



ÉVALUATION DES POLITIQUES SOCIALES

Du 10 au 14 septembre 2012 à Ouagadougou

Formation J-PAL Europe



© Aude Guerrucci

SOMMAIRE

1. Lettre de bienvenue	3
2. Présentation de J-PAL	5
3. Planning des cours	7
4. Informations pour le dîner du mardi	9
5. Présentation des professeurs	11
6. Présentation des assistants	13
7. Liste des participants	15
8. Instructions pour le travail en groupe	17
9. Groupes de travail	19
10. Étude de cas 1 – Mesurer l’impact	21
11. Étude de cas 2 – Lire les évaluations	27
12. Étude de cas 3 – Comment randomiser ?	35
13. Étude de cas 4 – Menaces à la validité	39
14. Exercice 1 – Mécanismes de l’assignation aléatoire	47
15. Exercice 2 – Taille de l’échantillon	55
16. Exercice 3 – Calcul de la taille de l’échantillon	57
17. Bibliographie	63
18. Liste de contrôle pour les évaluations aléatoires	65
19. Glossaire	75
20. Contacts J-PAL et IPA	81

BIENVENUE !

Chaque année, les membres du J-PAL (Jameel Poverty Action Lab) forment plusieurs dizaines de participants à l'évaluation aléatoire. Cette année ont déjà eu lieu des formations au Cap Town (Afrique du Sud), à Accra (Ghana), à Cambridge (USA), à Paris (France), à Rio de Janeiro (Brésil) et à New Delhi (Inde).

J-PAL Europe dispense ses cours en alternance d'une année sur l'autre en anglais et en français. Notre souci est de nous ouvrir à la diversité des publics francophones, et cette première formation que nous organisons en français en Afrique sub-saharienne revêt pour nous une importance particulière.

Parmi vous se trouvent des personnes issues d'institutions publiques, d'instituts de recherche, d'organisations internationales ou d'ONG. Vous venez majoritairement du Burkina Faso, mais également d'autres pays de la région ou d'Europe.

Certains d'entre vous ont pratiqué des projets de recherche pendant des années, d'autres ont des années d'expérience de travail opérationnel sur le terrain. Certains d'entre vous sont des experts en statistique, d'autres sont des experts en santé, en éducation ou en agriculture. Nos discussions pendant la semaine aborderont tous ces domaines et enrichiront notre réflexion.

Vous avez l'occasion de créer des liens avec d'autres professionnels de l'évaluation dans différents secteurs et différents contextes. Nous espérons que vous trouverez des intérêts communs et que de nouveaux partenariats et projets se noueront au cours de la semaine.

Avec cette formation, nous souhaitons vous préparer à faire vos propres évaluations aléatoires. Nous restons ensuite à votre disposition pour vous apporter aide et conseils, que ce soit pour des questions sur la taille de l'échantillon ou pour vous aider à réfléchir sur l'évaluation d'un programme particulier. N'hésitez pas à nous contacter !

Nous vous souhaitons une semaine intéressante et productive !



Hélène Giacobino
Directrice
J-PAL Europe

J-PAL : DES CHERCHEURS CONTRE LA PAUVRETÉ

Fondé en 2003 au MIT (Massachusetts Institute of Technology), J-PAL (Jameel-Poverty Action Lab) est un **réseau de chercheurs** qui compte désormais quatre antennes régionales : **J-PAL Europe** situé à Paris, en France, **J-PAL Asie du Sud**, à Chennai, en Inde, **J-PAL Amérique Latine** à Santiago, au Chili et **J-PAL Afrique du Sud** à Cape Town.

J-PAL est un réseau de 53 chercheurs du monde entier. Depuis 2003, ils conduisent plus de 250 évaluations aléatoires de programmes contre la pauvreté, dont une centaine est déjà terminée.

Notre postulat: l'efficacité des programmes sociaux serait nettement améliorée si les décisions politiques étaient fondées sur des preuves scientifiques rigoureuses.

Notre méthode: les évaluations aléatoires.



QUE FAIT J-PAL ?

J-PAL a trois objectifs principaux :

1. **Évaluer des programmes de développement** : Les chercheurs de J-PAL ont une **démarche expérimentale** comparable à celle des essais clinique en médecine: on choisit un groupe d'individus pris au hasard qui va participer à un programme, et on confronte ensuite les résultats obtenus avec ceux d'un groupe témoin, n'y ayant pas pris part. Ces enquêtes avec sélection aléatoire connaissent un véritable boom ces dernières années en économie du développement.

Ce procédé permet notamment de **mesurer rigoureusement** des paramètres jusqu'alors difficilement quantifiables. Ainsi, les chercheurs évaluent les effets de projets très divers: la micro-finance dans les zones rurales du Maroc, la diffusion de nouvelles techniques agricoles au Kenya, la lutte contre la corruption en Indonésie, les effets de la discrimination positive en Inde, mais également l'impact de la formation professionnelle en France.

2. **Développer les compétences** : J-PAL a formé **plusieurs centaines de responsables** d'une cinquantaine de pays différents, grâce à des cours dispensés notamment aux États-Unis, en

France, en Inde ou en Amérique Latine. Depuis, plusieurs de ces participants ont mené d'importantes évaluations aléatoires. J-PAL encourage et soutient ces initiatives.

3. **Diffuser les résultats obtenus aux décideurs publics et privés** : Grâce à ses nombreuses publications et à sa participation intensive à des séminaires et conférences, J-PAL diffuse les évaluations réalisées. Désormais les gouvernements de plusieurs pays utilisent ces résultats pour élaborer leurs décisions en matière de politique sociale. Ainsi, en matière de déparasitage une importante évaluation a été menée au Kenya: les résultats extrêmement positifs ont été mis à profit dans plusieurs pays.



QUI FINANCE J-PAL ?

Dans le monde entier, J-PAL coopère de différentes façons avec des agences publiques de développement, des organisations internationales, des ONG et des groupes du secteur privé, en évaluant des programmes : par exemple, avec l'Agence Française du Développement, le Haut Commissaire aux Solidarités actives contre la pauvreté, la Banque Mondiale, l'Institut Veolia Environnement, la National Science Foundation, le Millenium Challenge Corporation. Ces institutions financent également J-PAL.

La John D. and Catherine T. MacArthur Foundation, la Bill and Melinda Gates Foundation, la William and Flora Hewlett Foundation et la Nike Foundation font notamment également partie des **donateurs** soutenant J-PAL.

En 2005, le Laboratoire d'Action contre la Pauvreté (PAL) est devenu **J-PAL** en l'honneur de Monsieur Abdul Latif **Jameel**, à la suite de la donation importante faite par son fils, ancien élève du MIT et militant généreux de toutes les initiatives de lutte contre la pauvreté du monde entier.

PLANNING DES COURS

Lundi 10 septembre

09h00 – 09h30	<i>Accueil des participants / petit-déjeuner</i>	
09h30 – 11h00	Cours 1: Pourquoi évaluer ?	Harounan Kazianga
11h00– 12h30	Étude de cas 1 : Femmes politiques en Inde	
12h30– 13h30	<i>Déjeuner</i>	
13h30– 15h00	Cours 2 : Comment mesurer l'impact ?	William Parienté
15h00– 15h30	<i>Pause-café</i>	
15h30– 18h00	Travail de groupe	

Mardi 11 septembre

09h00 – 10h30	Étude de cas 2 : Apprendre à lire en Inde	
10h30 – 11h00	<i>Pause-café</i>	
11h00– 12h30	Cours 3: Pourquoi tirer au sort ?	Florencia Devoto
12h30– 13h30	<i>Déjeuner</i>	
13h30– 15h00	Cours 4 : Comment tirer au sort ?	William Parienté
15h00– 15h30	<i>Pause-café</i>	
15h30 – 16h30	Étude de cas 3 : Enseignants supplémentaires au Kenya	
16h30– 17h00	Exercice 1 : Techniques de tirage au sort Exercice 2 : La taille de l'échantillon	
17h00 – 18h00	Travail de groupe	
19h30 – 21h30	<i>Dîner de groupe</i>	

Mercredi 12 septembre

09h00 – 11h00	Cours 5 : Comment déterminer la taille de l'échantillon ?	Bruno Crépon
11h00 – 11h30	<i>Pause-café</i>	
11h30 – 12h30	Exercice 3 : Calcul de puissance	
12h30 – 13h30	<i>Déjeuner</i>	
13h30 – 15h00	Cours 6 : Risques et solutions	Loïc Watine
15h00 – 15h30	<i>Pause-café</i>	
15h30 – 18h00	Travail de groupe	

Jeudi 13 septembre

09h00 – 10h30	Étude de cas 4 : Déparasitage au Kenya	
10h30 – 11h00	<i>Pause-café</i>	
11h00 – 12h30	Cours 7 : Une évaluation de A à Z	Marc Gurgand
12h30 – 13h30	<i>Déjeuner</i>	
13h30 – 15h00	Cours 8 : Analyses, coût/efficacité et généralisations	Hélène Giacobino
15h00 – 15h30	<i>Pause-café</i>	
15h30 – 18h00	Travail de groupe	

Vendredi 14 septembre

09h00 – 10h30	Présentations de groupe	
10h30 – 11h00	<i>Pause-café</i>	
11h00 – 12h30	Présentations de groupe	
12h30 – 13h30	<i>Déjeuner</i>	
13h30 – 15h00	Présentations de groupe	
15h00 – 15h30	Conclusion de la formation	

DÎNER DU MARDI

Nous avons le plaisir de vous inviter pour dîner tous ensemble le mardi 11 septembre à 19h30.

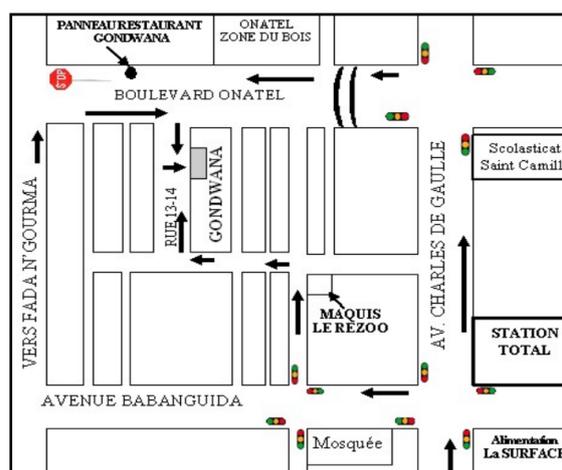
Pour vous rendre au restaurant :

Espace Gondwana

Rue du Docteur Balla Moussa Traoré
Ouagadougou

Tél. : +226 50 36 11 24

<http://www.africartisanat.com>



PRÉSENTATION DES PROFESSEURS



Harounan KAZIANGA est économiste du développement et actuellement Professeur-assistant au Département économique de la Spears School of Business de l'Université d'Oklahoma. Il a occupé des postes de post-doctorant à l'Université de Columbia et à Yale, et a été consultant pour la Banque Mondiale. Ses domaines d'intérêts sont l'agriculture et le développement humain en Afrique de l'Ouest.

harounan.kazianga@okstate.edu



Bruno CRÉPON Bruno Crépon est chercheur au Centre de Recherche en Économie et Statistiques (CREST) et un professeur associé à l'ENSAE et à l'École Polytechnique. Ses recherches portent sur les politiques publiques d'évaluation, avec une attention particulière sur les politiques publiques du marché du travail.

bruno.crepon@ensae.fr



Florencia DEVOTO est doctorante. Elle a rejoint J-PAL en 2006 et participe actuellement à trois évaluations au Maroc dans les domaines du microcrédit, de l'accès à l'eau et de l'éducation.

fdevoto@povertyactionlab.org



Hélène GIACOBINO est diplômée en Droit et a une Maîtrise d'Histoire et de Cinéma. Après une formation en tant qu'architecte et une autre en tant que psychologue, elle est devenue associée dans un cabinet d'Avocats pendant 15 ans. Elle a rejoint J-PAL Europe en 2009 en tant que Directrice stratégie et développement. Elle dirige désormais le bureau de Paris.

hgiacobino@povertyactionlab.org



Marc GURGAND est professeur associé à l'École d'économie de Paris, affilié et Directeur scientifique de J-PAL Europe. Ses recherches portent sur les politiques publiques du marché du travail, la scolarisation et les inégalités. Il mène actuellement une évaluation portant sur les systèmes d'orientation professionnelle pour les chômeurs et les bénéficiaires d'aide sociale. Il travaille aussi sur un programme étudiant les inégalités en Chine.

gurgand@pse.ens.fr



William PARIENTÉ a rejoint J-PAL Europe en 2006. Ses recherches actuelles portent sur l'accès au crédit, la pauvreté, et les enjeux de santé. Il travaille actuellement sur plusieurs évaluations aléatoires au Maroc, au Pakistan et en France.

william.pariente@uclouvain.be



Loïc WATINE travaille pour Innovations for Poverty Action (IPA), organisation-sœur de J-PAL, en tant que directeur-pays au Mali. Dans cette position qu'il occupe depuis 2010, il coordonne les activités liées à une dizaine d'évaluation aléatoires, dans les domaines de l'éducation, de la santé, et de la microfinance, et de l'agriculture. Il couvre également le Sénégal et le Burkina Faso. Il est ingénieur diplômé de l'École Polytechnique et a obtenu un master en administration publique et développement international de l'Université de Harvard.

lwatine@poverty-action.org

PRÉSENTATION DES ASSISTANTS



Emmanuel BAKIRDJIAN a rejoint J-PAL Afrique en juillet 2012. Il travaille actuellement sur trois évaluations d'impact focalisées sur le marché du travail en Afrique du Sud et sur le chômage des jeunes. Il participe également aux activités de renforcement de capacités sur le continent africain, en particulier dans les pays francophones.

ebakirdjian@povertyactionlab.org



Adrien BOUGUEN a rejoint J-PAL Europe en 2009. Il travaille actuellement sur l'évaluation d'un programme d'éducation mettant en place des cours de rattrapage pour les élèves de CP et sur un programme qui tente de résoudre le problème d'orientation des élèves de 3ème en les informant sur les possibilités de formation.

abouguen@povertyactionlab.org



Bastien MICHEL a rejoint J-PAL en 2008. Il travaille sur 3 projets en Inde : l'un vise à lutter contre les carences en fer en mettant du sel fortifié à la disposition des ménages, le second a pour but d'améliorer la qualité des repas fournis aux élèves d'école primaire, et le dernier vise à sensibiliser les populations au risque d'hypertension et à augmenter le recours à la médecine préventive par la population à risque.

bmichel@povertyactionlab.org



Élise PÉSONEL a rejoint J-PAL en 2011 et travaille actuellement sur deux évaluations aléatoires en France. L'une vise à mesurer l'impact d'un programme d'accompagnement sur l'insertion professionnelle d'étudiants boursiers, l'autre évalue les effets sur les choix d'orientation de lycéens d'une aide dans la construction de leur projet professionnel.

epesonel@povertyactionlab.org



Victor POULIQUEN a rejoint J-PAL en 2008 et travaille sur des évaluations dans les domaines de l'éducation et de la santé : un programme de transferts monétaires conditionnels dans l'éducation primaire au Maroc, un programme de bourses scolaires pour l'accès à l'éducation au Ghana, et différentes politiques de luttres contre le Sida au Kenya.

vpouliquen@povertyactionlab.org



Juliette SEBAN a rejoint J-PAL en 2008. Elle travaille actuellement sur les évaluations aléatoires de deux programmes sur l'insertion professionnelle des jeunes en France et l'entrepreneuriat, ainsi que sur l'impact de différents programmes de prévention HIV/SIDA au Cameroun.

jseban@povertyactionlab.org

LISTE DES PARTICIPANTS

Nom	Organisation	Pays
AUBIN Stéphanie	Southern African Confederation of Agricultural Unions (SACAU)	AFRIQUE DU SUD
BA Ousseynou	Millenium Challenge Account Sénégal	SÉNÉGAL
BACYE Z. François	Millenium Challenge Account Burkina Faso	BURKINA FASO
BAKO Dramane	Direction générale de la promotion de l'économie rurale (DGPER)	BURKINA FASO
BALIMA Mahamadi	Institut national de la statistique et de la démographie	BURKINA FASO
BATIONO Rakissiwindé	Conseils Burkinabé des chargeurs (CBC)	BURKINA FASO
BHAKKAN-MAMBIR Nadège	CEREGMIA	FRANCE
BOCK Rosilin	GIZ-Coopération Allemande à Mwangaza Action	BURKINA FASO
COMBARY Omer	Université Ouaga II	BURKINA FASO
COULIBALY Siriki Zanga Timgana	Afristat	MALI
DAHI Abdel Aziz	Centre national d'assurance maladie	MAURITANIE
DAO Seydou	Centre d'études et de recherche sur le développement (CERDI)	BURKINA FASO
DESRUMAUX Aurélie	Équilibres & Populations	FRANCE
GABAIX-HIALÉ Églantine	J-PAL Europe	FRANCE
GUISSOU S. Richard	Direction générale de la promotion de l'économie rurale (DGPER)	BURKINA FASO
ILBOUDO Francine	Direction générale de la promotion de l'économie rurale (DGPER)	BURKINA FASO

JIMÉNEZ-PONT Miguel	Centre du Commerce International	SUISSE
KABORE Marcel	Millenium Challenge Account Burkina Faso	BURKINA FASO
KABRE Patoinnewendé Alda	Département d'économie de l'école Polytechnique	FRANCE
KONCOBO Zakaria	Institut national de la statistique et de la démographie (INSD)	BURKINA FASO
KOUAME Salimata	Université de Ouagadougou	BURKINA FASO
LOADA Sibiri Martin	Action contre la faim (ACF)	BURKINA FASO
LOADA KABORE Kiswendsida Agnès	Centre pour la Gouvernance Démocratique (CGD)	BURKINA FASO
OUEDRAOGO Michael Gounwindpoulmdo	Institut de recherche empirique en économie politique (IREEP)	BURKINA FASO
PARE Lassina	Direction générale de la promotion de l'économie rurale (DGPER)	BURKINA FASO
PEDEHOMBGA Abdoulaye	Helen Keller International – Burkina	BURKINA FASO
SANOU Momini	Millenium Challenge Account Burkina Faso	BURKINA FASO
SAWADOGO Seydou	Mwangaza Action	BURKINA FASO
SECK Rokhaya	Association pour le développement de la langue Saafi (ADLAS)	SÉNÉGAL
TOUGOURI Julien	Millenium Challenge Account Burkina Faso	BURKINA FASO
YAMEOGO Sibiri François	Direction générale de la promotion de l'économie rurale (DGPER)	BURKINA FASO

Liste des contacts emails des participants :

A la fin de la formation, une liste de contacts contenant votre adresse email sera envoyée à tous les participants. Si vous ne souhaitez pas donner votre adresse email, merci de nous en informer et nous vous supprimerons de cette liste.

Support en ligne pour les participants des formations organisées par J-PAL :

L'année dernière, J-PAL a lancé le "RCT Help Online" (RHO). Cette liste de diffusion modérée a été créée pour lancer un forum de discussion entre tous les participants des formations J-PAL. Vous serez automatiquement invités par email à vous inscrire à cette liste une fois la formation terminée.

INSTRUCTIONS POUR LES TRAVAUX DE GROUPE

Vous allez être répartis en groupes de 4 à 6 personnes. Nous avons essayé au mieux de former les groupes avec des personnes présentant les mêmes domaines d'intérêt mais ayant des expériences différentes. Vous allez réaliser deux types d'activités au sein de ces groupes :

1. des études de cas et des exercices pratiques
2. la préparation d'une proposition d'évaluation

Études de cas et exercices



Chaque cas traite de sujets ayant fait l'objet du cours. Les études de cas fournissent le contexte des évaluations spécifiques qui sont citées dans les cours. En outre, chaque étude implique des discussions qui vous amèneront à réfléchir sur les sujets des cours suivants. Vous devrez lire l'étude correspondante et parcourir les sujets de discussion avant le travail en groupe.

Certaines études de cas sont également accompagnées d'exercices que vous devrez réaliser. Les fichiers Excel contenant ces exercices ainsi que le logiciel utilisé pour l'exercice 3 se trouvent sur la clé USB que vous avez reçue à votre arrivée.

Présentation de groupe

Au cours de la semaine, chaque groupe travaillera sur une proposition d'évaluation sur le sujet de son choix. Les cours et les études de cas aborderont divers aspects de l'évaluation ; ces aspects devront apparaître dans la présentation de groupe.

Pour vous aider à préparer votre présentation, une matrice sous PowerPoint est disponible sur la clé USB.

Le vendredi, chaque groupe présentera sa proposition, les professeurs et les autres participants pourront intervenir et poser des questions. C'est le moment idéal pour bénéficier de commentaires sur une évaluation que vous pouvez déjà être en train de préparer.



Pour vendredi, vous devrez donc préparer :

1. Une présentation de 20 minutes (avec 10 minutes supplémentaires pour les questions et les réactions).
2. Une note de 3 à 4 pages résumant la proposition.

La présentation et la note traiteront des sujets suivants :

1. L'objectif et la logique de l'évaluation : quelle est la question posée et en quoi est-elle importante ou intéressante ?
2. Conception du tirage au sort: comment les groupes Test et Témoin seront-ils déterminés et à quel niveau le tirage au sort interviendra-t-il ?
3. Problèmes de mesures : comment mesurerez-vous la réussite du programme ? Sur quelles variables les données seront-elles collectées ? Comment seront-elles collectées ? Outre les mesures de résultat final, recueillerez-vous des données sur le mécanisme faisant fonctionner le programme ? Si oui, quelles données recueillerez-vous à ce sujet ?
4. Quelle taille d'échantillon utiliserez-vous ? Quelle est la taille correcte de l'échantillon ?
5. Quels sont les risques pour l'intégrité de l'évaluation ? Comment essaierez-vous de les réduire ?
6. Comment les données seront-elles analysées ?

