Les essais randomisé : impasse pour le développement de l'Afrique

Seán M. Muller

School of Economics
University of Johannesburg
Afrique du Sud

Les essais contrôlés randomisés (ECR) sont censés être un outil important de réduction de la pauvreté dans le monde, et de contribution au développement économique et social. Après un examen minutieux, le fondement de cette affirmation se révèle remarquablement faible. Pire, en contraste marqué avec le battage médiatique autour de cette approche méthodologique, il y a des raisons substantielles de croire que l'utilisation des ECR pourrait en fait *nuire* aux perspectives et au bien-être des Africains et des pays africains.

Ce court article développe ces préoccupations comme suit[1]. La première section discute de certaines limites méthodologiques fondamentales des ECR et de l'incohérence intellectuelle qui en résulte pour les partisans de cette approche. Elle montre à quel point le projet *randomista est en lui-même défectueux*. La deuxième section examine ensuite comment les points de vue et les préjugés préexistants, qu'ils soient idéologiques ou sur le fonctionnement des processus sociaux et économiques, contredisent le cadre des ECR en tant qu'entreprise scientifique neutre. La dernière section illustre brièvement ces arguments avec deux exemples importants en Afrique du Sud où, je pense, l'approche randomista a fait et continue de faire beaucoup de mal.

Aperçu des limites méthodologiques

Les ECR sont une méthode pour obtenir des estimations quantitatives des effets causaux, et leur utilisation dans les essais de médicaments dans des contextes médicaux est bien connue. Le déploiement d'ECR pour répondre aux questions sociales et économiques n'est pas simple et peut, dans certaines instances, même être contraire à l'éthique dans de nombreux cas (voir Hoffmann, ce numéro). Au-delà cependant, on pourrait affirmer qu'en tant qu'accent méthodologique différent dans l'enquête intellectuelle, les ECR sont « pour la plupart inoffensifs » (Angrist et Pischke 2009). Mais ce qui caractérise l'augmentation spectaculaire de l'utilisation de cette méthode en économie, ce sont des affirmations de supériorité méthodologique dans le domaine politique, et un effort délibéré pour obtenir de l'influence (Banerjee 2007; Banerjee et Duflo 2009; 2011). C'est ce dernier projet qui a récemment reçu le prix Nobel d'économie 2019 (Nobel Media 2019) et ses promoteurs comme d'autres (Ravallion 2009; 2018; Deaton 2010) que j'appelle des « randomistas ».

L'argument de base du projet *randomista* ressemble à ceci:

1. Nous avons besoin d'estimations quantitatives fiables des effets de causalité pour prendre les bonnes décisions politiques.
2. Les hypothèses requises par d'autres méthodes économétriques pour obtenir de telles estimations sont invraisemblables.
3. Une simple analyse à l'aide d'un ECR peut identifier les effets d'intérêt politique sans nécessiter de connaissances préalables.
4. Par conséquent, les ECR annoncent une « révolution de la crédibilité » (Angrist et Pischke 2010) en économie et devraient être priorisés par les décideurs politiques en recherche de réponses simples à des questions importantes.

Chaque élément de l'argument est contesté et a fait l'objet de critiques de fond, mais pour les besoins actuels, je me concentre sur une contradiction dévastatrice au cœur de la défense des ECR dans la politique de développement.

Même si les ECR identifient effectivement les effets de causalité, la pertinence politique directe nécessite de passer d'une constatation empirique dans une population expérimentale à la recommandation d'une intervention dans une population générale plus large qui est celle dont les décideurs politiques se préoccupent[2]. Et le principal obstacle à cela est que le l'effet causal d'une intervention existe rarement, voire jamais, de manière isolée: il dépend des caractéristiques des personnes et du contexte dans lequel l'intervention est mise en œuvre. Ou en d'autres termes: l'effet causal d'une intervention dépend de son *interaction* avec d'autres facteurs. Ces facteurs peuvent varier dans le temps ainsi que dans l'espace géographique, de sorte que même dans la population expérimentale, il y aurait un résultat différent cinq ans plus tard.

Bien qu'assez intuitif et souligné par les premiers auteurs à considérer systématiquement le problème par rapport aux expérimentations sociales (Cook et Campbell 1979), il présente un sérieux dilemme pour les randomistas. Si l'effet causal dépend d'autres facteurs, alors pour l'extrapoler à une population différente il faut connaître les facteurs pertinents et disposer de données pour voir comment ils diffèrent entre les deux populations. Mais l'argument méthodologique pour les ECR est que les hypothèses sur la structure causale, utilisées par des méthodes non expérimentales pour obtenir des estimations supposées causales, ne sont pas crédibles.

Cette argumentation peut être formalisée (Muller 2015) pour montrer qu'en affirmant simplement, plutôt qu'en établissant, que les résultats des ECR s'appliquent à des populations plus larges, les randomistas approuvent une hypothèse avec l'exacte même structure que celle qu'ils rejettent lorsqu'il s'agit de méthodes non expérimentales. Au mieux, cela est intellectuellement incohérent, au pire, c'est fondamentalement malhonnête ; dans les deux cas, c'est une faille dévastatrice au cœur du projet randomista (Muller 2020).

Bien que l'étendue du problème n'ait pas été suffisamment reconnue, dans le sens de s'abstenir de faire des recommandations politiques ou des promesses sur la pertinence des politiques, certaines solutions indirectes ont été proposées[3] L'une d'entre elles est particulièrement intéressante: les praticiens d'essais randomisés utilisent leur « expertise » pour évaluer quand/si les résultats expérimentaux peuvent être appliqués à d'autres populations. L'invocation de l'expertise qualitative, notamment dans une perspective plus large de scientisme imitatif, intéressera certainement de nombreux lecteurs qui connaissent le mépris historique manifesté par les économistes pour les méthodes qualitatives, et les revendications fondées sur l'expertise individuelle plutôt que sur les « données », les « modèles » et l'« économétrie ». Mais cela ne résout pas non plus la contradiction intellectuelle, car si l'expertise qualitative peut être utilisée pour décider des perspectives d'extrapolation, pourquoi ne pourrait-elle également pas être utilisée pour décider des perspectives d'identification des effets causaux? Cela rendrait redondants non seulement les ECR, mais plus largement, les méthodes économétriques.

