérations de terrain des IMF ont largement dépassé les capacités d'analyse – de sorte que l'enthousiasme que suscite la microfinance dans la lutte contre la pauvreté ne repose sur aucun fait établi à l'issue d'une recherche rigoureuse. Dans l'état actuel des connaissances, il paraît difficile d'allouer les yeux fermés des fonds publics au développement de la microfinance* » (Zeller et Meyer, 2003). En outre, même si la microfinance a fait l'objet d'un bon nombre d'évaluations d'impact ces dix dernières années, celle évoquée ici est l'une des premières à faire appel à une méthode expérimentale, c'est-à-dire avec tirage aléatoire. Enfin, il s'avère particulièrement intéressant de mesurer l'impact du microcrédit dans les zones rurales, car beaucoup d'IMF ont du mal à atteindre ces populations parmi les plus démunies.

Créée en 1995, Al Amana est la première IMF du Maroc, avec quelque 250 000 clients. Jusqu'en 2006, sa clientèle se situait principalement en zones urbaines ou périurbaines (83 % des clients). Aujourd'hui, sa stratégie cible des zones rurales très étendues. Après l'ouverture d'une centaine de succursales faciles d'accès dans l'arrière-pays en 2004 et 2005, Al Amana a décidé d'étendre son action aux régions rurales enclavées. Désireuse d'obtenir une mesure scientifique de l'impact de la diffusion du microcrédit dans ce contexte difficile, la direction d'Al Amana a sollicité l'AFD pour un soutien financier à la réalisation de cette étude. Elle avait déjà sélectionné l'équipe de recherche chargée de l'évaluation (le Poverty Action Lab du MIT) en partenariat avec la toute nouvelle École d'économie de Paris (EEP).

Cette recherche entend analyser les impacts économiques et sociaux du microcrédit dans les régions rurales isolées du

Maroc, à l'aide d'une méthode expérimentale (École d'économie de Paris, 2006). L'assignation aléatoire au groupe expérimental – avec un groupe exposé au microcrédit dès le départ et un autre à une date ultérieure – donnera une estimation claire, transparente et rigoureuse des impacts. Le mode de déploiement des activités d'Al Amana offre un contexte idéal pour utiliser ce type de méthode.

L'évaluation porte sur 80 des 160 succursales qu'Al Amana a prévu d'ouvrir entre 2006 et 2008. Le principe de l'étude, conforme au plan initial, est d'identifier deux petites zones dans la région couverte par une succursale, l'une obtenant rapidement du microcrédit et l'autre un an plus tard. Trois enquêtes devaient être menées : l'enquête établissant les conditions de départ (de référence), l'enquête intermédiaire au bout d'un an et l'enquête finale deux ans plus tard. L'enquête finale devait permettre de mesurer les effets d'une distribution de crédit pendant deux ans par rapport aux effets obtenus au bout d'un an dans le groupe témoin, fournissant ainsi une analyse différentielle des effets de court et de moyen termes du traitement sur les populations.

Le *modus operandi* pour la constitution des deux groupes a été défini après des tests de faisabilité conduits dans neuf localités marocaines. La sélection de villages à partir d'une liste était impossible, du fait de la diversité des environnements ruraux, notamment en termes de régimes fonciers, de cultures, de reliefs et de climat. En outre, il fallait que les douars (villages) du groupe témoin soient suffisamment loin de la succursale de microcrédit et de toute autre source de crédit pour minimiser le risque de contamination. Les villages du groupe expérimental ont donc été sélectionnés de façon raisonnée selon des critères similaires. L'enquête de faisabilité a permis de définir une méthode d'appariement entre un village expérimental et un village témoin ayant les mêmes caractéristiques. Plusieurs variables ont été utilisées, dont la facilité d'accès, la population, les principales cultures pratiquées, etc. Le tirage aléatoire a été finalisé pour chaque succursale d'Al Amana sur une paire de villages similaires, l'un étant assigné par tirage au sort au groupe expérimental et l'autre au groupe témoin.

La faisabilité a beaucoup compté dans la définition des procédures d'expérimentation, notamment pour construire un modèle prédictif de la propension des villageois à souscrire un crédit. Afin de limiter le nombre d'interviews nécessaires pour atteindre la significativité statistique, il fallait impérativement sélectionner le plus grand nombre possible de futurs emprunteurs parmi les ménages interviewés au cours de l'enquête de référence. Pendant l'étude de faisabilité, 2 000 ménages répartis dans les neuf paires de villages ont été interviewés et la distribution de crédit a été suivie pendant les six mois suivants. Cette phase d'observation de la souscription d'un crédit a permis d'établir le modèle prédictif qui devait être proposé.