À quel usage réservez-vous les résultats ? En quoi les résultats affecteront-ils l'avenir des politiques et des programmes ?

GROUPES DE TRAVAIL

Groupe 1

Assistant : Emmanuel Bakirdjian

Stéphanie Aubin
Omer Combary
Sibiri Martin Loada
Lassina Pare
Julien Tougouri

Groupe 2

Assistant : Adrien Bouguen

Z. François Bacye
Mahamadi Balima
Rakissiwindé Bationo
Seydou Dao
Miguel Jiménez-Pont

Groupe 3

Assistant : Bastien Michel

Ousseynou Ba
Dramane Bako
Nadège Bhakkan-Mambir
S. Richard Guissou
Marcel Kabore

Groupe 4

Assistante : Élise Pésonel

Rosilin Bock
Églantine Gabaix-Hialé
Patoinnewendé Alda Kabre
Kiswendsida Agnès Loada Kabore
Michael Gounwindpoulmdo Ouedraogo
Rokhaya Seck

Groupe 5

Assistant : Victor Pouliquen

Siriki Zanga Timgana Coulibaly
Salimata Kouame
Francine Ilboudo
Momini Sanou
Sibiri François Yameogo

Groupe 6

Assistante : Juliette Seban

Abdel Aziz Dahi
Aurélié Desrumaux
Zakaria Koncobo
Abdoulaye Pedehombga
Seydou Sawadogo

ÉTUDE DE CAS 1

Les femmes, actrices en politique



Mesurer les effets d'une politique de quotas

Réfléchir aux questions de mesure et d'effets

Introduction

En 1992, l'Inde a modifié sa Constitution et a transféré le pouvoir de décision et d'exécution de programmes de développement des États aux Conseils Ruraux – ou *Gram Panchayat* (Conseils de Village). Ce sont désormais les Conseils de Village qui décident des programmes de développement à entreprendre ainsi que des budgets devant y être consacrés. Les États ont par ailleurs l'obligation de réserver à des femmes, un tiers des sièges aux Conseils de Village et un tiers des postes de Maire du Conseil de Village.

Dans la plupart des États, les villages devant appliquer cette règle du tiers, tant pour les sièges de conseillers que pour celui de Maire, sont choisis au hasard. Cela permet aux chercheurs d'évaluer de manière rigoureuse l'impact des politiques de quotas sur la politique et la gouvernance: les politiques suivies sont-elles différentes lorsqu'il y a plus de femmes en position de responsabilité ? Les politiques choisies par les femmes au pouvoir reflètent-elles les priorités politiques des femmes en général ?

Dans la mesure où la randomisation faisait partie du programme du gouvernement indien, l'organisation de l'évaluation s'est concentrée sur les données nécessaires pour mesurer l'impact du programme. Les questions étaient : quelles données collecter, quels instruments utiliser pour collecter les données et quelle taille l'échantillon doit-il avoir ?

Responsabilisation des Panchayati Raj

Les *Panchayat*, ou Conseils de Village, sont une tradition bien établie en Inde. Une assemblée (*yat*) de cinq aînés (*panch*) choisis par la communauté, se rassemble pour jouer le rôle de médiateur en cas de conflits entre les habitants ou entre les villages. De nos jours, les *Panchayat* sont officialisés et devenus les institutions locales de gouvernance.

Le mouvement d'officialisation a été lancé par les leaders de l'Indépendance qui étaient de fervents partisans de la décentralisation du pouvoir. Gandhi était en faveur de l'autogouvernement des villages, système par lequel chaque village se devait d'être « auto-suffisant et capable de gérer ses propres affaires ». Le Premier Ministre Nehru l'a suivi et a octroyé aux *Panchayat* des « pouvoirs importants » afin que les villages « aient une plus grande mesure de *swaraj* (autogouvernement) dans leurs propres villages ».

C'est ainsi que l'Article 40 de la Constitution Indienne – adopté au moment de l'Indépendance du pays – enjoint les États à faire en sorte que les *Panchayat* « fonctionnent en tant qu'unités d'autogouvernement ». Les recommandations préconisaient un système à trois niveaux, les Conseils de Villages constituant l'unité de base.¹ La majeure partie des États ont suivi la Directive ainsi que les recommandations et c'est ainsi que, dès le début des années 1950, les *Panchayat* ont été officialisés. Néanmoins, dans les années 1960, par manque de pouvoir réel et faute de soutien politique et financier de la part du gouvernement fédéral, les *Panchayat* ont disparu dans la plupart des États, et ce n'est qu'à partir des années 1990 qu'on a commencé à les voir renaître.

Cette renaissance est due à la Constitution. En 1992, l'Inde s'est dotée du 73^{ème} Amendement qui enjoint les États à établir le système *Panchayati Raj* à trois niveaux et à organiser des élections tous les cinq ans. Les Conseillers sont élus par la population et représentent un quartier. Les Conseillers élisent un Maire, pris parmi eux, qui s'appelle le *Pradhan*. Les décisions sont prises par

un vote à la majorité et le Maire n'a pas de droit de veto. Néanmoins, étant le seul membre du Conseil travaillant à plein temps, le Maire jouit d'un véritable pouvoir.

Le 73^{ème} amendement a pour objectif de décentraliser la fourniture de biens et de services publics essentiels au développement des zones rurales. Les États doivent déléguer leurs pouvoirs aux *Panchayat* pour tout ce qui concerne la prévision et l'exécution des programmes de développement local. Les financements continuent de provenir du gouvernement central, mais ils se sont plus affectés à des utilisations spécifiques. Ainsi, ce sont désormais les Conseils de Village qui décident des programmes à mettre en œuvre et des budgets à leur consacrer. Les Conseils de Villages peuvent sélectionner des programmes dans une liste de 29 secteurs, parmi lesquels : les services de santé (par exemple, services aux personnes âgées, aux veuves, maternité, soins aux femmes enceintes et santé des enfants) et les travaux publics (par exemple : raccordement au réseau d'eau potable, construction de routes, logement, bâtiments publics, électricité, irrigation et éducation).

Responsabilisation des femmes dans les Panchayati Raj

Les Conseils de Village sont importants et divers. Par exemple, dans l'État du Bengale Occidental, certains *Panchayat* comptent jusqu'à 12 villages et peuvent représenter jusqu'à 10.000 personnes provenant d'horizons très différents (religion, castes, ethnie et, bien entendu, sexe). Les positions politiques varient en fonction de l'identité de chaque groupe. Si les décideurs politiques ont une influence sur les choix politiques, alors les groupes sous-représentés dans les instances politiques peuvent se retrouver exclus, les Conseils de Village ignorant leurs priorités politiques. Quelques personnes se sont inquiétées du fait que les Conseils de Village, qui venaient tout juste d'accéder à de plus grandes responsabilités, puissent ne pas tenir compte des priorités de développement des groupes traditionnellement marginalisés, comme par exemple les femmes. Pour faire contre-pied à cela, le 73^{ème} Amendement a instauré deux conditions afin de s'assurer que les investissements reflètent bien les besoins de toutes les personnes dans le Conseil de Village.

La première condition assure la participation de la communauté. Si les investissements du Conseil de Village se doivent de refléter les priorités de la communauté, il faut que les Conseillers les connaissent au préalable. Ainsi, les Conseils de Village ont l'obligation d'organiser une Assemblée Générale tous les 6 mois, ou tous les ans, afin de dresser le rapport des activités de la période précédente et de soumettre le budget envisagé à l'approbation de la communauté. De plus, le Maire doit tenir une permanence, à des heures régulières, afin que tous ses administrés puissent officiellement lui faire part de leurs demandes ou déposer plainte. Ces deux conditions permettent aux citoyens de faire connaître leurs préférences politiques.

La seconde condition assure la représentation des femmes au sein du Conseil. Les États ont l'obligation de réserver au moins un tiers de tous les sièges et des postes de Maire à des femmes. Les États doivent s'assurer que les sièges réservés aux femmes sont « attribués conformément à un système tournant, dans les différentes circonscriptions du *Panchayat* » et que les postes de Maire réservés aux femmes sont « affectés à tour de rôle dans les différents *Panchayat* ». En d'autres termes, ils doivent s'assurer que les sièges et postes de Maire réservés aux femmes tournent équitablement entre les circonscriptions et au sein des Conseils de Village.

Quotas aléatoires en Inde : quels enseignements pouvons-nous en tirer ?

Notre équipe d'évaluation s'est vue confier la tâche d'évaluer l'impact de cette politique de quotas pour les femmes dans les Conseils de Village. L'évaluation doit aborder toutes les dimensions sur lesquelles ces quotas sont susceptibles de changer les communautés locales. Quelles données allez-vous collecter? Quels instruments allez-vous utiliser? Quelle sera la taille de votre échantillon?

Dans un premier temps, il convient de bien comprendre la politique de quotas. Quels étaient les besoins? Quels sont les avantages et les inconvénients de cette politique? Quels enseignements pouvons-nous en tirer?

Sujet de discussion N°1 : Quotas par sexe au sein des Conseils de Villages

1.	Quels étaient les principaux objectifs des Conseils de Village ?
2.	Les femmes sont sous-représentées en politique et au pouvoir. Seuls 10 % des membres de l'Assemblée Nationale indienne sont des femmes (17 % au niveau mondial). La sous-représentation des femmes importe t'elle ? Si oui, pourquoi ? Si non, pourquoi ?
3.	Quel est le résultat visé par les rédacteurs du 73ème Amendement en proposant les quotas réservés aux femmes ?

Les quotas par sexe entraînent généralement une très forte augmentation de la représentation politique des femmes. Le Rwanda par exemple est passé de la 24ème place pour ce qui concerne le "nombre de femmes élues au Parlement" à la 1ère place (avec 49 % de femmes élues) après l'introduction d'un système de quotas en 1996. Des changements similaires ont été notés en Argentine, au Burundi, au Costa Rica, en Irak, au Mozambique et en Afrique du Sud. Dans les faits, sur les 20 premiers pays dans ce classement, 17 ont mis en place une politique de quotas.

Imaginez que votre groupe est le Parlement National d'un pays et que vous êtes en train de décider la mise en place de quotas réservés aux femmes pour les élections au Parlement. Divisez votre groupe de manière aléatoire pour avoir deux sous-groupes : l'un en faveur des quotas et l'autre contre.

4.	Débattez du pour et du contre des quotas. A la fin du débat, vous disposerez d'une liste d'arguments pour et contre les quotas.
5.	Quels éléments souhaitez-vous collecter pour étayer les arguments de chaque groupe ?

Quelles données collecter ?

Tout d'abord, vous devez être très clairs sur les impacts potentiels du programme. Vous allez devoir collecter des données sur les dimensions qui seront, selon vous, les plus impactées. Quels sont les sujets principaux pour lesquels il convient d'évaluer la politique de quotas? Quels sont, selon vous, les domaines dans lesquels une politique de quotas va faire une différence?

Quels sont tous les effets possibles des quotas ?

Sujet de discussion N° 2 : Définissez un cadre logique afin de délimiter les résultats intermédiaires et finaux d'intérêt	
1.	Réflexion sur les effets possibles des quotas – tant positifs que négatifs. Suggestion : Utilisez les réponses au Sujet de Discussion N°1 comme point de départ.
2.	Pour chaque effet potentiel identifié, dressez la liste du ou des indicateurs que vous utiliseriez. Par exemple, si l'effet identifié est que les quotas vont avoir un effet sur la participation des femmes au processus politique, l'indicateur pourrait être le « Nombre de femmes participant à l'Assemblée Générale ».
3.	Supposez que vous disposez de ressources et de budget illimités et que vous êtes en mesure de collecter des données sur chacun de ces indicateurs, à la fois dans les communautés "réservées aux femmes" et dans les communautés "non réservées", et les comparer. Combien d'indicateurs souhaiteriez-vous collecter ?

Les effets multiples sont difficiles à interpréter : il faut donc définir une hypothèse

Les quotas en faveur des femmes peuvent avoir un impact sur un grand nombre d'effets, dans différents domaines. Par exemple, les quotas peuvent entraîner une amélioration de l'adduction au réseau d'eau potable au détriment de l'irrigation. Sans hypothèse *ex-ante* de la façon dont ces différentes variables risquent d'être affectées par une politique de quotas, il sera très difficile de comprendre les résultats obtenus. Ainsi par exemple: si vous preniez 500 villages et que vous les répartissiez de manière aléatoire, par groupes, un « groupe témoin » et un « groupe test », puis que vous calculiez les régressions afin de voir quelles sont les différences entre les villages sur 100 effets, vous attendriez vous à constater des différences entre ces villages ? Serait-il logique de justifier ces résultats *ex-post* ?

Il en va de même dans notre cas : si vous vous contentez de présenter votre rapport devant la commission qui vous a mandaté pour évaluer cette politique en expliquant que les quotas réservés aux femmes ont modifié certaines variables et n'ont pas eu d'effet sur d'autres, à quoi va leur servir votre rapport ? Comment les décideurs vont-ils savoir que ces différences sont le résultat de la politique mise en place, plutôt que celui du pur hasard ? Vous devez présenter une hypothèse claire montrant comment les quotas sont sensés modifier la prise de décisions politiques, ce qui va vous amener à prévoir la manière dont chaque effet sera affecté.

Sujet de Discussion N°2 : Définissez un cadre logique afin de déterminer les résultats intermédiaires et finaux	
4.	Identifiez des exemples d'hypothèses clés que vous voudriez tester. Sélectionnez-en une.
5.	Quels indicateurs, ou combinaisons d'indicateurs, allez-vous utiliser pour tester votre hypothèse clé?

Définir un cadre logique afin de déterminer les résultats intermédiaires et finaux

Un bon moyen pour identifier les effets importants consiste à exposer votre théorie du changement, c'est à dire à créer un cadre logique qui relie – étape par étape – l'intervention aux résultats clés finaux.

Sujet de Discussion N°2 : Définissez un cadre logique afin de délimiter les résultats intermédiaires et finaux d'intérêt	
6.	Dans le cas des quotas, quel est l'enchaînement de résultats possible ?
7.	Quelles sont les principales étapes clés qui vont permettre d'obtenir les résultats finaux ? Quelles sont les conditions qui doivent être remplies à chaque étape ?
8.	Quelles variables devez-vous essayer de rassembler à chaque étape de votre cadre logique ?
9.	Utilisez les résultats et les conditions pour établir un cadre logique possible qui va relier l'intervention aux résultats finaux

Cette étude de cas est tirée de l'article "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India," par Raghavendra Chattopadhyay et Esther Duflo (2004a), Econometrica 72(5), 1409-1443.

J-PAL remercie les auteurs qui ont autorisé l'utilisation de leur article.

ÉTUDE DE CAS 2

La campagne « Apprendre à lire »



Comment lire et évaluer les évaluations

Pourquoi le programme Learn to Read (L2R) ?

En 2004, dans le cadre d'une enquête à grande échelle, l'ONG Pratham a découvert que près de 15% des enfants (âgés de 7 à 14 ans) vivant dans les zones rurales de l'État d'Uttar Pradesh étaient incapables de déchiffrer une lettre et que seuls 39 % des enfants étaient capables de lire et de comprendre une histoire simple.

A la même époque, Pratham était en train de mettre au point le module "Learn to Read" (L2R) de sa campagne Read India. Ce module développait une pédagogie particulière pour l'enseignement des compétences de base en lecture, et s'accompagnait d'un effort d'organisation à la base en vue de recruter des bénévoles souhaitant enseigner.

Ce programme permettait aux membres des communautés de s'impliquer plus directement dans l'éducation de leurs enfants, grâce à des réunions de village au cours desquelles les salariés de Pratham donnaient aux villageois des informations sur le niveau d'alphabétisation dans leur village et sur le droit à l'éducation des enfants. Ces réunions ont aussi été pour Pratham, l'occasion de repérer des membres de la communauté souhaitant enseigner. Les bénévoles suivaient ensuite une session de formation sur la pédagogie, ensuite ils pouvaient prendre en charge une classe de soutien à la lecture en utilisant le matériel conçu et fourni par Pratham. Les salariés de l'ONG revenaient régulièrement dans les villages afin de vérifier que les classes de soutien avaient bien lieu ; c'était également pour eux l'occasion de proposer une formation complémentaire si cela s'avérait nécessaire.

Le projet Learn to Read a-t-il marché ?

Le programme Learn To Read de Pratham a-t-il marché ? De quoi avons-nous besoin pour être en mesure de mesurer si un programme a marché, s'il a eu un impact ?

D'une manière générale, lorsque l'on pose la question "le programme a-t-il marché ? », on se demande : est-ce que le programme a atteint ses objectifs et a changé certains effets pour les personnes ayant participé au programme ? Ces changements ne sont-ils pas dus à d'autres facteurs, ou à d'autres événements qui se sont produits au même moment ? Afin de montrer que c'est bien le programme qui a causé les changements observés, nous devons simultanément montrer que si le programme n'avait pas été mis en place, les changements observés ne se seraient pas produits.

Mais comment savons-nous ce qui se serait produit ? Si le programme s'est produit, il s'est produit. Mesurer ce qui se serait produit nécessite de franchir les frontières d'un monde imaginaire dans lequel le programme n'aurait jamais été offert à ces participants. Les résultats qui auraient été obtenus par les mêmes participants dans ce monde imaginaire est ce que l'on appelle le **contrefactuel**. Dans la mesure où nous ne pouvons pas observer le véritable contrefactuel, le mieux que nous puissions faire est l'estimer en l'imitant.

Le principal défi auquel nous sommes confrontés lors de l'évaluation de l'impact d'un programme consiste à construire, ou imiter, le scénario contrefactuel. Traditionnellement, pour ce faire, nous sélectionnons un groupe de personnes qui ressemblent autant que possible aux participants mais qui n'a pas bénéficié au programme. Ce groupe est appelé groupe témoin (ou groupe de comparaison). Étant donné que nous voulons être en mesure de dire que c'est bien le programme, et non d'autres facteurs, qui a entraîné les changements au niveau des effets, il est important que la seule différence entre le groupe témoin et le groupe de participants (ou groupe

test) soit bien que le groupe test n'a pas bénéficié du programme. Ensuite, nous estimons "l'impact" qui est la différence observée à la fin du programme entre les résultats obtenus par le groupe témoin et les résultats obtenus par les participants au programme (groupe test).

L'estimation de l'impact ne peut être fiable que si le groupe témoin réussit à imiter parfaitement le scénario contrefactuel. Si le groupe témoin n'est qu'une pauvre imitation du scénario contrefactuel, l'impact sera (dans la plupart des cas) mal évalué. Par conséquent, la décision d'utiliser telle ou telle méthode pour sélectionner le groupe témoin est l'une des décisions les plus importantes au moment de l'élaboration d'une évaluation d'impact.

Tout cela nous ramène à nos questions : est-ce que le programme Learn to Read a marché ? Quel a été son impact sur les niveaux de lecture des enfants ?

Dans notre étude de cas, l'objectif du programme est « d'améliorer le niveau de lecture des enfants » et la mesure de l'effet est le niveau de lecture. Ainsi, lorsque nous demandons si le projet Learn to Read a marché, la question que nous posons est : est-ce que le programme a amélioré le niveau de lecture des enfants ? L'impact est la différence entre les niveaux de lecture après que les enfants ont suivi les cours de soutien à la lecture et les niveaux de lecture qui seraient les leurs si ces cours n'avaient jamais existés.

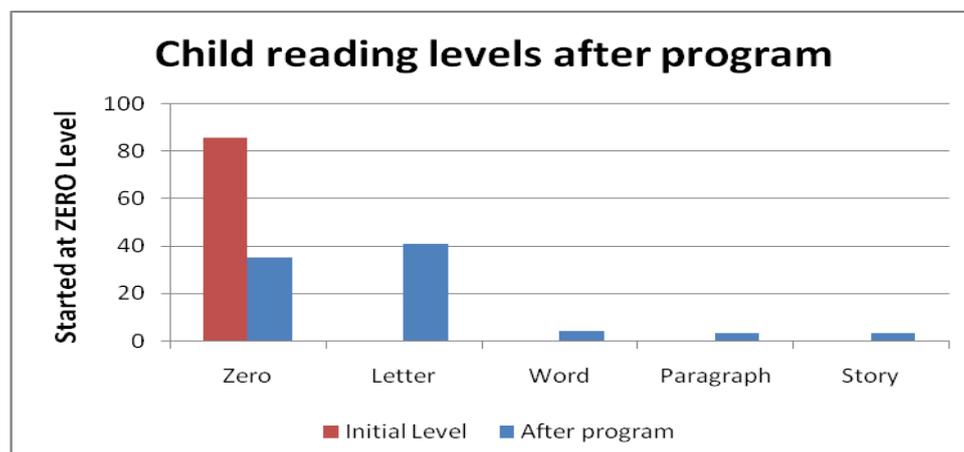
Quels groupes de comparaison pouvons-nous utiliser ?

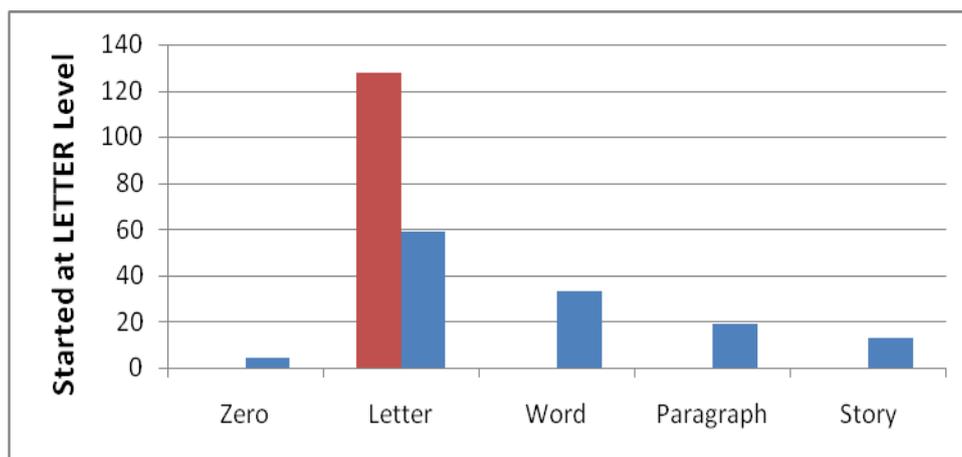
Estimation de l'impact du projet Learn to Read

Méthode n°1 :

Communiqué de presse : Read India aide les enfants à apprendre à lire.

Dans le cadre de sa campagne Read India, Pratham fête le succès de son programme "Learn to Read". L'organisation, dont l'objectif est d'améliorer les niveaux d'alphabétisation des enfants, enregistre des progrès significatifs grâce à de nouvelles méthodes pédagogiques, du matériel scolaire de meilleure qualité, mais surtout grâce à l'investissement des bénévoles. La réussite du programme "Learn to Read" (L2R) apporte la démonstration qu'en modifiant le programme et en mobilisant les énergies au sein des communautés, il est possible d'obtenir des résultats significatifs. Malgré des investissements considérables dans la construction d'écoles et dans la fourniture de déjeuners gratuits, le gouvernement n'est pas parvenu à obtenir des résultats similaires. En moins d'une année, le niveau de lecture des enfants qui ont participé à la campagne L2R a considérablement progressé.





Juste avant le début du programme, la moitié de ces enfants étaient incapables de déchiffrer plus de quelques mots en Hindi – et nombre d’entre eux étaient incapables d’en déchiffrer un seul. Après quelques mois seulement dans les cours de soutien à la lecture de Pratham, plus de la moitié des enfants ont vu leur niveau de lecture progresser d’au moins un échelon : un grand nombre d’entre eux étaient capables de déchiffrer des mots et plusieurs étaient en mesure de lire des paragraphes et des histoires ! *En moyenne, le niveau d’alphabétisation des élèves a progressé de près d’un échelon sur la période.*

Sujet de discussion n°1

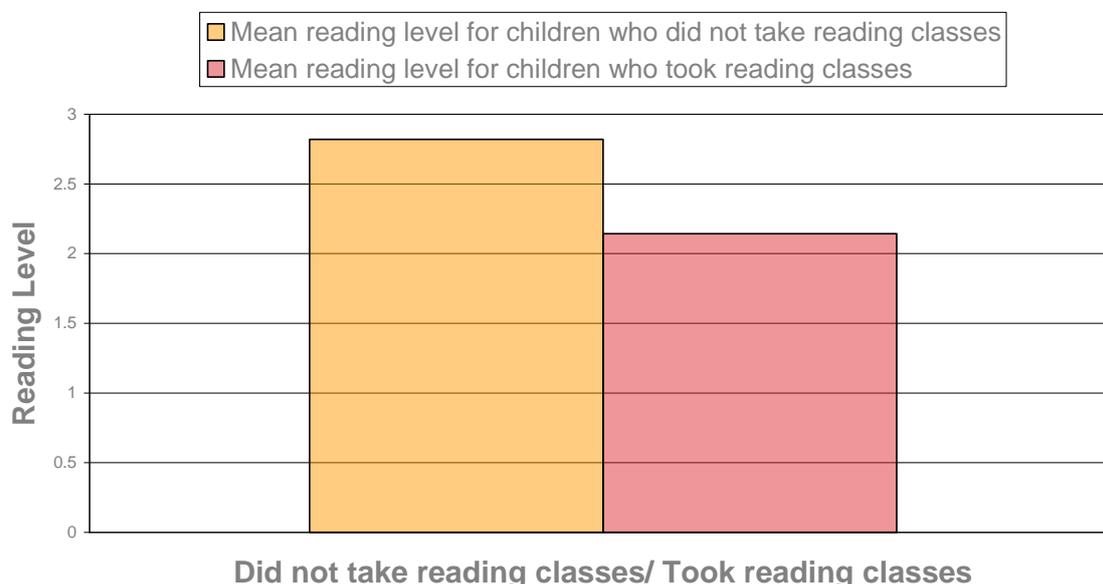
1. Quel est le type d’évaluation impliqué dans ce communiqué de presse ?
2. Quel est le contrefactuel dans cet exemple ?
3. Quels pourraient être les problèmes avec ce type d’évaluation ? (Utilisez des exemples concrets).

Méthode n°2 :

Editorial : Le projet « Read India » n’est pas à la hauteur des espérances

Pratham a mobilisé des millions de dollars et s’est développée très rapidement pour déployer son programme “Learn to Read” sur la totalité du territoire indien ; pour autant, la question reste posée : les élèves ont-ils réellement appris à lire ? Différents éléments récents semblent indiquer le contraire. Une équipe d’évaluateurs missionnés par Education for All a constaté que les élèves qui avaient suivi les cours de soutien à la lecture atteignaient des niveaux d’alphabétisation bien inférieurs aux niveaux obtenus par leurs homologues dans les villages n’ayant pas bénéficié du programme. Un an après la mise en place des classes de soutien à la lecture, les élèves de Pratham étaient seulement capables de déchiffrer des mots, alors que les autres étaient en mesure de lire des paragraphes entiers.

Comparison of reading levels of children who took reading classes Vs. reading levels of children who did not take them



Notes: Le Niveau de Lecture est une variable pour laquelle 0 indique que l'élève est incapable de lire quoi que ce soit, 1 indique que l'élève sait lire l'alphabet, 2 qu'il est capable de déchiffrer des mots, 3 qu'il est capable de lire des paragraphes et 4, qu'il est capable de lire des histoires complètes.

Sujet de discussion n°2

1. Quel est le type d'évaluation utilisé dans cet Editorial ?
2. Quel est le scénario contrefactuel ?
3. Quel pourrait être le problème avec ce type d'évaluation ? (Utilisez des exemples concrets)

Méthode n°3 :

Courrier des Lecteurs : EFA ferait mieux d'être équitable et précis dans ses évaluations

Récemment, les journaux ont publié de nombreux rapports tout à fait injustes au sujet des programmes déployés par l'ONG Pratham. Dans un récent article, les fonctionnaires de Education For All vont même jusqu'à dire que les classes de soutien organisées par Pratham sont néfastes aux enfants. Néanmoins, il apparaît que dans son analyse, Education for All n'utilise pas les bons outils pour mesurer l'impact du programme Learn to Read. EFA compare les niveaux de lecture des élèves de Pratham à ceux des autres élèves du village mais ne tient pas compte du fait que Pratham cible les enfants particulièrement en retard au début du programme. Si Pratham s'était contenté de recruter les enfants les plus en avancés, puis les avaient comparés à leurs homologues les plus en difficulté, elle serait en mesure de revendiquer un succès sans même avoir organisé la moindre classe ! Ce n'est pas ce qu'a fait Pratham.

Pratham ne s'attend pas à voir ses élèves analphabètes atteindre un niveau supérieur à celui des enfants les plus avancés du village. Tout ce que Pratham essaye de faire, est d'initier une amélioration de la situation. Ainsi, le critère de mesure devrait être *l'amélioration* des niveaux de lecture – et non pas le niveau final. Nous avons refait l'analyse de l'EFA, en utilisant une mesure plus appropriée des effets: nos résultats démontrent une progression deux fois plus rapide pour les enfants inclus dans le programme Pratham que pour les autres (le niveau de lecture a progressé de 0,6, contre 0,3 pour les enfants n'ayant pas bénéficié du programme). Du point de vue statistique, cette différence est très importante.

Sujet de discussion n° 3

1. Quel type d'évaluation utilise ce courrier ?
2. Quel est le scénario contrefactuel ?
3. Quel pourrait être le problème avec ce type d'évaluation ? (Utilisez des exemples concrets)

Méthode n°4 :

Les chiffres ne mentent pas ... sauf lorsque les statisticiens sont aveugles

Pratham crie victoire ... Ses opposants crient à l'imposture. Comme souvent dans un tel cas, en y regarder de plus près, on constate que la vérité se situe quelque part au milieu du gué.

La presse s'est fait l'écho de la guerre que se livrent les supporters et les détracteurs de Pratham. Pratham, et ses défenseurs, affirment que la campagne Read India a permis une forte augmentation de l'alphabétisation des enfants. Plusieurs détracteurs déclarent qu'en négligeant les écoles, les programmes de Pratham sont en réalité très néfastes pour les élèves. Malheureusement, les adversaires qui se livrent bataille utilisent des outils d'analyse qui posent tous de sérieux problèmes. La grande victime de cette bataille est le public qui aimerait obtenir une réponse à la question suivante : Pratham aide-t-elle les bénéficiaires qu'elle cible ?

Notre rapport utilise des méthodes statistiques complexes afin de mesurer le véritable impact des programmes de Pratham. Notre inquiétude était de savoir comment traiter les variables qui avaient créé la confusion au niveau des résultats précédents. Nous avons donc mené une enquête dans ces villages et collecté des données portant sur l'âge des enfants, leur niveau de classe et le niveau d'éducation des parents ; nous avons ensuite utilisé ces données pour prédire les résultats des enfants aux tests.

Le Tableau 1 fait ressortir des résultats positifs, des résultats négatifs ainsi que des "absences de résultat" en fonction de la variable étudiée. Les résultats de la colonne (1) indiquent que le programme de Pratham est néfaste pour les enfants. On constate une corrélation négative entre les classes ayant bénéficié du programme de Pratham et les résultats de lecture finaux (-0,68). La colonne (3), qui évalue les améliorations, suggère des résultats impressionnants (0,24). Mais si l'on regarde les résultats pour les enfants (soit du point de vue du niveau, soit du point de vue de l'amélioration), et si l'on *contrôle* en fonction des niveaux de lecture initiaux, de l'âge, du sexe et du niveau d'éducation des parents – tous ces paramètres étant des paramètres déterminants du niveau de lecture des enfants – nous ne constatons aucun impact des programmes de Pratham.

Table 1: Reading outcomes							
	Level				Improvement		
	(1)		(2)		(3)		(4)
Reading Classes	-0.68	**	0.04		0.24	**	0.11
	(0.0829)		(0.1031)		(0.0628)		(0.1081)
Previous reading level			0.71	**			
			(0.0215)				
Age			0.00				-0.01
			(0.0182)				(0.0194)
Sex			-0.01				0.05
			(0.0469)				(0.0514)
Standard			0.02				-0.08
			(0.0174)				(0.0171)
Parents Literate			0.04				0.13
			(0.0457)				(0.0506)
Constant	2.82		0.36		0.37		0.75
	(0.0239)		(0.2648)		(0.0157)		(0.3293)
School-type controls	No		Yes		No		0.37

Notes: The omitted category for school type is "Did not go to school". Reading Level is an indicator variable that takes value 0 if the child can read nothing, 1 if he knows the alphabet, 2 if he can recognize words, 3 if he can read a paragraph and 4 if he can read a full story

Ainsi, en contrôlant les variables appropriées, nous constatons que d'une part Pratham n'a pas été néfaste, contrairement ce que le disent ses opposants, et que d'autre part, elle n'a pas aidé les enfants dans leur apprentissage. Ainsi, Pratham n'a pas été en mesure de nous convaincre qu'elle est en mesure d'investir efficacement l'argent de ses donateurs.

Sujet de Discussion n° 4

1. Quel type d'évaluation ce rapport utilise-t-il ?
2. Quel est le scénario contrefactuel ?
3. Quel pourrait être le problème avec ce type d'évaluation ? (Utilisez des exemples concrets)

Cette étude de cas se base sur l'article intitulé "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in India," par Abhijit Banerjee (MIT), Rukmini Banerjee (Pratham), Esther Duflo (MIT), Rachel Glennerster (J-PAL), et Stuti Khemani (La Banque Mondiale). J-PAL remercie les auteurs qui ont autorisé l'utilisation de leur article.

ÉTUDE DE CAS 3

Enseignants supplémentaires



Comment randomiser ?

Poser les bonnes questions

En 2005, face à la surcharge des salles de classe et au manque d'enseignants, l'ONG International Child Support Africa (ICS) a proposé d'aider les écoles de l'ouest du Kenya en mettant à disposition des enseignants vacataires dans 140 écoles primaires. Dans le cadre d'un programme de deux ans, ICS a fourni les fonds nécessaires à ces écoles pour qu'elles puissent recruter un enseignant supplémentaire. A la différence des fonctionnaires du Ministère de l'Éducation, les enseignants vacataires étaient recrutés localement par les Conseils d'École. L'objectif d'ICS était que le programme améliore le niveau d'apprentissage des élèves grâce, notamment, à une diminution du nombre d'élèves par classe, et aussi parce que les enseignants recrutés devaient rendre compte directement aux communautés pour lesquelles ils travaillaient. Cependant, les enseignants vacataires sont généralement moins bien formés et perçoivent un salaire inférieur à celui de leurs homologues fonctionnaires. La question était donc de savoir si ces enseignants seraient suffisamment motivés, compte tenu de leur rémunération, ou suffisamment compétents, compte tenu de leur niveau de qualification.

Quels sont les protocoles expérimentaux qui ont permis de tester l'impact de cette intervention sur la réussite scolaire ? Parmi les changements apportés dans le paysage scolaire, quels sont ceux qui sont principalement responsables de l'amélioration des résultats des élèves ?

Des classes surchargées

A l'instar de nombreux autres pays en développement, le Kenya a fait d'énormes progrès pour atteindre les Objectifs de Développement du Millénaire, et notamment pour garantir l'accès universel à l'enseignement primaire. Du fait, en grande partie, de l'introduction de la gratuité de l'école à compter de 2003, le nombre d'enfants inscrits à l'école primaire a progressé de près de 30 %, passant de 5,9 millions d'enfants inscrits en 2002 à 7,6 millions en 2005.

Faute d'avoir été accompagné de financements gouvernementaux, ce progrès est à l'origine de toute une série de nouveaux défis au Kenya :

1) Taille des classes : Du fait des contraintes budgétaires, l'augmentation du nombre d'enfants inscrits à l'école primaire n'a pas été accompagnée par une augmentation proportionnelle du nombre d'enseignants (les salaires des enseignants représentent déjà la part la plus importante du budget du Ministère de l'Éducation). Cela s'est traduit par un nombre très important d'élèves dans les classes, en particulier dans les petites classes. Dans un échantillon de classes dans l'ouest du Kenya par exemple, les classes de première année de primaire comptent en moyenne 83 élèves. Ceci est inquiétant car tout le monde s'accorde à penser que les enfants jeunes ont particulièrement besoin d'être dans des classes peu chargées car c'est à cet âge là qu'ils commencent à s'acclimater à l'école. Le Syndicat National des Enseignants Kenyans estime que le pays aurait besoin de recruter 60.000 instituteurs supplémentaires, en plus des 175.000 enseignants déjà présents, afin de pouvoir être présent auprès de tous les enfants en âge d'aller à l'école primaire et de faire baisser le nombre d'élèves par classe.

2) Absentéisme des enseignants : Le problème du nombre d'élèves par classe est exacerbé par l'absentéisme des enseignants qui reste très élevé, atteignant jusqu'à 20 % dans certaines régions du Kenya.

Dans les faits, les enseignants absents ne sont pas remplacés et les enfants soit restent à traîner dans la cour, soit rentrent chez eux, soit sont envoyés dans une autre classe, souvent d'un niveau différent. Les petites écoles, qui sont la majorité dans les zones rurales des pays en développement, peuvent être purement et simplement fermées du fait de l'absence de l'enseignant. Avant de décider d'envoyer ou non leur enfant à l'école, les parents doivent

commencer par se demander si l'école sera ouverte. Cette situation se traduit naturellement par un fort absentéisme des élèves – même les jours où l'école est ouverte.

3) Hétérogénéité des classes : Au Kenya, les classes sont également très hétérogènes et les différences sont très importantes d'un élève à un autre en termes de niveau de préparation à l'école et d'aide à la maison.

L'approche qui consiste à regrouper les élèves par groupes en fonction de leurs capacités (ce que l'on appelle le *tracking* ou affectation par niveau) est une approche très controversée parmi les universitaires et les décideurs politiques. D'un côté, s'il est plus facile pour les enseignants d'enseigner à des groupes homogènes d'élèves, les groupes de niveau rendraient l'école plus efficace et permettraient d'améliorer les résultats aux examens. D'un autre côté, nombre de personnes défendent la thèse selon laquelle si les élèves apprennent en partie auprès de leurs pairs, les groupes de niveau désavantageraient les élèves les moins performants, ce qui aurait pour effet d'exacerber les inégalités.

4) Rareté du matériel scolaire : Compte tenu du coût élevé des matériels éducatifs et du nombre croissant d'élèves, les moyens sont rares dans l'éducation et il n'est pas rare de voir jusqu'à 4 élèves partageant un même manuel. De plus, lorsqu'une infrastructure est totalement surchargée, elle se détériore plus rapidement, de même les manuels s'abiment plus vite à partir du moment où ils sont utilisés par plusieurs élèves.

5) Faible taux d'élèves terminant leur cursus : En conséquence de cette situation, le nombre d'élèves arrivant au bout du cursus primaire est faible au Kenya : seuls 45,1 % des garçons et 43,3 % des filles vont jusqu'au bout de la première année.

Tous ces problèmes mis bout à bout posent de nouveaux défis aux communautés : comment assurer une éducation ayant un niveau de qualité minimum au Kenya, compte tenu des contraintes budgétaires que connaît le pays ?

Qui sont les Enseignants Vacataires ?

Dans plusieurs pays en développement, les gouvernements ont tenté de répondre à cette situation en recrutant localement des enseignants vacataires pour combler les postes qui n'étaient pas pourvus. Ces enseignants vacataires n'ont pas le statut de fonctionnaire : (1) ils sont nommés sur la base d'un contrat annuel renouvelable, sans garantie de renouvellement (à la différence des enseignants fonctionnaires) ; (2) ils sont souvent moins qualifiés que les enseignants traditionnels et peu susceptibles de détenir un véritable diplôme d'enseignement ; (3) leur salaire est inférieur à celui des enseignants officiels (ils perçoivent en général une rémunération qui équivaut à 1/5^{ème} du salaire des enseignants officiels) ; et (4) la plupart du temps, ils sont originaires de l'endroit où est l'école.

Les enseignants vacataires sont-ils efficaces ?

Le recours croissant à des enseignants vacataires a été l'une des plus importantes innovations politiques permettant d'assurer l'enseignement primaire dans les pays en développement ; c'est également celle qui a suscité le plus de controverses. Les personnes qui soutiennent cette politique considèrent que c'est un moyen efficace pour étendre l'accès à l'éducation et assurer un enseignement de qualité à un grand nombre d'élèves de première génération. Conscients du fait que la décision du Conseil d'École de renouveler leur contrat pour l'année suivante peut dépendre des résultats de leurs élèves, les enseignants vacataires sont motivés pour travailler plus dur que leurs homologues fonctionnaires. De plus, les enseignants vacataires viennent de la même zone

géographique que leurs élèves, ils ont la même culture et partagent bien souvent leur statut socio-économique. Les opposants arguent du fait que certes, il y a un enseignant dans chaque classe, mais que dans la mesure où ces enseignants sont sous-qualifiés et non formés, ils ne sont pas en mesure de produire des résultats en termes d'apprentissage. De plus, le recours à des enseignants vacataires dévalorise la profession d'enseignant et met à mal la motivation de tous les enseignants. Donc, même si cette approche peut être utile à court terme, elle peut se révéler nuisible pour le recrutement d'enseignants hautement qualifiés à l'avenir.

Si le recours aux enseignants vacataires a suscité de nombreuses polémiques, il n'existe que peu d'éléments tangibles permettant de connaître leur efficacité à améliorer les résultats d'apprentissage des élèves.

Évaluation aléatoire du programme "Des enseignants supplémentaires"

En janvier 2005, International Child Support Africa a lancé un programme de deux ans pour étudier l'effet des enseignants vacataires sur l'éducation au Kenya. Dans le cadre de ce programme, ICS a fourni des fonds à 140 Conseils d'Écoles afin qu'ils puissent recruter un enseignant vacataire pour prendre en charge une classe supplémentaire de première année. Le but de cette intervention était de répondre aux trois premiers défis, à savoir : taille des classes, responsabilité des enseignants et hétérogénéité des niveaux. L'évaluation a été conçue pour mesurer l'impact de la réduction du nombre d'élèves par classe, l'efficacité relative des enseignants vacataires et pour apprécier si la répartition des élèves en fonction de leurs niveaux pouvait avoir un impact à la fois sur les élèves en difficulté et sur ceux réussissant à l'école.

Comment traiter différentes questions étudiées grâce au protocole expérimental ?

On peut avoir recours à différentes stratégies de randomisation. Quelle stratégie de randomisation pourrait être utilisée pour évaluer les questions suivantes ? Pour chacune de ces questions, attachez-vous à l'unité de randomisation (niveau) adapté.

Sujet de discussion n°1 : Test de l'efficacité des enseignants vacataires

1. Quelle est l'efficacité relative des enseignants vacataires par rapport aux enseignants officiels ?

Sujet de discussion n°2 : Étude des approches plus générales permettant d'améliorer l'éducation

2. Quel est l'effet d'une baisse du nombre d'élèves par classe sur la réussite des élèves ?
3. Quel est l'impact du regroupement des élèves en fonction de leurs aptitudes sur la réussite des élèves ?