Le modèle permet donc d'identifier— sur la base d'un bref questionnaire (dix questions) appliqué à 100 ménages dans chaque village— les 25 ménages qui ont la plus forte probabilité de souscrire un crédit auprès d'Al Amana.

Les données collectées dans les neuf premiers sites ont montré que le processus d'évaluation se déroulait correctement. La démarche aléatoire fonctionnait de façon satisfaisante et les écarts de départ entre les ménages du groupe expérimental et du groupe témoin n'étaient pas significatifs. En outre, la collaboration entre Al Amana et l'équipe de chercheurs était exemplaire. Toutefois, quelques problèmes techniques sont apparus qui ont amené à un changement de méthodologie.

Les enquêtes réalisées la première année ont révélé que le nombre d'emprunteurs était plus faible que ce qui était prévu en théorie dans les régions retenues où Al Amana n'avait pas eu d'expérience préalable. Pour les premières succursales couvertes dans l'étude de faisabilité, le taux d'emprunt de l'ensemble des ménages était de 21 % seulement après 14 mois (36 % chez les ménages du groupe expérimental, sélectionnés grâce au modèle de propension). Dans les 23 succursales incluses par la suite dans l'étude (représentant 39 douars dans le groupe expérimental), la souscription moyenne de crédit était de seulement 7 % après cinq mois, avec des variations énormes entre les villages (de 0 à 55 %).

La faiblesse du nombre d'emprunteurs dans le groupe expérimental interviewé a compromis la possibilité de mesurer un impact limité de façon statistiquement significative. Avec un taux d'emprunt de 20 %, il serait par exemple impossible de détecter un changement de consommation inférieur à 21 %. Pour résoudre cette menace sur la bonne fin de l'évaluation, plusieurs mesures ont été adoptées afin de sensibiliser davantage les villageois des douars appartenant au groupe expérimental : multiplication des réunions d'information ; ouverture à tous des quotas de crédit réservés aux femmes ; et incitations du personnel d'Al Amana à s'occuper de ces villages isolés.

En outre, le protocole a dû être profondément révisé. Ces difficultés indiquaient qu'une période d'exclusion d'un an ne serait pas suffisamment longue pour obtenir une évolution significativement différente entre le groupe expérimental et le groupe témoin. Décision a alors été prise d'étendre la période d'exclusion du groupe témoin à deux ans, ce qui n'a pas été sans conséquences. Pour l'équipe locale d'Al Amana, cela signifiait qu'il fallait expliquer à la population du village témoin que l'accès au crédit avait été retardé. Cela signifiait aussi qu'Al Amana aurait moins de clients, ce qui entraînait un manque à gagner pour l'organisme. L'enquête à mi-parcours a été annulée et son budget a été réaffecté de façon à inclure 20 villages supplémentaires et d'avoir ainsi de plus grandes chances d'atteindre la significativité statistique dans la dernière étape.

Ces deux ajustements de fond ont été possibles grâce à la coopération remarquable entre Al Amana, l'EEP et l'AFD pour surmonter les difficultés rencontrées et garantir le bon achèvement de l'évaluation. Les parties se rencontrent régulièrement et écartent ainsi tout risque de malentendus. Le protocole de recherche est parfaitement transparent pour toutes les parties prenantes.

Les résultats seront disponibles en 2010.

2.4. Une évaluation d'impact expérimentale de la micro-assurance de santé au Cambodge

L'assurance de santé fait partie des grands enjeux auxquels sont actuellement confrontés les pays en développement. La France s'est récemment engagée à accroître ses investissements dans la protection sociale de ces pays. L'AFD, dont l'expérience en matière de financement de programmes de santé est encore relativement limitée, soutient pour l'instant deux programmes de micro-assurance de santé, l'un au Cambodge, l'autre au Laos.