Sujet de discussion n°3 : Peut-on aborder toutes les questions au moyen d'une évaluation unique

4. Une seule évaluation permet-elle d'aborder tous ces problèmes en même temps ?
5. Quelle stratégie de randomisation permettrait de le faire ?

Cette étude de cas est tirée de l'article intitulé "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya" par Esther Duflo, Pascaline Dupas et Michael Kremer. 2009

J-PAL remercie les auteurs qui ont autorisé l'utilisation de leur article.

ÉTUDE DE CAS 4

Déparasitage au Kenya



Gérer les menaces à l'intégrité d'une expérimentation

Sur la période 1998 – 2004, l'ONG International Child Support Africa a déployé une vaste campagne de déparasitage dans 75 écoles primaires dans l'ouest du Kenya. Dans le cadre de ce programme, 30.000 élèves inscrits dans ces écoles ont reçu un test contre les vers (ankylostomes, ascaris, trichures et vers de Guinée)

L'affectation aléatoire permet d'assurer que les groupes Test et Témoin sont comparables au début du programme, mais ne permet pas d'assurer qu'ils restent comparables à la fin du programme. L'affectation aléatoire ne permet pas non plus de s'assurer que les personnes suivent le test qui leur a été prescrit. Par ailleurs, la vie continue après la randomisation : d'autres événements, en dehors du programme, se produisent entre le moment de la randomisation et la fin du programme. Ces événements peuvent introduire de nouveaux biais de sélection ; ils amoindrissent la validité des estimations d'impact et font peser une menace sur l'intégrité de l'expérimentation.

Comment est-il possible de gérer les éléments qui menacent l'intégrité des expérimentations ?

Les vers : un problème courant - une solution qui ne coûte pas cher

Les vers sont à l'origine de 40 % des maladies tropicales qui sévissent dans le monde. Les infections sont monnaie courante dans les zones dans lesquelles le niveau d'hygiène est faible. Plus de 2 milliards de personnes sont infectées. Les enfants, qui continuent d'apprendre les bonnes pratiques d'hygiène, sont particulièrement vulnérables : 400 millions d'enfants en âge d'aller à l'école souffrent d'une infection chronique par des parasites intestinaux.

Les vers ont des effets qui vont au-delà de la santé des enfants. Les symptômes comprennent : agitation, diarrhées, douleurs abdominales et anémie. Au-delà des effets sur la santé et la nutrition, les parasitoses peuvent ralentir le développement physique et mental des enfants et faire baisser leur assiduité – et leurs performances – à l'école.

De mauvaises conditions sanitaires et de mauvaises habitudes d'hygiène facilitent la transmission. Les personnes infectées excrètent les œufs des vers dans leurs selles et leurs urines. Dans les zones dans lesquelles les conditions sanitaires sont mauvaises, les œufs contaminent ensuite les sols et l'eau. Les autres personnes sont infectées par l'ingestion de terre ou d'eau souillée (ankylostomes, ascaris et trichures) ou par pénétration des larves au travers de la peau après contact avec de la terre ou de l'eau contaminée. (ankylostomes ou vers de Guinée). Les enfants en âge d'aller à l'école sont plus susceptibles de disséminer les vers car leurs pratiques d'hygiène sont plus à risque (ils sont plus susceptibles de se baigner dans des eaux contaminées, plus susceptibles de ne pas utiliser les latrines, moins susceptibles de se laver les mains avant de manger). Ainsi, en traitant un enfant, non seulement on réduit la charge parasitaire chez l'enfant lui-même, mais de plus, on fait baisser la transmission de la maladie – ce qui profite à la communauté dans son ensemble.

Les tests éradiquent les vers présents dans l'organisme, mais ne préviennent pas les réinfections. Il existe des médicaments qui, administrés oralement, sont en mesure d'éradiquer 99 % de la charge parasitaire dans l'organisme : il s'agit de l'albendazole ou du métabendazole pour le test des ankylostomes, des ascaris et des trichures ; et du praziquantel pour le test des vers de Guinée. Ces médicaments sont sûrs et ne coûtent pas cher : une dose d'albendazole ou de métabenzole coûte moins de 3 cents de dollars US, et une dose de praziquantel coûte moins de 20 cents de dollars US. Les effets secondaires de ces médicaments sont mineurs et peu nombreux.

Les vers colonisent les intestins et les voies urinaires mais ne se reproduisent pas dans l'organisme. Les vers pénètrent dans l'organisme par contacts répétés avec de la terre ou de l'eau

contaminée. Dans les zones de forte prévalence, l'OMS recommande des campagnes de déparasitage préventives dans les écoles. Dans les écoles dans lesquelles la prévalence d'ankylostomes, d'ascaris et de trichures est supérieure à 50%, on recommande un test de masse à l'albendazole tous les 6 mois ; quant aux écoles dans lesquelles la prévalence de vers de Guinée est supérieure à 30 %, on recommande un test de masse au praziquantel à raison d'une fois par an.

Le Programme de Déparasitage à l'École

International Child Support Africa (ICS) a déployé un Programme de Déparasitage des Ecoles Primaires (Primary School Deworming Program - PSDP) dans le District de Busia, dans l'ouest du Kenya, une région fortement peuplée et dans laquelle la prévalence des vers est élevée. Le test en vigueur était conforme aux recommandations de l'OMS et les médicaments étaient administrés par des infirmiers du Ministère de la Santé, en présence de responsables de l'ONG ICS.

On s'attendait à ce que le programme PSDP ait un impact sur la santé, la nutrition et l'éducation des enfants. Pour mesurer les impacts du programme, ICS a collecté des données sur une série de paramètres : prévalence de l'infection aux vers, charge parasitaire (sévérité de l'infection), maladies déclarées par les malades eux-mêmes, assiduité à l'école et résultats aux examens.

Protocole d'évaluation : l'expérimentation telle que prévue

En raison de contraintes administratives et financières, le programme n'a pas pu être déployé immédiatement dans toutes les écoles. A la place, les 75 écoles ont été réparties de manière aléatoire en trois groupes de 25 écoles qui ont été intégrées progressivement au programme en trois ans. Les écoles du groupe 1 ont commencé à être traitées à partir de 1998 et 1999, celles du groupe 2 courant 1999 et les écoles du groupe 3 à compter de 2001. Ainsi en 1999, les écoles des groupes 1 et 2 constituaient le groupe test et les écoles du groupe 3 le groupe témoin.

Figure 1 : L'expérimentation prévue : calendrier de test du programme PSDP montrant les groupes de test en 1998 et 1999

	1998	1999	2001
Groupe 1	Test	Test	Test
Groupe 2	Témoin	Test	Test
Groupe 3	Témoin	Témoin	Test

Les menaces pesant sur l'intégrité de l'expérimentation

Sujet de discussion n°1 : Les menaces

La randomisation assure que les groupes sont équivalents et par conséquent comparables en début de programme. On estime donc l'impact comme étant la différence entre le résultat moyen obtenu par le groupe test et le résultat moyen obtenu par le groupe témoin. Pour être en mesure de dire que c'est bien le programme qui a produit l'impact, vous devez être en mesure d'affirmer que le programme a été l'unique différence entre le groupe test et le groupe témoin, sur toute la durée de l'évaluation.

1.	Que signifie le fait de dire que les groupes sont équivalents au démarrage du programme ?
2.	Etes-vous en mesure de vérifier que les groupes sont équivalents au démarrage du programme ? Comment ?
3.	Que peut-il se produire en cours d'évaluation qui puisse rendre les groupes non équivalents ?
4.	Comment est-ce que cette non équivalence en fin de programme peut menacer l'intégrité de l'expérimentation ?
5.	Vous avez randomisé et créé un groupe test et un groupe témoin équivalents. Si les groupes demeurent équivalents, que peut-il se passer d'autre après la randomisation qui puisse menacer le fait que le programme était l'unique différence entre les deux groupes ?
6.	Dans la première Etude de Cas, vous avez découvert d'autres méthodes permettant d'estimer l'impact d'un programme, telles que l'analyse par simple différence, par régressions multiples, par régressions multiples avec panel de données, par appariement. <ul style="list-style-type: none"> a. Pour chacune des menaces que vous venez d'identifier, dites si ces menaces existent pour chacune de ces autres méthodes, et si oui, comment ? b. Les menaces qui pèsent sur l'intégrité de l'expérimentation n'affectent-elles que la méthode par randomisation ?

Comment gérer l'attrition lorsque les groupes ne sont plus équivalents

L'attrition est le phénomène qui se produit lorsque des personnes rejoignent ou sortent de l'échantillon – groupe test ou groupe témoin – en cours d'expérimentation. Dans les essais cliniques, un exemple courant est le décès des patients. Cet exemple est même tellement courant que l'attrition est parfois appelée « mortalité expérimentale ».

Sujet de discussion n°2 : Comment gérer l'attrition

Vous voulez étudier les effets du déparasitage sur la santé. Vous êtes en train d'étudier la charge parasitaire (sévérité de l'infection), en la notant de la manière suivante : fortes infections : 3, infections modérées : 2, infections légères : 1.

Le programme est déployé à l'école : les données sont donc collectées dans les écoles, lieu où les enfants sont rassemblés (ce qui évite à la personne qui collecte les données d'avoir à se déplacer chez chaque enfant). La personne qui collecte les données mesure tous les enfants présents dans l'école, un jour choisi de manière aléatoire, les enseignants et les autorités scolaires ne sont pas prévenus à l'avance.

Il y a 30.000 enfants : 15.000 dans les écoles du groupe test et 15.000 dans les écoles du groupe témoin. Après randomisation, les groupes sont équivalents et des enfants des trois niveaux d'infection sont également représentés.

Le protocole est respecté à 100 % : tous les enfants dans le groupe test sont réellement traités et aucun enfant du groupe témoin n'est traité. Le déparasitage en début d'année scolaire se traduit par une charge parasitaire de 1 en fin d'année, en raison des réinfections. Les enfants qui ont une charge parasitaire de 3 ne sont présents à l'école que la moitié du temps et sortent de l'étude s'ils ne sont pas traités. Le tableau ci-dessous montre le nombre d'enfants dans chaque catégorie de charge parasitaire, avant et après le test.

Charge Parasitaire	Avant		Après	
	Test	Témoin	Test	Témoin
3	5,000	5,000	0	Sortis
2	5,000	5,000	0	5,000
1	5,000	5,000	15,000	5,000
Total d'enfants testés à l'école	15,000	15,000	15,000	10,000

- Avant l'examen, quelle est la charge parasitaire pour le groupe test ?
 - Avant l'examen, quelle est la charge parasitaire pour le groupe témoin ?
 - Les groupes sont-ils équivalents ?
- Après l'examen, quelle est la charge parasitaire pour le groupe test ?
 - Après l'examen, quelle est la charge parasitaire pour le groupe témoin ?
 - Quelle est la différence ?
- Calculez les différences de résultats au début et à la fin de l'année
 - Cette différence entre les résultats est-elle une estimation fiable de l'impact du programme ?
 - Si elle n'est pas fiable, surestime-t-elle l'impact, ou à l'inverse, le sous-estime-t-elle ?
- Dans la mesure où le groupe test a été traité, vous vous attendiez à constater une différence entre les groupes à la fin de l'année.

 - Si cette différence correspond à un effet, quelle est la source d'attrition, si attrition il y a ?
 - Comment pouvez-vous résoudre ce problème afin d'obtenir une meilleure estimation de l'impact du programme ?
- Après le test, quelle est la charge parasitaire pour le groupe témoin si vous testez également les 5000 enfants qui sont sortis de l'étude ? (on suppose que tous auraient une charge parasitaire de 3)
 - Calculez l'impact du programme
 - Quelle est la taille du biais d'attrition ?
- Le programme a également étudié les taux d'assiduité à l'école et les résultats scolaires.
 - Une attrition différentielle pourrait-elle biaiser l'un ou l'autre de ces résultats ?
 - L'impact serait-il surestimé ou sous-estimé ?

Comment gérer les situations de contamination partielle (le groupe test n'est pas traité ou le groupe témoin est traité)

Il est possible que certaines personnes affectées au groupe test finissent par ne pas être traitées. Par exemple, pour un programme de soutien scolaire après l'école, il est possible que les enfants devant aller au cours de soutien n'y aillent pas. Il est également possible que les personnes affectées au groupe témoin finissent par avoir accès au cours, soit dans le cadre du programme, soit auprès d'un autre fournisseur. Il est encore possible que les enfants du groupe témoin obtiennent une aide supplémentaire de la part des enseignants, ou acquièrent les matériels et les méthodes du programme auprès de leurs camarades de classe. D'une manière ou d'une autre, ces personnes ne respectent pas l'affectation qui était la leur dans l'expérimentation prévue. C'est ce que l'on appelle "la contamination partielle", la "diffusion" ou pire, la « contamination ». A la différence des expériences de laboratoire qui sont étroitement contrôlées, la diffusion est chose très courante dans les programmes sociaux. Après tout, la vie continue et les personnes restent ce

qu'elles sont et vous ne pouvez pas contrôler ce qu'elles décideront de faire pendant toute la durée de l'expérimentation. La seule chose que vous puissiez faire, c'est planifier votre expérimentation et leur offrir des tests. Dans de telles circonstances, comment allez-vous pouvoir gérer les menaces qui découlent de cette contamination partielle ?

Sujet de discussion n°3 : Comment gérer les problèmes de contamination partielle ?

Tous les enfants des familles les plus pauvres ne possèdent pas de chaussures ; leur charge parasitaire est par conséquent de 3. Bien que leurs parents n'aient pas payé les frais d'inscription à l'école, les enfants sont autorisés à rester à l'école pendant toute l'année scolaire. On a demandé aux parents de donner leur accord pour le test ; pour ce faire, les parents devaient se rendre à l'école et signer un formulaire d'autorisation dans le bureau du Directeur. Dans la mesure où ils n'avaient pas payé les frais d'inscription, les parents les plus pauvres étaient réticents à l'idée de se rendre à l'école. Ainsi, aucun des enfants ayant une charge parasitaire de 3 n'a en réalité été traité et leur note de charge parasitaire était toujours de 3 à la fin de l'année. Aucun des enfants affectés au groupe témoin n'a été traité. Tous les enfants inclus dans l'échantillon au début de l'année ont été suivis, soit à l'école, soit à domicile.

<i>Charge parasitaire</i>	<i>Avant</i>		<i>Après</i>	
	Test	Témoin	Test	Témoin
3	5,000	5,000	5,000	5,000
2	5,000	5,000	0	5,000
1	5,000	5,000	10,000	5,000
Nombre total d'enfants testés	15,000	15,000	15,000	15,000

1.
 - a. Calculez l'estimation de l'impact à partir des affectations de départ
 - b. Que mesure cette estimation de « l'intention de traiter » ?
 - c. Il s'agit d'une mesure précise de l'effet du programme, mais est-ce une bonne mesure ? Quand est-elle utile ? Quand ne l'est-elle pas ?

Vous souhaitez connaître les effets du test sur les enfants effectivement traités

2. Cinq de vos collègues viennent vous voir ; tous sont d'accord pour dire que vous devriez calculer l'effet du traitement en n'utilisant que les 10.000 enfants qui ont été traités.
 - a. Quel est l'impact si vous n'utilisez que les enfants traités ?
 - b. Le conseil est-il bon ? Pourquoi ? Pourquoi pas ?
3. Un autre collègue vous dit que ce n'est pas une bonne idée de ne pas tenir compte du tout des enfants non traités, et que vous devez les inclure mais les considérer comme faisant partie du groupe témoin.
 - a. Quelle est l'estimation de l'impact obtenu dans ce cas ?
 - b. Le conseil est-il bon ? Pourquoi ? Pourquoi pas ?
4. Un autre collègue vous suggère d'utiliser les taux de contamination, c'est à dire la proportion de personnes ayant respecté le test dans chaque groupe. Vous allez devoir diviser « l'intention de traiter » par la différence entre les taux de contamination.
 - a. Quels sont les taux de contamination pour le groupe test et pour le groupe témoin ?
 - b. Quelle est l'estimation de l'impact obtenue dans ce cas ? Le conseil est-il bon ? Pourquoi ? Pourquoi pas ?
5. Le programme a entraîné une prise de conscience autour du problème des vers, alors certains parents du groupe témoin ont acheté les médicaments et ont traité leurs enfants à la maison. Dans l'ensemble, ce sont 2000 enfants du groupe témoin qui ont été traités.

Quelle est l'estimation de l'impact du « traitement sur les traités » ?

Comment gérer les externalités – lorsque le groupe témoin, lui-même non traité, bénéficie du traitement donné au groupe test

Les personnes du groupe témoin peuvent bénéficier indirectement des effets bénéficiant aux personnes traitées. Par exemple, un programme de distribution de moustiquaires imprégnées pour lutter contre la transmission du paludisme peut bénéficier indirectement aux personnes qui elles-mêmes ne dorment pas sous une moustiquaire. Ces effets sont appelés des externalités.

Sujet de discussion n°4 : Comment gérer les externalités

Dans le cadre du programme PSDP, la randomisation a été faite au niveau des écoles.

Dans la zone de l'évaluation, les personnes habitent dans des fermes rapprochées les unes des autres. On a constitué des groupes de fermes, par zones d'un rayon de 3 km. Le diagramme ci-dessous montre trois de ces groupes, A, B et C. Les fermes sont suffisamment proches les unes des autres pour que les enfants jouent ensemble. Par ailleurs, les familles avaient le choix entre plusieurs écoles primaires.

Il y a 3 écoles dans les groupes A et B et 5 dans le groupe C. Il est courant que des enfants de fermes voisines, voire même que des frères et sœurs, fréquentent des écoles différentes. Les écoles du groupe 1 faisaient partie du groupe test et les écoles des groupes 2 et 3 faisaient partie du groupe témoin.

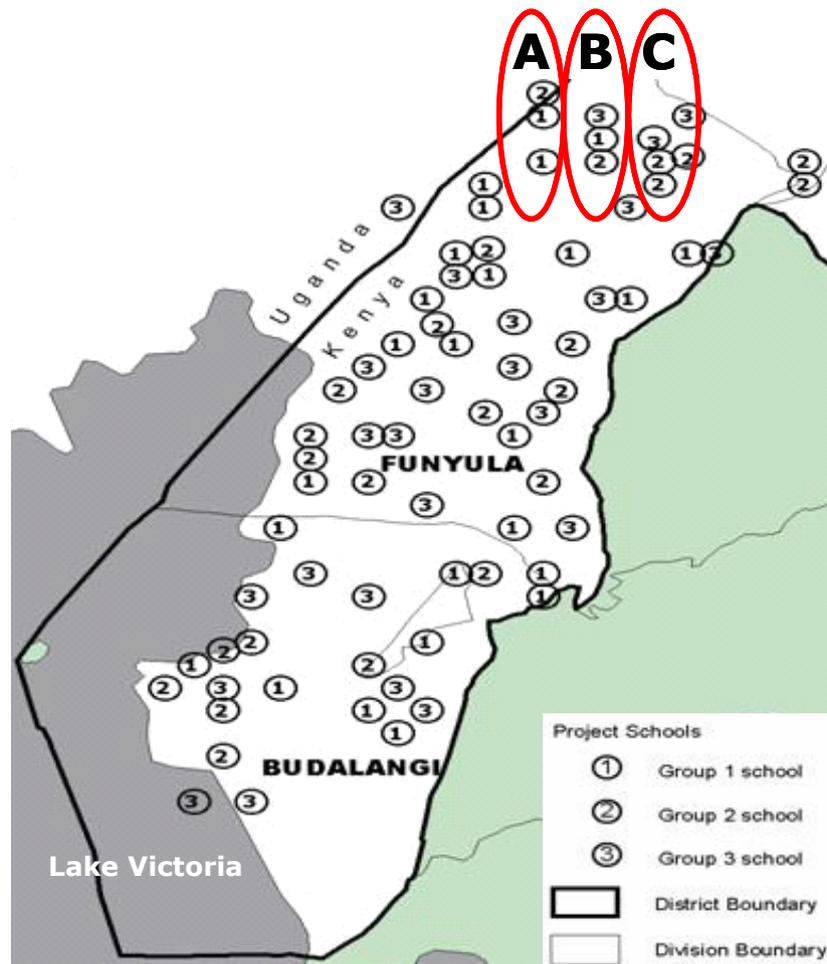
Dans chaque école, il y avait 100 enfants. Le protocole a été respecté à 100 % (tous les enfants du groupe test ont été traités et aucun enfant du groupe témoin n'a reçu de traitement).

-
1. Vous estimez l'impact en comparant les charges parasitaires moyennes dans les écoles du groupe test et dans celles du groupe témoin.

Cette estimation surestime-t-elle ou sous-estime-t-elle l'impact du traitement ?

2.
 - a. La densité de traitement est la proportion de personnes traitées par rapport au nombre de personnes non traitées dans un groupe de personnes données.
 - b. Quelle est la densité de traitement dans les écoles du groupe test la première année ?
 - c. Quelle est la densité de traitement dans les écoles du groupe témoin ?
 - d. Quelles sont les densités de traitement dans les zones A, B et C, pendant la première année ?
 - e. Quelles sont les densités de traitement dans les zones A, B et C, pendant les années 2 et 3 ?
 3.
 - a. S'il y a des externalités, à quoi sont-elles dues selon vous ?
 - b. Avez-vous la possibilité de mesurer les effets des externalités dans les écoles ?
 - c. Si vous pensez ne pas pouvoir mesurer les effets des externalités, de quoi auriez-vous besoin pour le faire ?
 - d. Pouvez-vous mesurer les effets des externalités d'une école à une autre ?
 4. Classez les zones A, B et C en fonction de l'importance des effets des externalités du traitement pour les années 1, 2 et 3.
 5.
 - a. Si vous aviez randomisé au niveau individuel, qu'auriez-vous pu faire pour mesurer les effets des externalités d'une personne à une autre ?
 - b. Si vous aviez randomisé au niveau des écoles, que pouvez-vous faire pour mesurer les effets des externalités d'une école à l'autre ?
 - c. Que suggère cette stratégie générale ?
-

Sujet de discussion n°4 : Comment gérer les externalités



* Les données de géo localisation GPS ont été collectées avant le mois de mai 2000, alors que les Etats-Unis continuaient de réduire la précision internationale des systèmes GPS. Il est possible que les résultats ne soient précis qu'à quelques centaines de mètres près. Ainsi, l'une des écoles du Groupe 3 semble être en Ouganda alors qu'en réalité, elle se situe sur le côté Kenyan de la frontière. Quant à l'école qui semble être au milieu du lac Victoria, elle est en réalité située sur une petite île.

*Cette étude de cas est tirée de l'article intitulé "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," par Edward Miguel et Michael Kremer, "Econometrica 72(1): 159-217, 2004
J-PAL remercie les auteurs qui ont autorisé l'utilisation de leur article.*

EXERCICE 1

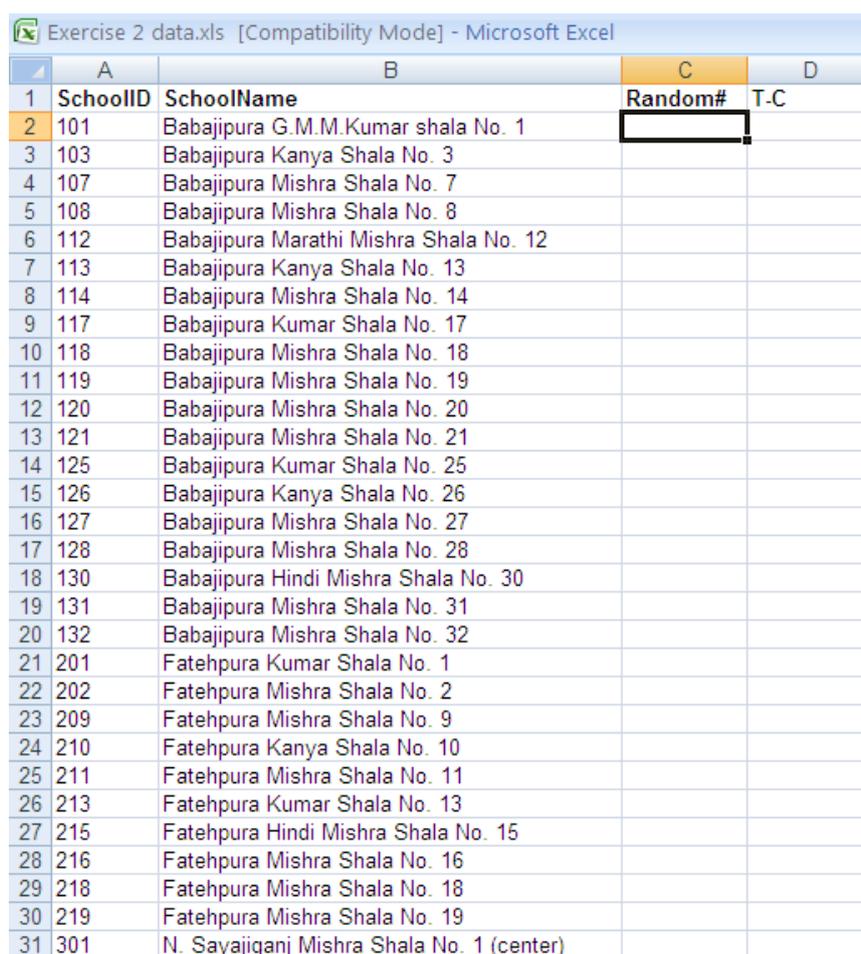
Les mécanismes de l'assignation aléatoire en utilisant MS Excel ®

Partie 1: L'assignation aléatoire simple

Comme beaucoup de programmes, MS Excel a une fonction génératrice de nombre aléatoire.

Supposons que nous ayons une liste d'écoles et que nous souhaitons assigner la moitié des écoles au groupe traitement et l'autre moitié au groupe contrôle.

1. Ouvrir le fichier Excel avec la liste des écoles



	A	B	C	D
1	SchoolID	SchoolName	Random#	T-C
2	101	Babajipura G.M.M.Kumar shala No. 1		
3	103	Babajipura Kanya Shala No. 3		
4	107	Babajipura Mishra Shala No. 7		
5	108	Babajipura Mishra Shala No. 8		
6	112	Babajipura Marathi Mishra Shala No. 12		
7	113	Babajipura Kanya Shala No. 13		
8	114	Babajipura Mishra Shala No. 14		
9	117	Babajipura Kumar Shala No. 17		
10	118	Babajipura Mishra Shala No. 18		
11	119	Babajipura Mishra Shala No. 19		
12	120	Babajipura Mishra Shala No. 20		
13	121	Babajipura Mishra Shala No. 21		
14	125	Babajipura Kumar Shala No. 25		
15	126	Babajipura Kanya Shala No. 26		
16	127	Babajipura Mishra Shala No. 27		
17	128	Babajipura Mishra Shala No. 28		
18	130	Babajipura Hindi Mishra Shala No. 30		
19	131	Babajipura Mishra Shala No. 31		
20	132	Babajipura Mishra Shala No. 32		
21	201	Fatehpura Kumar Shala No. 1		
22	202	Fatehpura Mishra Shala No. 2		
23	209	Fatehpura Mishra Shala No. 9		
24	210	Fatehpura Kanya Shala No. 10		
25	211	Fatehpura Mishra Shala No. 11		
26	213	Fatehpura Kumar Shala No. 13		
27	215	Fatehpura Hindi Mishra Shala No. 15		
28	216	Fatehpura Mishra Shala No. 16		
29	218	Fatehpura Mishra Shala No. 18		
30	219	Fatehpura Mishra Shala No. 19		
31	301	N. Sayajiganj Mishra Shala No. 1 (center)		

2. Assigner un nombre aléatoire à chaque école

La fonction ALEA () (ou RAND () en anglais) est le générateur de nombres aléatoires d'Excel. Pour l'utiliser, dans la colonne C, écrire =ALEA () dans chaque cellule adjacente à chaque nom. Vous pouvez aussi écrire cette fonction sur la ligne 2 et simplement la copier/coller à l'ensemble de la colonne, ou l'étirer.

	A	B	C	D
1	SchoolID	SchoolName	Random#	T.C
2	101	Babajipura G.M.M.Kumar shala No. 1	=RAND()	
3	103	Babajipura Kanya Shala No. 3		
4	107	Babajipura Mishra Shala No. 7		
5	108	Babajipura Mishra Shala No. 8		
6	112	Babajipura Marathi Mishra Shala No. 12		
7	113	Babajipura Kanya Shala No. 13		
8	114	Babajipura Mishra Shala No. 14		
9	117	Babajipura Kumar Shala No. 17		
10	118	Babajipura Mishra Shala No. 18		
11	119	Babajipura Mishra Shala No. 19		
12	120	Babajipura Mishra Shala No. 20		
13	121	Babajipura Mishra Shala No. 21		
14	125	Babajipura Kumar Shala No. 25		
15	126	Babajipura Kanya Shala No. 26		
16	127	Babajipura Mishra Shala No. 27		
17	128	Babajipura Mishra Shala No. 28		
18	130	Babajipura Hindi Mishra Shala No. 30		
19	131	Babajipura Mishra Shala No. 31		
20	132	Babajipura Mishra Shala No. 32		
21	201	Fatehpura Kumar Shala No. 1		
22	202	Fatehpura Mishra Shala No. 2		
23	209	Fatehpura Mishra Shala No. 9		
24	210	Fatehpura Kanya Shala No. 10		
25	211	Fatehpura Mishra Shala No. 11		
26	213	Fatehpura Kumar Shala No. 13		
27	215	Fatehpura Hindi Mishra Shala No. 15		
28	216	Fatehpura Mishra Shala No. 16		
29	218	Fatehpura Mishra Shala No. 18		
30	219	Fatehpura Mishra Shala No. 19		
31	301	N. Sayajiganj Mishra Shala No. 1 (center)		

Ecrire = ALEA() donne un nombre aléatoire à 15 chiffres compris entre 0 and 1.

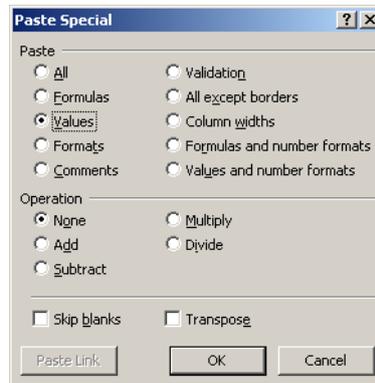
	A	B	C	D
1	SchoolID	SchoolName	Random#	T.C
2	101	Babajipura G.M.M.Kumar shala No. 1	0.80541713	
3	103	Babajipura Kanya Shala No. 3	0.53078382	
4	107	Babajipura Mishra Shala No. 7	0.92449824	
5	108	Babajipura Mishra Shala No. 8	0.81342515	
6	112	Babajipura Marathi Mishra Shala No. 12	0.59650637	
7	113	Babajipura Kanya Shala No. 13	0.58563987	
8	114	Babajipura Mishra Shala No. 14	0.6486176	
9	117	Babajipura Kumar Shala No. 17	0.46206529	
10	118	Babajipura Mishra Shala No. 18	0.18134939	
11	119	Babajipura Mishra Shala No. 19	0.69772005	
12	120	Babajipura Mishra Shala No. 20	0.83992642	
13	121	Babajipura Mishra Shala No. 21	0.85501349	
14	125	Babajipura Kumar Shala No. 25	0.30572517	
15	126	Babajipura Kanya Shala No. 26	0.53388093	
16	127	Babajipura Mishra Shala No. 27	0.46003571	
17	128	Babajipura Mishra Shala No. 28	0.27464658	
18	130	Babajipura Hindi Mishra Shala No. 30	0.02073858	
19	131	Babajipura Mishra Shala No. 31	0.77709404	
20	132	Babajipura Mishra Shala No. 32	0.2362122	
21	201	Fatehpura Kumar Shala No. 1	0.91552715	
22	202	Fatehpura Mishra Shala No. 2	0.95669543	
23	209	Fatehpura Mishra Shala No. 9	0.48508217	
24	210	Fatehpura Kanya Shala No. 10	0.62054343	
25	211	Fatehpura Mishra Shala No. 11	0.17807564	
26	213	Fatehpura Kumar Shala No. 13	0.36389518	
27	215	Fatehpura Hindi Mishra Shala No. 15	0.03446481	
28	216	Fatehpura Mishra Shala No. 16	0.51526826	
29	218	Fatehpura Mishra Shala No. 18	0.17860571	
30	219	Fatehpura Mishra Shala No. 19	0.04501407	
31	301	N. Sayajiganj Mishra Shala No. 1 (center)	0.93881649	

3. Copier les cellules de la colonne C et coller les valeurs sur les mêmes cellules

La fonction =ALEA() crée des nouveaux nombres aléatoires à chaque fois que vous ferez un changement dans la feuille Excel. En effet, Excel recalcule toutes les valeurs de chaque cellule à chaque changement sur la feuille. (Vous pouvez aussi relancer le calcul, et donc re-randomiser en pressant F9.)

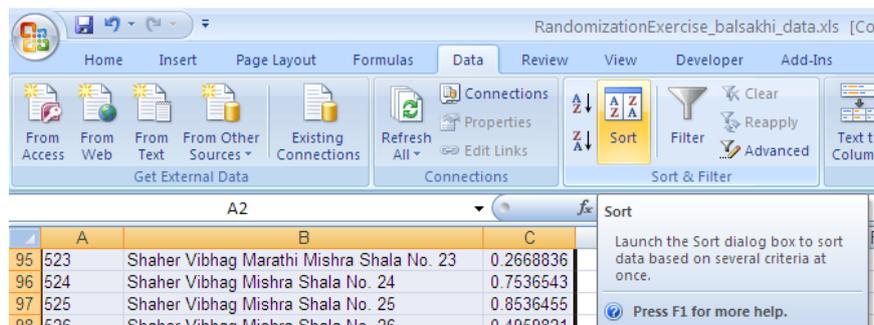
Cela peut toutefois porter à confusion. Une fois que vous avez généré la colonne de nombres aléatoires, vous n'avez pas besoin de re-randomiser. Vous avez déjà une colonne de nombres aléatoires. Pour empêcher Excel de recalculer, il faut remplacer les « fonctions » par les « valeurs ».

Pour faire cela, sélectionner toutes les valeurs de la colonne C et cliquer droit n'importe où dans la colonne sélectionnée et choisir « Copier ». Puis recliquer droit n'importe où dans la colonne sélectionnée et choisir « Collage Spécial ». La fenêtre « Collage Spécial » apparaît, cliquer alors sur « Valeurs ».

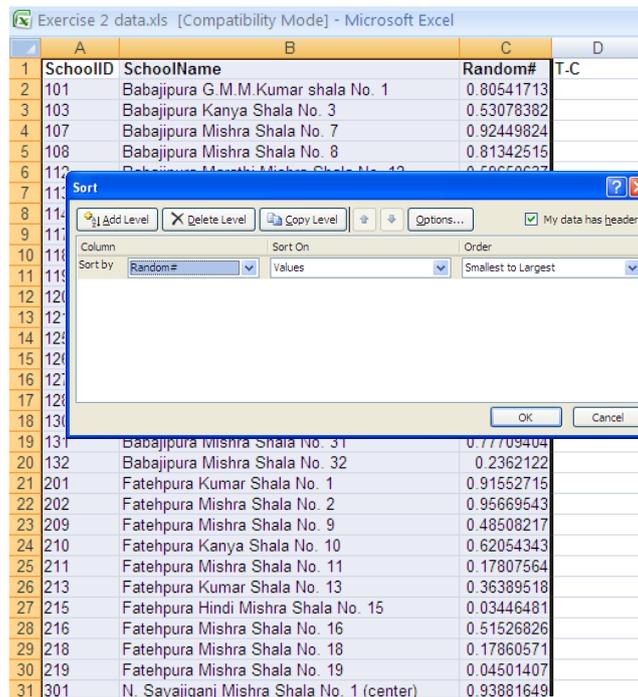


4. Trier l'ensemble des colonnes en ordre ascendant ou descendant de la colonne C

Sélectionner les colonnes A, B, et C et cliquer sur « Trier » :



La fenêtre « Tri » apparaît.



Dans la colonne « Trier par », sélectionner « random # ». Cliquer sur OK. Cette manipulation trie la liste par ordre ascendant ou descendant du chiffre aléatoire, selon ce que vous avez choisi.

Et voilà! Vous avez une liste triée aléatoirement.

SchoolID	SchoolName	Random#	T.C
130	Babajipura Hindi Mishra Shala No. 30	0.02073858	
215	Fatehpura Hindi Mishra Shala No. 15	0.03446481	
219	Fatehpura Mishra Shala No. 19	0.04501407	
211	Fatehpura Mishra Shala No. 11	0.17807564	
218	Fatehpura Mishra Shala No. 18	0.17860571	
118	Babajipura Mishra Shala No. 18	0.18134939	
132	Babajipura Mishra Shala No. 32	0.23621222	
128	Babajipura Mishra Shala No. 28	0.27464658	
125	Babajipura Kumar Shala No. 25	0.30572517	
213	Fatehpura Kumar Shala No. 13	0.36389518	
127	Babajipura Mishra Shala No. 27	0.46003571	
117	Babajipura Kumar Shala No. 17	0.46206529	
209	Fatehpura Mishra Shala No. 9	0.48508217	
216	Fatehpura Mishra Shala No. 16	0.51526826	
103	Babajipura Kanya Shala No. 3	0.53078382	
126	Babajipura Kanya Shala No. 26	0.53388093	
113	Babajipura Kanya Shala No. 13	0.58563987	
112	Babajipura Marathi Mishra Shala No. 12	0.59650637	
210	Fatehpura Kanya Shala No. 10	0.62054343	
114	Babajipura Mishra Shala No. 14	0.6486176	
119	Babajipura Mishra Shala No. 19	0.69772005	
131	Babajipura Mishra Shala No. 31	0.77709404	
101	Babajipura G.M.M.Kumar shala No. 1	0.80541713	
108	Babajipura Mishra Shala No. 8	0.81342515	
120	Babajipura Mishra Shala No. 20	0.83992642	
121	Babajipura Mishra Shala No. 21	0.85501349	
201	Fatehpura Kumar Shala No. 1	0.91552715	
107	Babajipura Mishra Shala No. 7	0.92449824	
301	N. Sayajiganj Mishra Shala No. 1 (center)	0.93881649	
202	Fatehpura Mishra Shala No. 2	0.95669543	

5. Assigner les écoles aux groupes traitement et contrôle

Parce que votre liste est triée de manière aléatoire, le fait qu'une école soit dans le haut ou dans le bas de la liste, est totalement aléatoire. Ainsi, si vous assignez la première moitié (moitié haute) au groupe traitement et la seconde moitié au groupe contrôle, les écoles ont bien été assignées aléatoirement.

Dans la colonne D, écrire « T » pour la première moitié des lignes (lignes 2 à 61) et « C » pour la seconde moitié (lignes 62 à 123).

	A	B	C	D
1	SchoolID	SchoolName	Random#	T.C
2	130	Babajipura Hindi Mishra Shala No. 30	0.02073858	T
3	215	Fatehpura Hindi Mishra Shala No. 15	0.03446481	T
4	219	Fatehpura Mishra Shala No. 19	0.04501407	T
5	211	Fatehpura Mishra Shala No. 11	0.17807564	T
6	218	Fatehpura Mishra Shala No. 18	0.17860571	T
7	118	Babajipura Mishra Shala No. 18	0.18134939	T
8	132	Babajipura Mishra Shala No. 32	0.2362122	T
9	128	Babajipura Mishra Shala No. 28	0.27464658	T
10	125	Babajipura Kumar Shala No. 25	0.30572517	T
11	213	Fatehpura Kumar Shala No. 13	0.36389518	T
12	127	Babajipura Mishra Shala No. 27	0.46003571	T
13	117	Babajipura Kumar Shala No. 17	0.46206529	T
14	209	Fatehpura Mishra Shala No. 9	0.48508217	T
15	216	Fatehpura Mishra Shala No. 16	0.51526826	T
16	103	Babajipura Kanya Shala No. 3	0.53078382	T
17	126	Babajipura Kanya Shala No. 26	0.53388093	C
18	113	Babajipura Kanya Shala No. 13	0.58563987	C
19	112	Babajipura Marathi Mishra Shala No. 12	0.59650637	C
20	210	Fatehpura Kanya Shala No. 10	0.62054343	C
21	114	Babajipura Mishra Shala No. 14	0.6486176	C
22	119	Babajipura Mishra Shala No. 19	0.69772005	C
23	131	Babajipura Mishra Shala No. 31	0.77709404	C
24	101	Babajipura G.M.M.Kumar shala No. 1	0.80541713	C
25	108	Babajipura Mishra Shala No. 8	0.81342515	C
26	120	Babajipura Mishra Shala No. 20	0.83992642	C
27	121	Babajipura Mishra Shala No. 21	0.85501349	C
28	201	Fatehpura Kumar Shala No. 1	0.91552715	C
29	107	Babajipura Mishra Shala No. 7	0.92449824	C
30	301	N. Sayajiganj Mishra Shala No. 1 (center)	0.93881649	C
31	202	Fatehpura Mishra Shala No. 2	0.95669543	C

Retrier la liste dans l'ordre des identifiants des écoles (school id). Vous verrez que les écoles ont bien été assignées aléatoirement entre les groupes contrôle et traitement.

	A	B	C	D
1	SchoolID	SchoolName	Random#	T.C
2	101	Babajipura G.M.M.Kumar shala No. 1	0.80541713	C
3	103	Babajipura Kanya Shala No. 3	0.53078382	T
4	107	Babajipura Mishra Shala No. 7	0.92449824	C
5	108	Babajipura Mishra Shala No. 8	0.81342515	C
6	112	Babajipura Marathi Mishra Shala No. 12	0.59650637	C
7	113	Babajipura Kanya Shala No. 13	0.58563987	C
8	114	Babajipura Mishra Shala No. 14	0.6486176	C
9	117	Babajipura Kumar Shala No. 17	0.46206529	T
10	118	Babajipura Mishra Shala No. 18	0.18134939	T
11	119	Babajipura Mishra Shala No. 19	0.69772005	C
12	120	Babajipura Mishra Shala No. 20	0.83992642	C
13	121	Babajipura Mishra Shala No. 21	0.85501349	C
14	125	Babajipura Kumar Shala No. 25	0.30572517	T
15	126	Babajipura Kanya Shala No. 26	0.53388093	C
16	127	Babajipura Mishra Shala No. 27	0.46003571	T
17	128	Babajipura Mishra Shala No. 28	0.27464658	T
18	130	Babajipura Hindi Mishra Shala No. 30	0.02073858	T
19	131	Babajipura Mishra Shala No. 31	0.77709404	C
20	132	Babajipura Mishra Shala No. 32	0.2362122	T
21	201	Fatehpura Kumar Shala No. 1	0.91552715	C
22	202	Fatehpura Mishra Shala No. 2	0.95669543	C
23	209	Fatehpura Mishra Shala No. 9	0.48508217	T
24	210	Fatehpura Kanya Shala No. 10	0.62054343	C
25	211	Fatehpura Mishra Shala No. 11	0.17807564	T
26	213	Fatehpura Kumar Shala No. 13	0.36389518	T
27	215	Fatehpura Hindi Mishra Shala No. 15	0.03446481	T
28	216	Fatehpura Mishra Shala No. 16	0.51526826	T
29	218	Fatehpura Mishra Shala No. 18	0.17860571	T
30	219	Fatehpura Mishra Shala No. 19	0.04501407	T
31	301	N. Sayajiganj Mishra Shala No. 1 (center)	0.93881649	C

Partie 2: L'assignation aléatoire stratifiée

La stratification est le fait de diviser l'échantillon en plusieurs sous-groupes et ensuite d'assigner les individus au sein de chaque sous-groupe aux groupes traitement et contrôle. On fait souvent cela pour s'assurer que les sous-groupes sont bien représentés au sein des groupes traitement et contrôle. Ceci facilite les analyses futures par sous-groupe.