Avant d'augmenter ses investissements dans ce secteur, l'AFD a voulu conforter ses connaissances dans ce domaine et a décidé d'évaluer l'impact du programme d'assurance de santé SKY qu'elle finance au Cambodge. Lancé en 1998 par le Groupe de recherche et d'échanges technologiques (GRET), SKY offre aux ménages des soins de base et d'urgence entièrement gratuits et sans limite dans des dispensaires ainsi qu'un certain nombre d'autres services de santé, moyennant une prime mensuelle fixe. SKY entend permettre ainsi aux familles d'accéder aux soins sans risque de se ruiner.

En 2005, l'AFD a signé un protocole d'accord pour mener des évaluations de projets avec Scientific Evaluation and Global Action (SEGA), un centre de recherche de l'université de Californie (Berkeley et San Francisco). À la suite d'une première proposition méthodologique de SEGA pour l'évaluation d'impact du projet SKY, une mission d'identification a eu lieu fin 2006. Il s'agissait de définir plus précisément la méthodologie et le champ de l'évaluation mais aussi d'engager un processus d'appropriation des résultats attendus par les décideurs cambodgiens.

Un tirage aléatoire est au cœur de l'évaluation d'impact du projet SKY (Levine et al., 2007). Il sera mis en œuvre parallèlement au déploiement de SKY dans la province de Takéo, dont le démarrage est prévu à la fin du premier semestre 2008. Dans le plan d'étude retenu, l'outil méthodologique prin-

cipal consiste à distribuer aléatoirement des bons de réductions de la prime d'assurance de façon à faire varier la probabilité de souscription d'une assurance chez les ménages d'un village et à isoler l'impact de l'assurance de santé sur les résultats que l'on veut observer.

À la suite d'une première réunion de village, qui permettra de distribuer au hasard les bons de réduction, l'enquête de référence sera administrée à un sous-ensemble de ménages tirés au sort, stratifié selon la valeur du bon de réduction. Grâce aux données fournies par l'enquête de référence et aux dossiers de SKY sur les ménages ayant choisi de souscrire une assurance, le processus d'auto-sélection dans un programme d'assurance pourra être clarifié. On connaîtra ainsi les caractéristiques susceptibles d'inciter les ménages à souscrire une assurance de santé. En outre, les bons de réduction étant attribués de façon aléatoire, la manière dont le montant de la prime est corrélé avec les caractéristiques de départ des ménages assurés et non assurés pourra être établie.

Douze mois après les réunions de village, des enquêtes de suivi de tous les ménages interviewés seront organisées. Ces données ainsi que celles de l'enquête de référence fourniront des informations sur l'impact de SKY sur le recours aux soins et sur l'utilisation du système de santé. Elles devraient également préciser l'impact de l'assurance de santé sur les résultats économiques des ménages. Un an plus tard, une seconde enquête de suivi reprendra la plupart de ces sujets, en mettant à nouveau l'accent sur les évolutions en termes de résultats et de dépenses de santé.

Les effets à plus long terme de l'assurance sont aussi très intéressants, surtout au niveau du taux de défection des ménages ayant reçu des bons de réduction d'une valeur élevée. Si les tests pilotes montrent que la plupart des personnes ayant gagné un bon de réduction dans la première période renouvellent leur assurance pour une période additionnelle de six mois, alors le schéma qui précède suffira pour étudier les effets pendant au moins les 12 premiers mois.

Toutefois, comme dans tout projet, l'évaluation pourrait ne pas évoluer comme prévu. Le risque principal réside dans le nombre de ménages qui ne renouvellent pas leur adhésion à SKY après la période initiale de six mois. Actuellement, le taux de défection après ce délai de six mois est d'environ 17 % (sur la base des statistiques de SKY). Si le taux de défection dépasse nettement ce seuil chez les souscripteurs ayant reçu un coupon élevé, l'écart entre la situation des assurés du groupe expérimental initial (bon de valeur élevée) et celle du groupe témoin (bon de valeur faible) pourrait bien devenir minime. Il faudra alors probablement appliquer une méthode d'appariement quasi-expérimentale pour mesurer l'impact de SKY.

Outre les enquêtes auprès des ménages, l'équipe de chercheurs effectuera également une évaluation qualitative de SKY. Cette analyse examinera son impact sur le système de santé, y compris sur les recettes des établissements publics de santé, sur les changements dans la mise à disposition de médicaments et d'équipement médical et sur l'évolution des revenus et des comportements professionnels des personnels soignants.