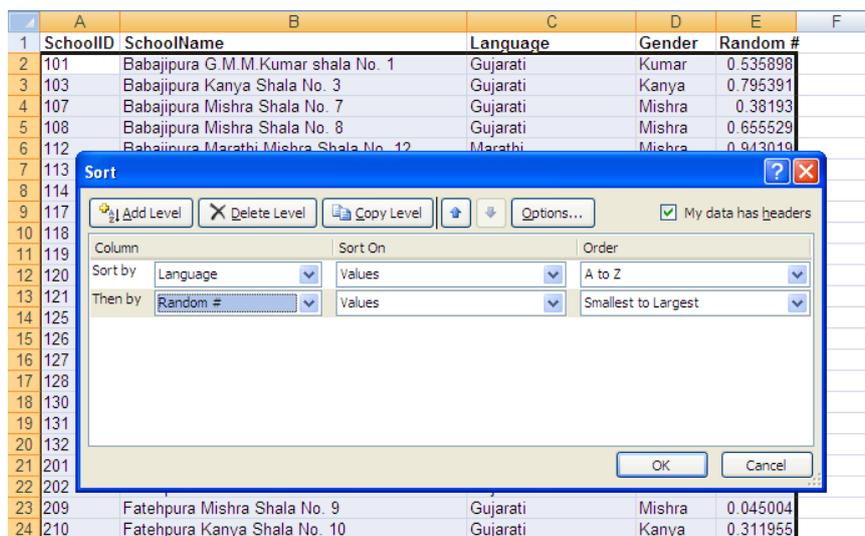
Par exemple, si vous souhaitez tester l'efficacité d'un nouveau programme d'éducation séparément pour les écoles dans lesquelles l'enseignement se fait en Hindi et celles où il se fait en Gujarati, vous pouvez stratifier par la variable "Langage d'instruction" afin d'assurer qu'il y ait un nombre identique d'écoles de chaque langage au sein des groupes traitement et contrôle.

1. Reprendre la liste des écoles dans l'onglet "stratifié"

Ce fichier comporte les variables de stratifications potentielles. Mécaniquement, la seule différence dans le tirage aléatoire est qu'au lieu de simplement trier par le nombre aléatoire, vous allez d'abord trier par langage, et ensuite par le nombre aléatoire. Bien évidemment, la première étape est de s'assurer que vous avez les variables par lesquelles vous souhaitez stratifier.

2. Trier par la variable de stratification et ensuite par le nombre aléatoire.

Supposons que vous avez toutes les variables dont vous avez besoin : dans la base, cliquer sur « Trier ». La fenêtre "Tri" apparaît. Trier par « Langage » puis cliquer sur « Ajouter un niveau » et sélectionner "Random #".



	A	B	C	D	E	F
1	SchoolID	SchoolName	Language	Gender	Random #	
2	101	Babajipura G.M.M.Kumar shala No. 1	Gujarati	Kumar	0.535898	
3	103	Babajipura Kanya Shala No. 3	Gujarati	Kanya	0.795391	
4	107	Babajipura Mishra Shala No. 7	Gujarati	Mishra	0.38193	
5	108	Babajipura Mishra Shala No. 8	Gujarati	Mishra	0.655529	
6	112	Babajipura Marathi Mishra Shala No. 12	Marathi	Mishra	0.943019	
7	113					
8	114					
9	117					
10	118					
11	119					
12	120					
13	121					
14	125					
15	126					
16	127					
17	128					
18	130					
19	131					
20	132					
21	201					
22	202					
23	209	Fatehpura Mishra Shala No. 9	Gujarati	Mishra	0.045004	
24	210	Fatehpura Kanya Shala No. 10	Gujarati	Kanya	0.311955	

The 'Sort' dialog box is open, showing the following settings:

- Sort by: Language (Values, A to Z)
- Then by: Random # (Values, Smallest to Largest)
- My data has headers:

3. Assigner les statuts de Traitement et Contrôle pour chaque groupe

Au sein de chaque groupe de langue, taper "T" pour la première moitié des lignes et « C » pour la seconde.

	A	B	C	D	E	F
100	132	Babajipura Mishra Shala No. 32	Gujarati	Mishra	0.8931975	C
101	615	Wadi Mishra Shala No. 15	Gujarati	Mishra	0.9142383	C
102	618	Wadi Kumar Shala No. 18	Gujarati	Kumar	0.9229356	C
103	408	Raopura Kanya Shala No. 8	Gujarati	Kanya	0.9285077	C
104	502	Shaher Vibhag Mishra Shala No. 2	Gujarati	Mishra	0.9549163	C
105	311	Sayajiganj Mishra Shala No. 11	Gujarati	Mishra	0.9595266	C
106	344	Sayajiganj Mishra Shala No. 44	Gujarati	Mishra	0.9688854	C
107	347	Sayajiganj Hindi Mishra Shala No. 47	Hindi	Mishra	0.0163449	T
108	332	Sayajiganj Hindi Mishra Shala No. 32	Hindi	Mishra	0.1528766	T
109	342	Sayajiganj Hindi Mishra Shala No. 42	Hindi	Mishra	0.2646791	T
110	215	Fatehpura Hindi Mishra Shala No. 15	Hindi	Mishra	0.3142377	T
111	326	Sayajiganj Hindi Mishra Shala No. 26	Hindi	Mishra	0.4291559	T
112	638	Wadi Hindi Mishra Shala No. 38	Hindi	Mishra	0.6772441	C
113	130	Babajipura Hindi Mishra Shala No. 30	Hindi	Mishra	0.7053783	C
114	315	Sayajiganj Hindi Mishra Shala No. 15	Hindi	Mishra	0.7955641	C
115	626	Wadi Hindi Mishra Shala No. 26	Hindi	Mishra	0.8918818	C
116	346	Sayajiganj Hindi Mishra Shala No. 46	Hindi	Mishra	0.9051467	C
117	303	N. Sayajiganj Marathi Mishra Shala No. 3	Marathi	Mishra	0.0354843	T
118	523	Shaher Vibhag Marathi Mishra Shala No. 23	Marathi	Mishra	0.1834626	T
119	409	Raopura Marathi Mishra Shala No. 9	Marathi	Mishra	0.7676874	T
120	611	Wadi Marathi Mishra Shala No. 11	Marathi	Mishra	0.8847497	T
121	329	Sayajiganj Marathi Mishra Shala No. 29	Marathi	Mishra	0.8992905	C
122	112	Babajipura Marathi Mishra Shala No. 12	Marathi	Mishra	0.9430188	C
123	327	Sayajiganj Marathi Mishra Shala No. 27	Marathi	Mishra	0.9515261	C
124	617	Wadi Marathi Mishra Shala No. 17	Marathi	Mishra	0.9648498	C

On aurait également pu stratifier par genre, ou par genre et langue.

EXERCICE 2

La taille de l'échantillon

Avec cet exercice, nous allons explorer visuellement des échantillons aléatoires de différentes tailles pour une population donnée. Nous allons notamment tenter de démontrer que les tailles d'échantillon les plus importantes reflètent mieux la population étudiée.

1. Ouvrir le fichier "Exercice 1_samplingDistribution_NEW.xlsm".
2. Si nécessaire, sélectionnez "Autoriser les macros" (Enable Macros).
3. Naviguer dans la feuille "Randomize" qui vous permet de choisir un échantillon aléatoire de taille "Sample Size" à partir des données de la feuille "Control".
4. Entrez "10" pour "Sample Size" et cliquez sur "Randomize". Observez la distribution des différentes caractéristiques au sein des groupes Test, Témoin et au sein de la population globale étudiée (distribution attendue). Avec un échantillon de si petite taille, pour les résultats en lecture, la différence entre les moyennes observées dans notre échantillon et la moyenne attendue est assez élevée. Cliquez sur "Randomize" plusieurs fois et regardez les changements de répartition.
5. Maintenant, essayez "50" pour la taille de l'échantillon. Que se passe-t-il ? Tirer au sort plusieurs fois et regardez la différence des pourcentages des résultats en lecture.

Augmentez la taille de l'échantillon à "500", puis "2000" et "10 000", et regardez les modifications. Que constatons-nous lorsque l'échantillon est plus important ? Comment cela influence-t-il notre groupe Test et notre groupe Témoin ? La différence entre les groupes Test, Témoin et la population globale étudiée pour les variables observées diminue-t-elle toujours lorsque nous augmentons la taille de l'échantillon ?

EXERCICE 3

Calcul de la taille d'un échantillon avec un protocole en grappes

Dans l'étude de cas des Panchayats, votre équipe d'évaluation s'était appliquée à mesurer l'effet d'un *programme (traitement)* de réservation de sièges pour les femmes au sein des conseils de villages, sur les résultats au niveau du village et du foyer.

Toutefois, le tirage au sort des chefs des Conseils femmes a été effectué au niveau du Gram Panchayat et donc sur une grappe de villages.

Il est possible que l'impact observé du programme soit corrélé pour des villages appartenant au même Panchayat.

Par exemple, tous les villages d'un Panchayat vont subir les mêmes précipitations et autres chocs économiques. Cela signifie que lorsqu'un village du Panchayat s'en sort particulièrement bien pour cette raison aléatoire (par exemple, le choc d'une bonne précipitation), il en sera de même pour tous les villages de ce Panchayat.

Cela va conduire à plus de bruit et donc à une erreur type plus grande que dans le cas habituel de l'échantillonnage indépendant : en effet, nous avons moins de variations que ce que nous pensons. Nous devons en tenir compte lorsque nous planifions aussi bien la taille de l'échantillon que la meilleure façon d'échantillonner les villages et les Panchayats.

Cet exercice vous aidera à comprendre comment y parvenir. Devez-vous échantillonner tous les villages d'un Gram Panchayat ? Devez-vous échantillonner 1 village parmi 100 Gram Panchayats ? Ou bien devez-vous échantillonner 3 villages parmi 40 Panchayats ? Comment décider ?

Nous étudierons ces questions en déterminant la taille de l'échantillon qui nous permet de détecter un effet spécifique ayant une puissance d'au moins 80 %. Pour mémoire, la puissance est la probabilité que l'on puisse détecter l'effet du programme.

Cet exercice vous montre de quelle manière la puissance de votre échantillon change avec le nombre de grappes, la taille des grappes, la taille de l'effet attendu du programme et le coefficient de corrélation intra-classe. Nous utiliserons un logiciel créé par Steve Raudebush avec l'aide financière de la Fondation William T. Grant. Vous trouverez des informations supplémentaires sur les plans en grappes sur leur site Internet.

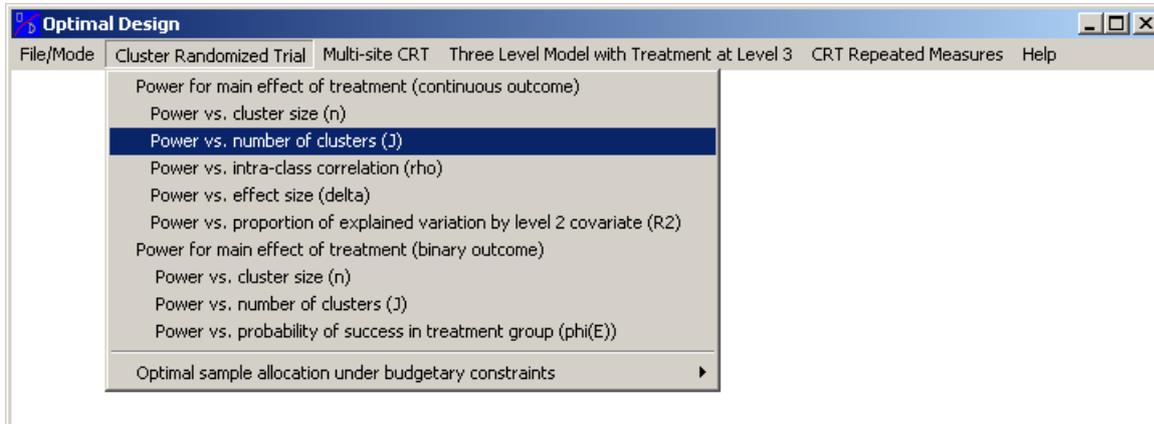
Partie 1 : Utilisation du logiciel de développement organisationnel (DO)

Tout d'abord, téléchargez le logiciel DO à partir du site Internet (un manuel pour ce logiciel est également disponible) ou à partir de la clé USB. que nous vous avons remise à votre arrivée.

http://sitemaker.umich.edu/group-based/optimal_design_software

Après le téléchargement, ouvrez le fichier exécutable et vous verrez un écran identique à celui ci-dessous. Sélectionnez l'option du menu : Cluster Randomized Trial (Étude aléatoire en grappes).

Sous ce menu, vous trouverez plusieurs options permettant de créer des graphiques.



Sélectionnez l'option Power vs number of clusters (J) (Puissance par rapport au nombre de grappes (J)). Une autre fenêtre de menu s'ouvre :

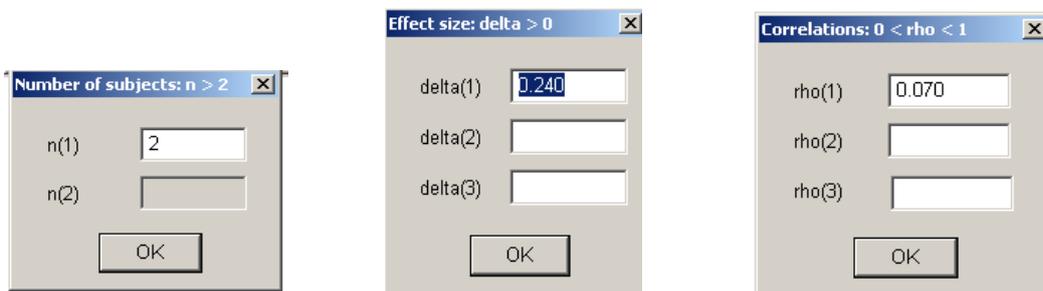


Sélectionnez α (alpha). Vous constaterez qu'il est déjà défini à 0,050 pour un niveau de signification de 95 %.

Cliquez ensuite sur n , le nombre de villages par grappe. Supposons que vous souhaitez savoir quel type de puissance vous pouvez obtenir si vous échantillonnez 2 villages à partir de chaque Gram Panchayat. Remplissez $n(1)$ par le chiffre 2 et cliquez sur OK.

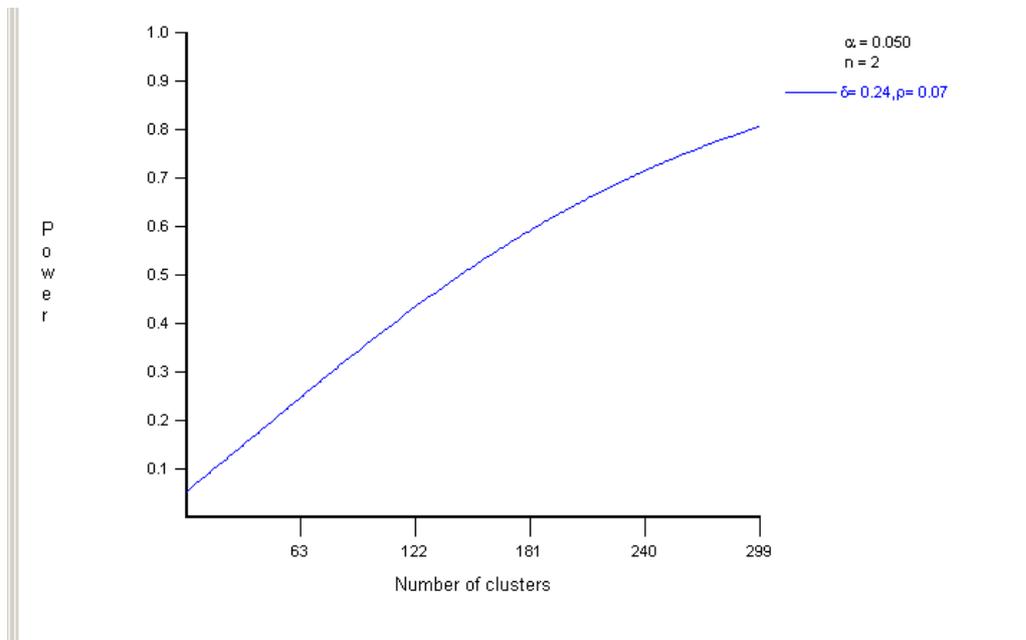
Nous devons maintenant déterminer δ (delta), l'ampleur de l'effet standardisé (l'ampleur de l'effet divisée par l'écart type de la variable d'intérêt). Supposons que vous souhaitez savoir s'il existe une augmentation de 30 % de l'investissement en eau potable dans les Gram Panchayats réservés aux femmes sur deux ans. Dans une petite enquête pilote, vous avez déterminé que, dans les Panchayats qui ne sont pas réservés aux femmes, il existe en moyenne 15 exemples d'investissement ces deux dernières années, et que l'écart type du nombre d'investissements est de 19. Nous souhaitons détecter une ampleur d'effet de 30 % de 15, ce qui correspond à 4,5. Nous divisons ensuite ce nombre par l'écart type pour obtenir δ égal à $4,5/19$, soit 0,24.

Sélectionnez δ dans le menu. Dans la boîte de dialogue qui apparaît, la valeur de delta(1) est prédéfinie à 0,200. Remplacez cette valeur par 0,24, et laissez vide la case correspondant à delta (2). Cliquez sur OK.

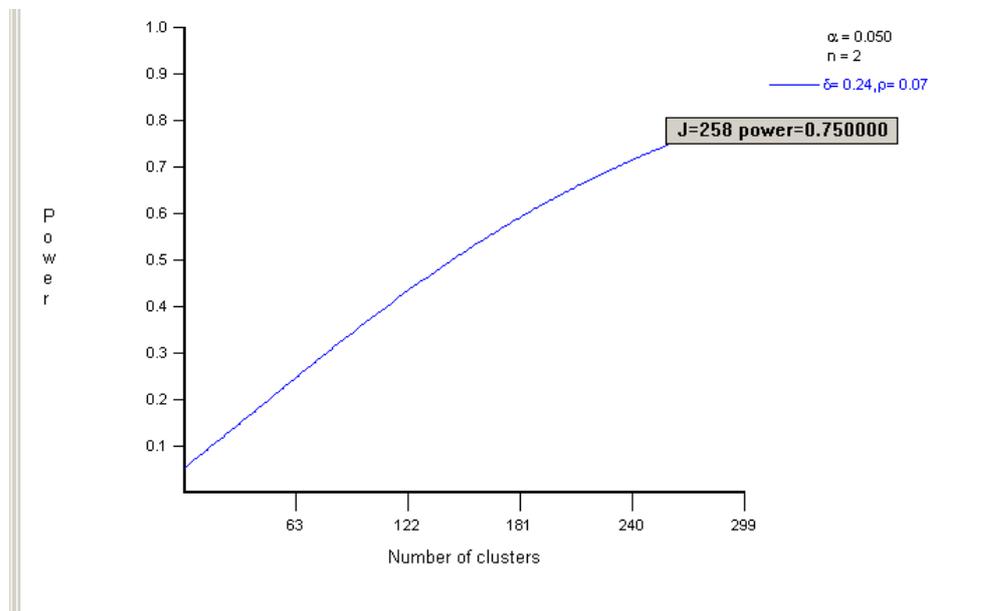


Nous devons enfin choisir rho ρ qui représente le *coefficient de corrélation intra-classe*. Rho nous renseigne sur l'importance de la corrélation des résultats pour les unités d'une même grappe. Si, en fait, les villages au sein d'un Gram Panchayat sont totalement indépendants par rapport à la variable d'intérêt, dans le cas présent, l'eau potable, alors rho ρ vaudra 0. Si, au contraire, les villages du Gram Panchayat sont identiques (aucune variation) par rapport aux résultats de l'eau potable, alors rho ρ vaudra 1. Vous avez déterminé dans votre enquête pilote que rho ρ vaut 0,07. Remplissez rho(1) par 0,07 et laissez rho(2) vide.

Vous devez voir apparaître un graphique semblable à celui ci-dessous. De quelle manière le nombre de grappes modifie-t-il la puissance de l'échantillon ? Combien de grappes devez-vous échantillonner pour obtenir une puissance de 80 % ?



Vous pouvez cliquer sur le graphique avec la souris pour voir la puissance exacte et le nombre exact de grappes pour un point particulier :



Vous avez vu combien de grappes sont nécessaires pour une puissance de 80 %, en échantillonnant deux villages par Panchayat. Supposons maintenant que vous souhaitez échantillonner 5 villages de chaque Gram Panchayat.