En résumé, le traitement aléatoire devrait permettre aux chercheurs d'estimer de façon fiable les impacts de l'assurance de santé, en les isolant des autres caractéristiques variant entre les ménages assurés et les non-assurés. Une enquête de référence pré-intervention portant sur environ 3 000 ménages et plus de 15 000 personnes ainsi que des enquêtes de suivi des mêmes ménages seront menées sur une période expérimentale de quatre ans. L'enquête couvrira les multiples aspects que le programme tente de modifier : état de santé, recours aux soins, vulnérabilité des actifs, décisions d'investissement et d'épargne et gestion du risque. Grâce au plan de recherche aléatoire, l'évolution des résultats dans le temps des ménages assurés et des ménages non assurés deviendra comparable, ce qui permettra d'estimer les effets de causalité de l'assurance de santé. Cette évaluation d'impact se caractérise par les différents partenariats noués au Cambodge et au niveau international, de façon à assurer qu'elle réponde aux besoins du programme, des institutions qui la financent (l'AFD

notamment) et d'autres partenaires œuvrant à la fourniture de soins de santé aux pauvres.

Tout au long de la rédaction de la proposition méthodologique, SEGA a travaillé en relation étroite avec l'équipe de l'AFD et du GRET ainsi qu'au Cambodge, avec les chercheurs partenaires de Domrei Research and Consulting. L'apport du GRET a été crucial à tous les niveaux de la proposition et notamment pour déterminer la faisabilité et la pertinence des projets de recherche. Ces relations sont en train de se développer avec la structuration et la mise en œuvre du projet de recherche et l'introduction des instruments d'enquête.

En outre, SEGA a reçu une subvention BASIS de l'USAID, qui lui permettra de développer des activités de renforcement des capacités de chercheurs et de praticiens cambodgiens. Ces activités impliquent un partenariat entre les chercheurs de l'université royale de Phnom Penh (RUPP). Tout au long de l'évaluation, SEGA offrira aux étudiants de la RUPP une formation à la conception de programmes d'évaluation et à la méthodologie associée, afin que les évaluations d'impact soient, à l'avenir, dirigées localement. En outre, une partie du budget alloué par l'AFD prévoit des réunions à Paris, pour diffuser la méthodologie (plan de recherche et techniques économétriques notamment) et les résultats de SEGA.

La première mission au Cambodge a permis de confirmer que l'évaluation du programme SKY venait directement étayer les objectifs du ministère cambodgien de la Santé. Les informations relatives à l'efficacité de SKY lui permettront en particulier de structurer la réforme du système national de santé, engagée depuis 1999 (OIT, 2005). Au cours de cette mission, l'équipe d'évaluation et l'AFD ont rencontré des représentants du ministère de la Santé pour identifier les problèmes particulièrement importants à aborder dans l'évaluation de SKY. La contribution des représentants du ministère de la Santé à la conception de l'étude sera sollicitée tout au long de cette évaluation ; en retour, ils seront régulièrement informés des résultats, qui seront disponibles en 2010.

3. Premiers enseignements

3.1. L'apprentissage institutionnel

Pour l'AFD, toutes ces expériences ont été l'occasion d'un apprentissage progressif. Elles lui ont aussi permis de modifier son approche des évaluations d'impact, l'amenant en particulier à adopter des méthodes qui imposent la réalisation d'enquêtes de référence avant le démarrage du projet. Il a fallu pour ce faire convaincre les différents partenaires de l'utilité de ces enquêtes et identifier les coéquipiers universitaires très en amont du projet. La préparation d'une telle recherche demande évidemment beaucoup de temps, puisqu'il faut réunir décideurs politiques, chercheurs, enquêteurs, promoteurs du développement et bailleurs de fonds pour trouver un accord sur les principes et les modalités de l'exercice. Dans les cas marocain et cambodgien, la phase de préparation a duré un an et demi.

L'implication des partenaires locaux s'est elle aussi faite progressivement. De toute évidence, les évaluations d'impact sont un exercice exigeant qui requiert l'implication totale de l'équipe du projet évalué. Sa contribution est essentielle pour adapter les questionnaires au contexte et ajuster la taille de l'échantillon à la participation attendue. Elle est particulièrement décisive pendant la phase de mise en œuvre de l'évaluation d'impact, pour s'assurer que le plan est correctement appliqué, surtout lorsqu'il s'agit d'éviter la contamination du groupe témoin. Cela implique aussi en général que le projet et ses différents partenaires soient prêts à modifier le protocole d'intervention pour permettre un traitement aléatoire ou pour construire un groupe de référence de qualité.

Toutefois, l'objectif stratégique poursuivi par les évaluations d'impact – contribuer à l'élaboration des politiques – va au-

delà de la seule implication de l'équipe de projet : les résultats rigoureusement obtenus ne servent à rien si les décideurs nationaux n'en tiennent pas compte. Des efforts particuliers ont été consentis dans le cas de la Guinée et du Cambodge afin d'obtenir, du début à la fin du processus, l'adhésion des politiques.

Les expériences en cours indiquent que les évaluations d'impact sont autant un outil de recherche qu'un outil d'évaluation. Si les résultats attendus en retour s'apparentent davantage à l'évaluation, la nature de l'investigation est celle d'une recherche scientifique. Pour l'AFD, ces évaluations d'impact ouvrent de nouvelles perspectives de collaboration avec des partenaires universitaires de haut niveau, en vue d'analyser ses opérations. Le grand intérêt suscité en interne par chacune de ces évaluations, chez les gestionnaires de projets comme dans la division Évaluation, est en train de façonner une culture de l'évaluation propice au lancement de nouvelles évaluations « scientifiques ».

3.2. L'évaluation d'impact pour un bailleur de fonds bilatéral

L'AFD ne peut se targuer de lointains antécédents en matière d'évaluation d'impact : les évaluations réalisées en 2008 (Cambodge et Maroc) ne portent que sur deux opérations, alors qu'elle en finance plus de 500 ! Toutefois, ces deux projets vont absorber 25 % du budget d'évaluation externe de la division Évaluation et entre 10 et 15 % de ses ressources humaines. Pourquoi une telle disproportion ?

Tout bailleur de fonds bilatéral doit d'abord s'interroger sur l'intérêt d'une évaluation d'impact. Ensuite, il doit décider s'il internalise ou externalise le pilotage de ce travail. S'il retient l'option interne, alors il doit réfléchir au lieu de rattachement idoine d'une telle activité : la division chargée de l'évaluation, le département en charge des connaissances ou le département responsable de l'élaboration des politiques ?

L'AFD s'est attelée au problème en confiant, nous l'avons vu, plusieurs évaluations d'impact à sa division Évaluation. L'évaluation d'impact a trouvé tout naturellement sa place au sein du Département de la recherche, qui a récupéré la fonction d'évaluation.

Si l'on en juge par la brève expérience de l'AFD, l'implication interne dans le pilotage d'une évaluation d'impact semble pouvoir nourrir efficacement le débat institutionnel sur les mesures d'impact. En outre, l'adoption d'une gestion axée sur les résultats ne peut être envisagée sans avoir connu les implications d'une évaluation rigoureuse d'impact.

Une implication directe dans la gestion de l'évaluation d'impact est décisive pour bien comprendre les règles, les difficultés objectives et les défis posés par ce type d'exercice – ce que les exemples évoqués ici ont parfaitement montré. Le débat autour de l'internalisation ou de l'externalisation à une équipe dédiée pourrait bien être stérile : l'appropriation de l'évaluation d'impact est une condition préalable pour une participation fructueuse à des réseaux internationaux, où l'on partage non seulement les résultats, mais aussi les processus et les méthodes.

3.3. Perspectives

Notre analyse des évaluations d'impact expérimentales lancées jusqu'ici montre qu'en dépit de leur intérêt, ces essais présentent quelques limites dès lors qu'ils concernent des projets du type de ceux que l'AFD finance. Des difficultés récur-

rentes apparaissent en effet avec l'application d'une méthodologie quantitative (associée à une étude de référence) aux projets de l'AFD : l'autosélection des bénéficiaires et la lenteur avec laquelle les populations souscrivent aux services spécifiques (microcrédit ou micro-assurance) lorsque ceux-ci leur sont proposés pour la première fois. En outre, il y a contradiction entre le maintien d'un groupe témoin en dehors de toute contamination et la mesure des impacts d'un nouveau projet – qui peuvent être longs à se matérialiser – sur des individus ou sur une population dans son ensemble.

Les résultats des évaluations d'impact expérimentales lancés récemment ne seront pas disponibles avant plusieurs années. Chaque expérience a requis, avant son démarrage, une phase d'identification de 18 mois. Pour plusieurs raisons – leur coût, le temps à consacrer à la préparation et à la mise en œuvre, les risques objectifs évoqués dans les plans d'évaluation et le fait que l'on ignore l'impact qu'ils peuvent avoir sur les décideurs politiques – l'AFD a décidé d'attendre les résultats avant de s'engager dans d'autres exercices de ce type.