Remplacez n par 5. À quoi ressemble le nouveau graphique ? Quel est le nouveau nombre de grappes nécessaires pour une puissance de 80 % ?

Enfin, voyons de quelle façon le coefficient de corrélation intra-classe (ρ) modifie la puissance d'un échantillon donné.

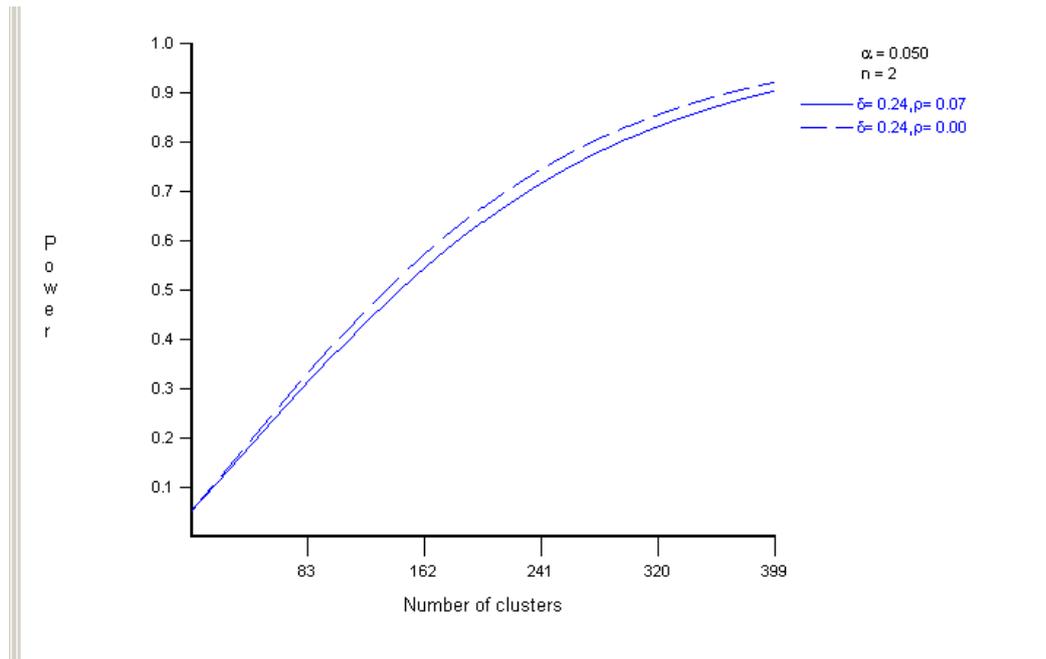
Laissez $\rho(1)$ à 0,07, mais pour comparer, remplacez $\rho(2)$ par 0,00.

Vous devez voir apparaître un graphique semblable à celui de la page suivante.

- La courbe bleue en trait plein représente les paramètres que vous avez définis, sur la base de vos estimations de pré-enquête, de l'effet des quotas pour les femmes sur l'eau potable.

- La courbe bleue en pointillés permet de comparer, pour voir quelle puissance vous pouvez obtenir à partir de votre échantillon si ρ est égal à 0

- Observez attentivement le graphique. De quelle manière la puissance de l'échantillon change-t-elle avec le coefficient de corrélation intra-classe (ρ) ?



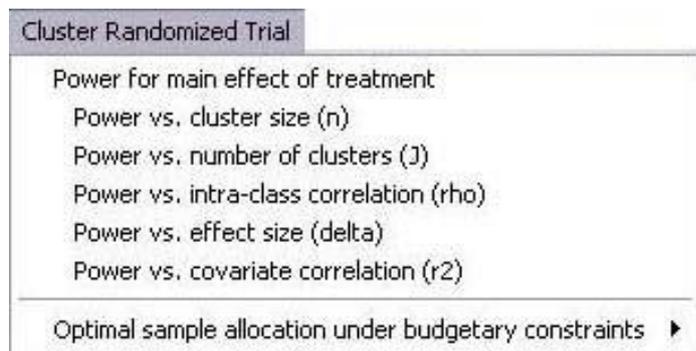
Vous pouvez redimensionner les axes x et y à l'aide des options du menu :



Partie 2 : Autres options du menu Étude aléatoire en grappes

Pour visualiser les autres options du menu, fermez le graphique en cliquant sur **X** dans le coin en haut à droite de la fenêtre interne. Sélectionnez à nouveau le menu Etude aléatoire en grappes.

Essayez de créer des graphiques pour voir de quelle manière la puissance change avec la taille des grappes (n), le coefficient de corrélation intra-classe (ρ) et l'ampleur de l'effet (δ). Vous devrez de nouveau saisir vos paramètres de pré-enquête chaque fois que vous ouvrirez un nouveau graphique.



BIBLIOGRAPHIE

Papiers à lire avant le début de la formation

Banerjee, Abhijit, Shawn Cole, Esther Duflo, et Leigh Linden, "Remedying Education: Evidence from Two Randomized Evaluations in India," *Quarterly Journal of Economics*, Vol. 122, No. 3, (Août 2007), pp. 1235-1264.

<http://www.mitpressjournals.org/doi/abs/10.1162/qjec.122.3.1235> ou

<http://www.povertyactionlab.org/publication/remedying-education-evidence-two-randomized-experiments-india>

Duflo, Esther, Michael Kremer, et Jonathan Robinson, "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya," Poverty Action Lab Working Paper

<http://www.povertyactionlab.org/publication/how-high-are-rates-return-fertilizer-evidence-field-experiments-kenya>

Duflo, Esther, Pascaline Dupas, et Michael Kremer, "Peer Effects and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya," Poverty Action Lab Working Paper Series

<http://www.povertyactionlab.org/publication/peer-effects-teacher-incentives-and-impact-tracking-evidence-randomized-evaluation-kenya>

Duflo, Esther, Rema Hanna, Stephen Ryan, "Monitoring Works: Getting Teachers to Come to School," Center for Economic Policy Research, Discussion Paper no. 6682, (Février 2008).

<http://www.povertyactionlab.org/publication/incentives-work-getting-teachers-come-school>

Karlan, Dean et Jonathan Zinman, "Observing unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment," (à paraître dans *Econometrica*)

http://karlan.yale.edu/p/OU_deco8_v1.pdf

Kling, Jeffrey R., Jeffrey B. Liebman et Lawrence F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75 (Janvier 2007), 83-119.

Bibliographie générale

Banerjee, Abhijit, "Inside the Machine: Toward a New Development Economics," Boston Review, (Mars/Avril 2007).

Banerjee, Abhijit V. et Esther Duflo, "The Experimental Approach to Development Economics," *Annual Review of Economics*, (2009).

Banerjee, Abhijit and Esther Duflo, "Giving Credit where it is Due," Unpublished Paper.

Deaton, Angus, "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development," NBER Working Paper, No. w14690, (Janvier 2009).

Duflo, Esther “Field Experiments in Development Economics,” *Advances in Economic Theory and Econometrics*, Eds. Richard Blundell, Whitney Newey, Torsten Persson, Cambridge University Press, Volume 2(42), voir aussi BREAD Policy Paper No. 002, 2005.

Duflo, Esther , Rachel Glennerster, Michael Kremer, “Using Randomization in Development Economics Research: A Toolkit” *Handbook of Development Economics*, Volume 4 (2008).

Imbens, Guido “Better LATE than Nothing,” NBER Working Paper No. w14896, (Avril 2009).

Imbens, Guido et Jeffrey M. Wooldridge, “Recent Developments in the Econometrics of Program Evaluation,” *Journal of Economic Literature*, Vol. 47, No. 1, (Mars 2009), pp. 5-86.

Karlan, Dean, and Jonathan Zinman. 2009. “Expanding Microenterprise Credit Access: Using Randomized Supply Decisions to Estimate the Impacts in Manila.” *Innovations in Poverty Action and Financial Access Initiative Working Paper*.

- Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, et Cynthia Kinnan. 2009. “The Miracle of Microfinance? Evidence from a Randomized Evaluation.” *MIT Department of Economics*.
- McKenzie, David, S. de Mel et C. Woodruff, 2009. “Are Women More Credit Constrained? Experimental Evidence on Gender and Microenterprise Returns
<http://siteresources.worldbank.org/DEC/Resources/Are_Women_More_Credit_Constrained.pdf>
”, *American Economic Journal: Applied Economics*, à paraître.
- Behrman, Jere, Susan Parker, et Petra Todd. 2009. “Medium Term Impacts of the Oportunidades Conditional Cash Transfer Program on Rural Youth in Mexico” et, dans Stephan Klasen et Felicitas Nowak-Lehmann, Eds., *Poverty, Inequality and Policy in Latin America*, Cambridge , MA: MIT Press.
- Baird, Sarah, Joan Hamory, et Edward Miguel, 2008. “Tracking, Attrition, and Data Quality in the Kenyan Life Panel Survey Round 1 (KLPS-1), mimeo, George Washington University and UC Berkeley.
- “Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment” (avec Norbert Schady et Renos Vakis), *American Economic Journal: Applied Economics*, à paraître.
- Edward Miguel et Michael Kremer, “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities,” *Econometrica* 72(1): 159-217, 2004

Liste de contrôle
pour vérifier la validité et les résultats
d'une évaluation aléatoire
d'un projet ou programme social



Février 2010

Cette publication a été produite par la coalition pour des stratégies fondées sur des preuves [[Coalition for Evidence-Based Policy](#)], avec l'aide financière des William T. Grant Foundation, Edna McConnell Clark Foundation et Jerry Lee Foundation.

Cette publication appartient au domaine public. Sa reproduction à des fins pédagogiques, en tout ou en partie, est autorisée.

Vous pouvez nous envoyer vos commentaires et vos suggestions à propos de ce document (jbaron@excelgov.org).

Liste de contrôle pour vérifier la validité et les résultats d'une évaluation aléatoire d'un projet ou programme social.

Cette liste de contrôle répertorie les éléments-clés à rechercher lors de la lecture des résultats d'une évaluation aléatoire d'une stratégie, d'un projet ou d'un programme social (« intervention »), afin d'évaluer s'il a produit des preuves valides de l'efficacité de l'intervention. Cette liste de contrôle suit étroitement les directives du Bureau de l'administration et du budget américain [U.S. Office of Management and Budget], ou OMB, et de l'Institut des sciences de l'éducation du Ministère de l'éducation américainⁱⁱ [U.S. Education Department's Institute of Education Sciences], ou IES ; Cependant, les opinions exprimées dans le présent document ne reflètent pas nécessairement les opinions de l'OMB ou de l'IES.

Cette liste de contrôle se limite aux éléments-clés et ne tente pas de traiter toutes les éventualités pouvant avoir un impact sur la validité des résultats d'une étude. Elle est conçue pour aider à se faire une opinion éclairée, et non pour s'y substituer ; laquelle peut s'avérer nécessaire, par exemple pour juger si un écart par rapport à un ou plusieurs des éléments de la liste de contrôle est suffisamment grave pour affecter les résultats de l'étude.

Une courte annexe traite du *nombre* d'essais sur échantillon aléatoire et contrôlé, bien conçus, qui sont nécessaires pour produire des preuves solides de l'efficacité d'une intervention.

Liste de contrôle pour la conception générale de l'étude

L'affectation aléatoire a été réalisée au niveau appropriée : groupes (par ex. classes, projets d'habitation), individus (par ex. élèves, locataires) ou les deux.

L'affectation aléatoire d'individus représente généralement l'approche la plus efficace et la plus économique. Cependant, il peut s'avérer nécessaire d'affecter des groupes de façon aléatoire (à la place ou en plus des individus) afin d'évaluer (i) les interventions pouvant avoir d'importants effets de réaction en chaîne ou de propagation aux non-participants, et (ii) les interventions effectuées auprès de groupes entiers, comme des classes, des projets d'habitation ou des communautés. (Voir référence 2 pour de plus amples détails).ⁱⁱⁱ

La taille de l'échantillon de l'étude est adéquate, c'est-à-dire suffisamment importante pour détecter les effets significatifs de l'intervention.

La taille adéquate d'un échantillon dépend des caractéristiques spécifiques de l'intervention, de l'échantillon de la population et de la conception de l'étude, comme cela est expliqué par ailleurs.^{iv} Voici deux éléments qui peuvent vous aider à juger si l'étude que vous lisez dispose d'un échantillon d'une taille adéquate :

- Si les résultats de l'étude indiquent que l'intervention a produit des effets *statistiquement significatifs* (comme cela est expliqué ultérieurement dans cette liste

de contrôle), on peut probablement considérer que l'échantillon était suffisamment important.

- Si les résultats de l'étude indiquent que l'intervention n'a pas produit d'effets statistiquement significatifs, le rapport sur l'étude devrait inclure une analyse montrant que l'échantillon était suffisamment important pour détecter les effets significatifs de l'intervention. (une analyse de ce type s'appelle une « analyse de puissance »).v

La référence 5 contient des exemples indicatifs des tailles d'échantillons provenant d'essais sur échantillon aléatoire et contrôlé bien conçus, effectués dans divers secteurs des politiques sociales.vi

Liste de contrôle pour s'assurer que les groupes d'intervention et de référence sont restés équivalents pendant l'étude

Le rapport d'étude inclut une analyse montrant qu'il existe peu ou aucune différence systématique entre les groupes d'intervention et de référence avant l'intervention (par ex., en matière d'âge, de sexe, de revenus, d'éducation).

Peu ou aucun membre du groupe de référence n'a participé à l'intervention ou bénéficié de l'intervention autrement (c'est-à-dire que le recoupement ou la contamination des membres du groupe de référence a été minimal).

L'étude a recueilli les résultats de la même façon et au même moment pour les membres des groupes d'intervention et de référence.

L'étude a obtenu des résultats pour une proportion élevée des membres de l'échantillon initialement choisis de façon aléatoire (c'est-à-dire que la déperdition de l'échantillon de l'étude est faible).

Globalement, une étude devrait obtenir des résultats pour au moins 80 % des membres de l'échantillon initialement choisis de façon aléatoire, y compris les membres affectés au groupe d'intervention qui n'ont pas participé ou qui n'ont pas terminé l'intervention. En outre, le taux de suivi doit être plus ou moins le même pour les groupes d'intervention et de référence.

Le rapport d'étude devrait inclure une analyse montrant que la déperdition de l'échantillon (le cas échéant) n'a pas affecté l'équivalence des groupes d'intervention et de référence.

L'étude, dans son évaluation des effets de l'intervention, a maintenu les membres de l'échantillon dans le groupe original auquel ils avaient été affectés de façon aléatoire.

Cela s'applique même aux :

- Membres du groupe d'intervention qui n'ont pas participé ou qui n'ont pas terminé l'intervention (leur maintien dans le groupe d'intervention est conforme à l'approche du type « vouloir traiter ») ; et
- Membres du groupe de référence qui ont pu participer à l'intervention ou en

bénéficiaire (c'est à dire les membres « recoupés » ou « contaminés » appartenant au groupe de référence).^{vii}

Liste de contrôle pour les mesures du résultat de l'étude

L'étude utilise des mesures de résultats « valides », c'est-à-dire des mesures extrêmement corrélées aux résultats réels que l'intervention cherche à affecter.

Par exemple :

- Les tests que l'étude a utilisés pour mesurer les résultats (par ex. les tests de niveau d'instruction ou de bien-être psychologique) sont des tests dont l'aptitude à mesurer les résultats réels est bien établie.
- S'il a été demandé aux membres de l'échantillon de fournir eux-mêmes les résultats (par ex. comportement criminel), leurs rapports ont été corroborés par des mesures indépendantes et/ou objectives dans la mesure du possible (par ex. rapports de police).
- Les mesures des résultats n'ont pas favorisé le groupe d'intervention par rapport au groupe de référence ou vice-versa. Par exemple, une étude d'un programme informatisé ayant pour but d'enseigner les mathématiques aux jeunes élèves ne devrait pas mesurer les résultats à l'aide d'un test informatisé, car le groupe d'intervention aura probablement de plus grandes facilités à utiliser un ordinateur que le groupe de référence.^{viii}

L'étude a mesuré les résultats qui sont importants sur le plan stratégique ou pratique, pas seulement des résultats intermédiaires pouvant prévoir ou non des résultats importants.

Quelques illustrations : (i) l'étude d'un programme de prévention des grossesses devrait mesurer les résultats tels que les grossesses réelles et pas seulement l'attitude des participants vis-à-vis du sexe ; et (ii) l'étude d'un programme de rééducation des dyslexiques devrait mesurer des résultats tels que la compréhension et la maîtrise de la lecture et pas seulement la capacité à déchiffrer les mots.

Quand cela le justifiait, les membres de l'équipe chargée de l'étude qui ont recueilli les résultats ont fait cette tâche « à l'aveugle », c'est-à-dire qu'ils ne savaient pas qui appartenait au groupe d'intervention et au groupe de référence.

Cette approche « à l'aveugle » est importante quand l'étude mesure des résultats par le biais d'entrevues, de tests ou d'autres instruments qui ne sont pas totalement structurés, laissant à la personne qui effectue les mesures une marge de jugement subjectif. L'approche « à l'aveugle » protège contre la partialité ou le parti pris de la personne qui effectue les mesures (par ex. en tant que promoteur de l'intervention) et l'influence que cela pourrait avoir sur ses mesures des résultats. L'approche « à l'aveugle » est

importante, par exemple, quand une étude mesure l'incidence des coups sur le terrain de jeu par le biais de l'observation du terrain de jeu ou quand une étude mesure les capacités d'identification des mots des jeunes élèves par le biais de tests administrés personnellement.

L'étude a privilégié les données sur les résultats à long terme de l'intervention (par ex. un an après la fin de l'intervention, voire même plus longtemps).

Cela permet aux spécialistes et aux responsables du développement des stratégies de juger de la durabilité des effets de l'intervention dans le temps. Dans la plupart des cas, ce sont les effets à long terme, plutôt que les effets immédiats, qui sont les plus importants sur le plan pratique et stratégique.

Liste de contrôle pour le rapport sur les effets de l'intervention de l'étude

Si l'étude conclut que l'intervention a un effet sur les résultats, le rapport comprend (i) la dimension de l'effet et si la dimension de l'effet est d'importance stratégique ou pratique ; et (ii) les tests montrant que l'effet est statistiquement significatif (c'est-à-dire qu'il est peu probable qu'il soit dû au hasard).

Ces tests, visant à prouver que les résultats sont statistiquement significatifs, doivent prendre en compte les caractéristiques-clés de la conception de l'étude, notamment :

- Whether Si les individus (par ex. élèves) ou les groupes (par ex. classes) ont été affectés de façon aléatoire ;
- i l'échantillon a été trié en groupes avant l'affectation de façon aléatoire (c'est-à-dire « stratifié », « bloqué » ou « associé ») ; et
- Si l'étude prévoit que les évaluations de l'effet de l'intervention ne s'appliquent qu'aux sites (par ex. projets d'habitation) de l'étude ou si elles sont généralisables à une population plus large.

S

Le rapport d'étude inclut les effets de l'intervention sur tous les résultats que l'étude a mesurés, pas seulement ceux pour lesquels l'effet est positif.

Cela permet de juger si les effets positifs représentent une exception ou une tendance régulière.

Annexe : Combien d'essais à échantillon aléatoire et contrôlé sont nécessaires pour produire des preuves solides d'efficacité ?

Pour avoir une assurance solide qu'une intervention est efficace si elle est fidèlement reproduite, il faut généralement rechercher des preuves, en particulier dans les cas suivants :

- L'efficacité de l'intervention a été démontrée, par le biais d'essais à échantillon aléatoire et contrôlé bien conçus, sur plus d'un site de mise en œuvre.

Une telle démonstration peut être composée de deux essais ou plus, effectués sur des sites de mise en œuvre différents, ou d'un essai plus large sur plusieurs sites.

- Le ou les essais ont évalué l'intervention dans le cadre et les conditions d'une communauté réelle où le(s) dit(s) essai(s) doit / doivent normalement être mis en œuvre (par ex. des cliniques communautaires de lutte contre la toxicomanie, des écoles publiques, des sites de programmes de formation professionnelle).

Cela est fait par opposition aux conditions extrêmement contrôlées que l'on trouve sur les sites spécialisés que les chercheurs mettent en place dans une université à des fins d'étude ou dans les cadres dans lesquels les chercheurs eux-mêmes se chargent de l'intervention.

- Il n'y a pas d'évidence contraire, comme des essais à échantillon aléatoire et contrôlé de l'intervention montrant l'absence d'effets.

Références

ⁱ Les Conseils de Villages, appelés Gram Panchayats, forment l'unité de base du Panchayat Raj. Les Présidents des Conseils de Villages, élus par les membres du Conseil de Village, sont membres du Conseil Sous-Districale Conjoint (*panchayat samiti*). Tout en haut du système, on trouve le Conseil Districale (*zilla parishad*) constitué des Présidents de Conseils Conjoints.

ⁱⁱ U.S. Office of Management and Budget ou OMB (Bureau de l'administration et du budget américain), What Constitutes Strong Evidence of Program Effectiveness (Ce qui constitue des preuves solides d'efficacité d'un programme), http://www.whitehouse.gov/omb/part/2004_program_eval.pdf, 2004 ; U.S. Department of Education's Institute of Education Sciences ou IES (Institut des sciences de l'éducation du Ministère de l'éducation américain), Identifying and Implementing Educational Practices Supported By Rigorous Evidence (Identification et mise en œuvre de pratiques pédagogiques soutenues par des preuves rigoureuses), <http://www.ed.gov/rschstat/research/pubs/rigorousvid/index.html>, décembre 2003 ; What Works (Ce qui fonctionne), Clearinghouse of the U.S. Education Department's Institute of Education Sciences (Bureau central de l'Institut des sciences de l'éducation du Ministère de l'éducation américain), Key Items To Get Right When Conducting A Randomized Controlled Trial in Education (Éléments-clés pour les essais sur échantillon aléatoire et contrôlé dans le domaine de l'éducation), préparé par la Coalition for Evidence-Based Policy (Coalition pour des stratégies fondées sur des preuves), http://www.whatworkshelpdesk.ed.gov/guide_RCT.pdf, 2005.