L'AFD tient aussi à pouvoir identifier des méthodes susceptibles d'apporter des réponses à un maximum de questions utiles pour les politiques. Dans ce but, les évaluations en Guinée et au Cambodge ont fait ou feront appel à un large éventail de méthodes : des données relatives aux processus et aux impacts sur les principales parties prenantes autres que les clients directs du projet seront collectées et analysées. Mais un nouveau problème se pose : comment trouver des chercheurs spontanément prêts à utiliser toute la gamme des méthodes indispensable à l'évaluation d'impact ? Il s'avère nécessaire d'affiner la méthode principale utilisée. Les gestionnaires de projets et l'AFD doivent s'assurer que la recherche de réponses à un maximum de questions pertinentes passe avant les choix méthodologiques.

De fait, la division Évaluation de l'AFD s'est attelée à la conception de démarches innovantes pour mettre au point une méthodologie rigoureuse d'évaluation d'impact, en utilisant des raisonnements contrefactuels et en quantifiant le change-

ment net dans les résultats sans faire nécessairement appel à des techniques expérimentales ou quasi expérimentales. Il faut tout un arsenal de méthodes scientifiques d'évaluation d'impact pour apporter des réponses à l'ensemble des questions soulevées par les interventions de développement, sans pour autant négliger la question de l'attribution. Comme l'a montré l'évaluation d'impact en Guinée, la division Évaluation de l'AFD estime que la recherche qualitative peut répondre de façon rigoureuse aux défis posés par les évaluations d'impact.

Face aux demandes internes croissantes en ce domaine, l'AFD prévoit d'adopter une approche pragmatique, qui non seulement favorise la généralisation d'évaluations de qualité, mais intègre également, quand c'est possible, la question de l'impact. Si les gestionnaires de projets sont encore réticents à financer des études de référence et la collecte de données spécifiques, ils le sont encore plus dès lors qu'il faut s'engager dans un processus de long terme et allouer un budget conséquent à un programme de recherche sur l'impact de leurs opérations. On voit par là combien il est important de réunir davantage d'informations sur les bénéfices réels des évaluations d'impact avant de pouvoir les développer à grande échelle.

Références

AFD (2007), *2007-2011 Strategic Plan of the Agence Française de Développement*, Paris.

Bamberger, M., Rugh, J., Mabry, L. (2006), *Real World Evaluation : working under budget, time, data and political constraints*, Sage publications.

CGD (2006), *When will we ever learn? Improving lives through impact evaluations*, Report of the Evaluation Gap Working Group.

Delarue, J. (2007), *Mise au point d'une méthode d'évaluation systématique de l'impact des projets de développement agricole sur le revenu des producteurs, Etude de cas en région kpèlè (République de Guinée)*, PhD thesis, AGROPARISTECH, Paris.

Gittinger, J.P. (1982), *Economic Analysis of Agricultural Projects (EDI Series in Economic Development)*, John Hopkins University Press, Baltimore.

Gubert, F. and F. Roubaud (2005), *Analyser l'impact d'un projet de microfinance : l'exemple d'ADéFI à Madagascar*, Notes et Documents, N°. 19, AFD, Paris.

ILO (2005), *Informal Economy Poverty, Employment in Cambodia, Mongolia, Thailand: The First Eighteen Months – A Progress Review*, ILO, Geneva. www.ilo.org/public/english/bureau/integration/download/publicat/4_3_370_informal_economy_final_report_-_with_cover_sheet.pdf

Levine, D. I., R. Polimeni and R. Arunachalam (2007), *A randomized controlled trial of micro-health insurance in Cambodia, A preliminary proposal to AFD, SEGA, UC Berkeley, Berkeley*.

Paris School of Economics (2006), *Evaluation de l'impact d'un programme de microcrédit en milieu rural*, Research project, Paris School of Economics, Paris.

Paris School of Economics, (2007), *Evaluation de l'impact du microcrédit en milieu rural*, Note written for the Steering Committee 1st August 2007, Paris School of Economics, Paris.

Zeller M. and R.L. Meyer (dir.pub.) (2003), *The triangle of microfinance: financial sustainability, outreach and impact*, John Hopkins University Press, Baltimore.