ⁱⁱⁱ L'affectation aléatoire des groupes, à la place ou en plus des individus, peut être nécessaire dans les situations suivantes (ou similaires) :

(a) L'intervention peut avoir d'importants effets de réaction en chaîne ou de propagation aux individus autres que ceux qui en bénéficient.

Par exemple, s'il existe de bonnes raisons de penser qu'un programme de lutte contre la toxicomanie destiné aux jeunes dans un projet d'habitation public produira une réduction importante de la consommation de drogues non seulement chez les participants au programme, mais également chez leurs pairs dans le projet d'habitation (par le biais de l'influence des pairs), il est probablement nécessaire d'affecter de façon aléatoire des projets d'habitation entiers aux groupes d'intervention et de référence afin de déterminer les effets du programme. Une étude qui n'affecterait de façon aléatoire que de jeunes individus du projet d'habitation au groupe d'intervention par rapport au groupe de référence sous-estimerait les effets du programme, dans la mesure où le programme réduit la consommation de drogues chez les jeunes du groupe d'intervention ET du groupe de référence du projet.

(b) L'intervention est effectuée pour des groupes, comme des classes ou des écoles (par ex. un programme de réforme des programmes à l'échelle de classes ou d'écoles entières), et l'étude essaie de distinguer les effets de l'intervention des effets des autres caractéristiques du groupe (par ex. la qualité du professeur).

Par exemple, dans le cadre de l'étude d'un nouveau programme de classe, les classes de l'échantillon seront généralement différentes de deux façons : (i) elles utilisent le nouveau programme ou non, et (ii) le professeur de chaque classe. Ainsi, si l'étude (par exemple) affecte de façon aléatoire des élèves individuels à deux classes qui utilisent le nouveau programme par rapport à deux classes qui ne l'utilisent pas, l'étude ne pourra pas distinguer les effets du nouveau programme des effets des autres caractéristiques des classes, comme la qualité des professeurs. Une telle étude devrait donc probablement affecter de façon aléatoire des professeurs et des classes dans leur totalité (un échantillon suffisant de chaque) aux groupes d'intervention et de

référence, afin de s'assurer que les deux groupes sont équivalents, non seulement en matière de caractéristiques des élèves mais aussi de caractéristiques des classes et des professeurs.

Pour des raisons similaires, une étude portant sur une réforme à l'échelle des écoles entières devra probablement affecter de façon aléatoire des écoles entières aux groupes d'intervention et de référence, afin de s'assurer que les deux groupes sont équivalents, non seulement en matière de caractéristiques des élèves mais aussi de caractéristiques des écoles (par ex. qualité des professeurs et taille moyenne des classes).

^{iv} What Works (Ce qui fonctionne), Clearinghouse of the U.S. Education Department's Institute of Education Sciences (Bureau central de l'Institut des sciences de l'éducation du Ministère de l'éducation américain), *Key Items To Get Right When Conducting A Randomized Controlled Trial in Education (Éléments-clés pour les essais sur échantillon aléatoire et contrôlé dans le domaine de l'éducation)*, op. cit., no. 1.

^v Les ressources pouvant être utiles pour passer en revue ou effectuer une analyse de puissance incluent : les services de conseils gratuits de la William T. Grant Foundation en matière de conception d'essais sur groupe échantillon aléatoire, à http://sitemaker.umich.edu/group-based/consultation_service ; Steve Raudenbush et coll., *Optimal Design Software for Group Randomized Trials (Logiciel de conception optimale pour les essais sur groupe échantillon aléatoire)*, à http://sitemaker.umich.edu/group-based/optimal_design_software ; Peter Z. Schochet, *Statistical Power for Random Assignment Evaluations of Education Programs (Puissance statistique pour les évaluations à affectation aléatoire des programmes pédagogiques)* (<http://www.mathematica-mpr.com/publications/PDFs/statisticalpower.pdf>), préparé pour U.S. Department of Education's Institute of Education Sciences ou IES (Institut des sciences de l'éducation du Ministère de l'éducation américain), 22 juin 2005 ; et Howard Bloom, *Randomizing Groups to Evaluate Place-Based Programs (Affectation aléatoire des groupes pour l'évaluation des programmes fondés sur les emplacements)*

(http://www.wtgrantfoundation.org/usr_doc/RSCChapter4Final.pdf), préparé pour une conférence de la Society for Research on Adolescence (Société de recherche sur l'adolescence), 2 mars 2004.

^{vi} Voici des exemples illustratifs de tailles d'échantillons provenant d'essais sur échantillon aléatoire et contrôlé bien conçus dans divers domaines des politiques sociales : (i) 4 028 demandeurs et bénéficiaires de l'aide sociale ont été affectés de façon aléatoire dans un essai du **Programme de formation en compétences élémentaires et opportunités professionnelles de Portland, Oregon (un programme de transition de l'aide sociale au travail)**, afin d'évaluer les effets du programme sur l'emploi et les revenus – voir <http://evidencebasedprograms.org/Default.aspx?tabid=157> ; (ii) de 400 à 800 femmes ont été affectées de façon aléatoire dans chacun des trois essais du Nurse-Family Partnership (un programme de visites d'infirmières à domicile destiné aux femmes enceintes à faibles revenus), afin d'évaluer les effets du programme sur une gamme de facteurs affectant la mère et l'enfant, comme la négligence ou la violence faite aux enfants, les arrestations criminelles et la dépendance à l'aide sociale – voir <http://evidencebasedprograms.org/Default.aspx?tabid=35> ; 206 élèves de seconde ont été affectés de façon aléatoire dans un essai « Check and Connect » (un programme de prévention de l'abandon des études pour les élèves à risques), afin d'évaluer les effets du programme sur l'abandon des études – voir <http://evidencebasedprograms.org/Default.aspx?tabid=163> ; 56 écoles comprenant près de 6 000 élèves ont été affectées de façon aléatoire dans un essai « LifeSkills Training » (un programme de prévention de la toxicomanie, de l'alcoolisme et du tabagisme), afin d'évaluer les effets du programme sur la consommation de drogues, d'alcool et de tabac des élèves

voir <http://evidencebasedprograms.org/Default.aspx?tabid=116>.

^{vii} L'étude, après obtention des évaluations des effets de l'intervention sur les membres de l'échantillon maintenus dans leurs groupes d'origine, peut parfois utiliser un ajustement prenant en compte ceux qui n'ont pas participé afin d'évaluer les effets sur les membres du groupe d'intervention qui ont véritablement participé à l'intervention. Une variante de cette technique peut parfois être utilisée pour prendre en compte le recoupement ou la contamination. Voir Larry L. Orr, *Social Experimentation: Evaluating Public Programs With Experimental Methods (Expérimentation sociale: Évaluation des programmes publics à l'aide de méthodes expérimentales)*, Sage Publications, Inc., 1999, p. 62 et 210 ; et Howard S. Bloom, "Accounting for No-Shows in Experimental Evaluation Designs" (Prendre les non-participants en compte dans les conceptions d'évaluations expérimentales), *Evaluation Review*, vol. 8, avril 1984, pp. 225-246.

^{viii} De façon similaire, l'étude d'un programme de prévention des crimes impliquant une supervision étroite par la police des participants au programme ne devrait pas utiliser les taux d'arrestation comme mesure des résultats, car cette supervision peut générer plus d'arrestations dans le groupe d'intervention.

GLOSSAIRE

Attrition

L'attrition est la diminution de la taille de l'échantillon que l'on suit au cours du temps due aux « perdus de vue ». Le taux d'attrition est le pourcentage des individus de l'échantillon initial qui sont absents de l'échantillon final.

Attriters = « Perdus de vue »

Certains sujets entrés dans l'étude ne peuvent, pour des raisons diverses, faire l'objet d'une évaluation à la fin de la période d'observation prévue par le protocole. Ces cas correspondent à ce que l'on appelle des perdus de vue. Par exemple, un élève inscrit dans une école participant à l'étude, mais qui est absent le jour de l'examen final, est « perdu de vue ».

Les perdus de vue peuvent entraîner un biais dans l'analyse des résultats. Exclure ces perdus de vue revient à supposer qu'ils se sont comportés comme les sujets suivis jusqu'au terme de l'étude. Une attitude raisonnable en l'absence complète d'information est de considérer tous les perdus de vue comme des échecs s'ils appartiennent au groupe traité et comme des succès, s'ils appartiennent au groupe témoin. Si le résultat est toujours favorable pour le nouveau traitement étudié, ce dernier est incontestablement efficace (théorie du biais maximal).

Biais (Bias)

Un biais est une erreur systématique présente dans une enquête et qui en fausse le résultat. Une telle erreur peut être due à la structure de l'étude, aux observateurs, aux techniques d'observation utilisées, aux individus observés ou au fait que l'échantillon étudié n'est pas représentatif de la population.

Biais de sélection (Selection Bias)

Il y a un biais de sélection dès lors que les sujets observés dans l'étude ne constituent pas un groupe représentatif ou un groupe adapté à la résolution des questions posées dans le protocole. Les biais de sélection se produisent lors de la formation des échantillons et sont liés aux critères d'inclusion des individus (ou unités statistiques) dans l'étude.

Les principaux biais de sélection sont :

- (1) le biais de recrutement qui apparaît lorsque la probabilité de faire entrer des sujets dans l'étude est liée à l'un ou plusieurs des facteurs étudiés
- (2) l'auto-sélection : il s'agit d'un biais qui peut se produire si les sujets entrent dans l'étude en fonction d'une décision qui leur est au moins en partie propre et qui peut être liée aux phénomènes étudiés. Ce type de biais se retrouve très fréquemment au niveau des non-répondants (ou des perdus de vue) : ainsi, si une personne refuse de répondre, ce peut être parce qu'elle se sent particulièrement concernée par le phénomène étudié ou que celui-ci l'empêche matériellement de répondre (maladie, décès,...).

Choc (Shock)

Un événement imprévu ou du au hasard, qui affecte un individu ou une unité d'observation. Par exemple, dans une étude sur la rentabilité d'une micro-entreprise commerciale, un choc peut-être une hausse soudaine de la demande pour le produit vendu par le commerçant étudié. Dans une étude sur les performances scolaires, un choc pour une école peut être la prise de retraite d'un « bon » directeur d'école et son remplacement par un « mauvais » directeur d'école.

Coefficient de variation (Coefficient of Variation)

Le coefficient de variation correspond à l'écart type d'une distribution exprimé en pourcentage de la moyenne de la distribution.

Comparabilité (Comparability)

Comparaison (Comparison)

Conformité (Compliance)

Un individu « compliant » est celui qui se soumet au programme ou aux enquêtes, c'est-à-dire accepte de se conformer au protocole de l'étude. Un individu non-« compliant » est celui qui refuse ou ne parvient pas à suivre le protocole de l'étude.

Coût-efficacité (Cost-effectiveness)

= rapport coût-efficacité ou analyse coût-bénéfice

Parmi les programmes ayant le même objectif, le programme qui a le meilleur rapport coût-bénéfice ou coût efficacité est celui qui atteint l'objectif au moindre coût.

DALY (disability adjusted life years)

= années de vie ajustées sur l'incapacité

Les DALY, ou années de vie ajustées sur l'incapacité, sont un indicateur composite de la charge de morbidité faisant intervenir le temps. Cet indicateur ajoute aux années perdues pour cause de décès prématuré les années de bonne santé perdues suite aux maladies et à l'incapacité associée.

Échantillon stratifié (Stratified sample)

Type d'échantillon qui résulte d'une étude préliminaire des caractéristiques de la population qui constituent autant de strates et d'une sélection au hasard des sujets qui font partie de ces strates.

Effet Hawthorne (Hawthorne Effect)

L'effet Hawthorne décrit la situation dans laquelle les résultats d'une expérience ne sont pas dus aux facteurs expérimentaux mais au fait que les sujets ont conscience de participer à une expérience dans laquelle ils sont testés, ce qui se traduit généralement par une plus grande motivation = Dans une recherche expérimentale, biais des participants du groupe expérimental qui améliorent leur performance du seul fait de savoir qu'ils sont soumis à la variable indépendante.

Effet John Henry (John Henry Effect)

Dans une recherche expérimentale, biais des participants du groupe témoin qui vont au-delà de leur rendement habituel parce qu'ils se savent comparés avec le groupe expérimental qui a été soumis au programme.

Enquête initiale (Baseline)

Peut aussi être utilisé pour : (...) de référence, ou, (...) de départ, ou, initial, par exemple baseline test = test de référence, test de départ, test initial...

Équilibre général (General Equilibrium)

Un équilibre général est un équilibre économique sur l'ensemble des marchés. L'équilibre général donne le niveau de production et de consommation des biens et les prix dans une économie.

Expérimentation (Experiment)

Externalité (externalities ou Spillovers)

= externalités, effets externes

L'externalité, ou effet externe, désigne une situation économique dans laquelle l'action d'un agent influe positivement ou négativement sur le bien-être d'un autre agent, sans que cette influence ne se traduise par une variation du niveau des prix.

Extrapoler (Extrapolate)

Utiliser les relations estimées dans l'échantillon étudié pour faire des prédictions sur les individus non-représentés dans l'échantillon d'étude.

Équilibre partiel (Partial Equilibrium)

Un équilibre partiel est un équilibre économique restreint à un seul marché. Un raisonnement en équilibre partiel suppose donc que l'on considère comme négligeables les effets induits sur le reste de l'économie par l'allocation des facteurs de production issue de l'équilibre sur le marché considéré.

Erreur de type I (Type I Error)

C'est la probabilité de rejeter l'hypothèse nulle alors qu'elle est vraie.

Erreur de type II (Type II Error)

C'est la probabilité de ne pas rejeter l'hypothèse nulle alors qu'elle est fausse.

Grappe (Cluster)

Analyse par grappes = Cluster analysis

Échantillonnage par grappes = Cluster sampling

Expérimentation aléatoire en grappe = Clustered randomized experiment

Par opposition aux évaluations aléatoires dans lesquelles des individus sont assignés aléatoirement à différents groupes d'intervention, dans une évaluation aléatoire par grappe les unités sociales (par exemple un village, une école, une région) sont assignées aléatoirement à différents groupes d'intervention.

Hypothèse alternative (Alternative hypothesis)

Hypothèse qui postule l'existence d'une différence entre A et B, donc $A > B$, $A < B$ ou $A \neq B$. Correspond à l'hypothèse de recherche, qui doit être prouvée vraie, en montrant au moyen de tests statistiques que l'hypothèse nulle ($A=B$) est fausse.

Hypothèse de recherche (Research Hypothesis)

Hypothèse mise à l'épreuve par la recherche. Ce que soutient le chercheur. Sur le plan statistique, correspond habituellement à [l'hypothèse alternative](#).

Hypothèse nulle (Null hypothesis)

Hypothèse qui postule que $A = B$. Dans un test, l'hypothèse nulle est supposée être vraie tant que les tests statistiques ne la rejettent pas.

Intervention / programme (Treatment)**ITT (Intention to Treat Estimation)**

= Analyse en intention de traitement

Dans une expérience randomisée contrôlée, l'analyse des résultats consiste à effectuer une comparaison entre l'efficacité des programmes dans les différents groupes de sujets. Une analyse en intention de traitement porte sur l'ensemble des sujets inclus dans l'essai, sans qu'il soit tenu compte du programme effectivement reçu. Ainsi, un sujet inclus à tort ou ayant quitté l'échantillon prématurément, devra quand même être pris en compte dans l'analyse.

L'ITT est utile car dans certains cas l'élimination d'observations de l'analyse finale risque d'entraîner une différence systématique entre les groupes comparés (biais de sélection), ce qui n'autorise plus à porter un jugement de causalité (c'est-à-dire à attribuer aux seuls effets des traitements les résultats observés).

Mise en œuvre (Implementation)

= exécution

Niveau de tirage au sort ou de randomisation (Level of randomization)

Par exemple, la randomisation peut être au niveau d'individuel (élève) ou de groupe (école)

Prospéctif (Prospective)

Par opposition aux études rétrospectives, l'étude prospective consiste à recueillir des informations originales et adaptées à l'objectif poursuivi sur un ou plusieurs groupes de sujets et de suivre le ou les groupes au cours du temps pour étudier une question prédéterminée. Les études prospectives sont généralement plus longues et difficiles à mettre en œuvre que les études rétrospectives ; elles ont par contre l'avantage de limiter les biais et les erreurs. Les études randomisées contrôlées sont typiquement des études prospectives.

Rejet/Rejeter (Reject)

C'est le résultat d'un test d'une hypothèse nulle où la puissance statistique nous permet de rejeter l'hypothèse nulle.

Résultat (Outcome)**Évaluation aléatoire (Randomized evaluation)**

On parle d'évaluation aléatoire ou avec répartition aléatoire des sujets, lorsque la comparabilité des groupes de sujets est assurée par une répartition aléatoire des programmes ou interventions. Une méthode de tirage au sort est en effet indispensable car c'est le seul moyen d'éviter des biais de sélection lors de la répartition des sujets entre les groupes.

Rétrospectif (Retrospective)

L'étude rétrospective cherche à établir des liens entre un phénomène observé et des événements antérieurs. Les informations utilisées dans une étude prospective sont donc constituées à partir de documents existants, avec leurs imperfections et leurs lacunes (par exemple, des enquêtes réalisées pour étudier un autre phénomène).

Dans le cas d'une évaluation d'impact, une étude rétrospective utilise des données déjà collectées pour évaluer un programme déjà réalisé.

Sélection de l'échantillon (Sample Selection)**Stratification (Stratification)**

La stratification est l'acte de grouper les observations en « strates » selon certaines caractéristiques observées avant de sélectionner un échantillon. La sélection de l'échantillon se fait ensuite en sélectionnant des individus au sein de chaque strate. Cela garantit l'équilibre et la représentativité de l'échantillon. Cette méthode est particulièrement importante lorsque la taille de l'échantillon à sélectionner est réduite.

Stratifier (Stratify)**Témoin (Control)****ToT (Treatment on Treated)**

= Analyse « sous traitement » ou *per protocol*

Analyse qui juge des effets d'un programme sur des groupes effectivement « traités » dans les termes du protocole, mais dont la comparabilité originelle n'est pas garantie, du fait des possibles biais de sélection liés aux « perdus de vue ». Par opposition, l'ITT inclut la totalité de l'échantillon initial dans l'analyse, quelles qu'aient été les modalités réelles de traitement.

Validité externe (External validity)

Échelle ou indice qui permet d'évaluer une recherche, et plus particulièrement ses conclusions externes ; la validité des résultats de la recherche pour un groupe qui n'a pas été étudié (le reste de la population). Le mot *externe* renvoie ici à la capacité de généraliser les résultats à un groupe qui n'a pas été étudié (la population) mais qui possède les caractéristiques du groupe qui a été étudié (l'échantillon).

Validité interne (Internal validity)

Échelle ou indice qui permet d'évaluer une recherche, et plus particulièrement ses conclusions internes. La validité interne repose sur les diverses caractéristiques d'une recherche qui permettent de vérifier l'hypothèse (est-elle vraie ou fausse?). Elle permet d'affirmer avec plus ou moins de certitude que X est bien la cause de Y et ainsi de répondre à la question : le variable indépendante (ou le traitement) a-t-elle eu l'effet escompté? Le mot *interne* renvoie ici à la portée des conclusions que l'on tirent de l'étude de l'échantillon (en non de la population).

CONTACTS DES BUREAUX J-PAL & IPA OFFICE



Projets J-PAL à travers le monde

Bureaux J-PAL

Bureaux Global au MIT

30 Wadsworth St., E53-320
Cambridge, MA 02142, ÉTATS-UNIS
Tél. : +1 (617) 324-6566
Email : info@povertyactionlab.org
Site internet : www.povertyactionlab.org

Bureau Europe à l'École d'économie de Paris

66bis avenue Jean Moulin
75014 Paris, FRANCE
Tél. : +33 (0)1 71 19 40 70
Email : jpaleurope@povertyactionlab.org
Site internet : www.povertyactionlab.org/europe

Bureau Asie du Sud à l'IFMR

Institute for Financial Management and Research
IITM Research Park, A1, 10th Floor
Kanagam Road
Taramani, Chennai 600113, INDE
Tél. : +91 44 3247 50 56
Email : jpalsa@povertyactionlab.org
Site internet : www.povertyactionlab.org/south-asia

Bureau Amérique latine et Caraïbes à Pontificia
Universidad Católica de Chile
Instituto de Economía
Av. Vicuna Mackenna 4860
Santiago, CHILI
Tél. : +(56-2) 354-1291
Email : jpallac@povertyactionlab.org
Site internet : www.povertyactionlab.org/LAC

Bureau Afrique à SALDRU
University of Cape Town
Private Bag X3
Rondebosch 7701, AFRIQUE DU SUD
Tél. : +27 21 650 5981
Email : jpalafrica@povertyactionlab.org
Site internet : www.povertyactionlab.org/africa

IPA

Bureau principal :
Innovation for Poverty Action
101 Whitney Ave
New Haven CT 06510, ÉTATS-UNIS
Tél. : +(1) 203 772 2216
Email: contact@poverty-action.org
Site internet : www.poverty-action.org

Pour obtenir les contacts des autres bureaux IPA :
Consultez la page : www.poverty-action.org/about/contact