Proposer l'expertise qualitative des randomistas comme solution au problème d'extrapolation attire également l'attention sur une autre caractéristique dangereuse de leur projet: non seulement les ECR sont placés en haut d'une hiérarchie de preuves, mais les randomistas cherchent à se placer en haut d'une hiérarchie des connaissances (« hiérarchie épistémique »).

Bien sûr, le fait de privilégier les opinions des économistes sur d'autres qui ont, sans doute, une plus grande expertise n'est pas un phénomène nouveau ; dans le passé, il a conduit à des accusations d'« impérialisme économique ». Cependant, la promesse de réponses simples basées sur des « expériences scientifiques » combinée à une poussée d'influence bien financée rend le projet randomista encore pire que ses prédécesseurs.

Expériences idéologiques

Être au sommet d›une hiérarchie d'expertise est un moyen sûr d'obtenir une influence politique, mais il n'y a aucune raison prima facie de croire qu'un universitaire américain qui dirige un ECR a plus de connaissances sur un système de santé local que, par exemple, un médecin compétent qui a travaillé dans ce système pendant des décennies. Un autre problème lié aux ECR a reçu peu d'attention, mais il est très important pour les pays en développement qui essaient d'élaborer des stratégies et des trajectoires appropriées pour le développement social et économique[4]. Le problème est le suivant: *le choix même d'une intervention sur laquelle fonder un ECR est lui-même le résultat d'une conception préexistante de la manière dont le monde fonctionne et de comment il devrait fonctionner.*

Prenons l'exemple suivant. Un décideur dans le district de Busia au Kenya est préoccupé par les résultats d'apprentissage dans les écoles publiques car une

proportion relativement faible d'enfants acquiert des compétences de base en lecture, écriture et calcul. Le décideur demande l'aide d'un randomista pour améliorer les résultats et le randomista, naturellement, propose qu'un ECR, ou plusieurs ECR, soient exécutés pour savoir « ce qui fonctionne » [5]. Mais d'où vient l'intervention qui formera la base de l'ECR?

Une réponse pourrait être de dire: «essayons quelque chose qui semble avoir fonctionné ailleurs». Mais cela pose, car dans le cadre totalement absurde de l'approche randomista, rien ne peut fonctionner sans avoir été vérifié par les résultats d'un ECR. Ainsi, dans le cas de base, le randomista doit tirer une intervention de l'ensemble d'interventions *qu'il pense quelles pourraient fonctionner*[6].

D'un point de vue purement méthodologique, cela est intéressant car, comme indiqué ci-dessus, le motif méthodologique du projet randomista est né du scepticisme pour les connaissances causales *ex ante.* Pourtant, le simple choix d'une intervention impose les convictions de chercheur dans au moins trois cas:

1. Dans la détermination de l'ensemble des interventions qui pourraient fonctionner en théorie.
2. De déterminer le sous-ensemble de 1 qui est considéré comme pratiquement réalisable.
3. De hiérarchiser les possibilités afin de sélectionner une seule intervention ou, au mieux, une poignée d'interventions.

Pour nous concentrer sur l'exemple que nous avons choisi, supposons que le décideur ou le chercheur ait les preuves d'un taux d'absentéisme élevé des enseignants, ce qui est considéré comme une cause probable de mauvais résultats. Quelle intervention expérimentale pourrait-on instituer? Le chercheur A qui considère que les fonctionnaires des pays en développement sont intrinsèquement paresseux peut préférer un système incitatif punitif basé sur une surveillance accrue. Si cela est pratiquement impossible en raison de la résistance des enseignants ou d'autres parties prenantes, un système basé sur les récompenses peut être la deuxième meilleure option. D'un autre côté, le chercheur B (qui estime que le manque de ressources et les environnements de travail de mauvaise qualité affectent négativement la motivation) peut proposer une intervention qui augmente considérablement les ressources scolaires.

Notez que l'intervention expérimentale préférée de chaque chercheur peut même ne pas faire partie de l'ensemble d'interventions possibles ou théoriquement efficaces de son homologue. Cela se répercutera également sur l'interprétation des résultats de tout ECR. Le chercheur A interprètera les preuves d'ECR de l'inefficacité d'une intervention basée sur les ressources comme confirmant simplement ce à quoi il s'attendait déjà, tandis que le chercheur B pourrait l'interpréter comme reflétant le fait que, pour avoir un impact, l'augmentation des ressources scolaires peut prendre du temps. De même, B interprétera l'inefficacité d'une intervention incitative comme reflétant le fait que l'absentéisme est causé par d'autres facteurs, tandis que A peut l'interpréter comme reflétant la nécessité de modifier la conception du mécanisme d'incitation.

Le point important est qu'il n'y a rien de neutre dans les ECR: les interventions choisies pour les tests sont le résultat de décisions prises par des chercheurs menant des expérimentations et des institutions qui les financent, et refléteront donc leurs notions préconçues sur le fonctionnement du monde et les solutions à considérer comme plausibles.

Deux exemples d'Afrique du Sud

Deux exemples spécifiques d'Afrique du Sud illustrent la pertinence de ces arguments et les dangers du projet randomista pour les pays en développement: l'utilisation trompeuse d'un ECR pour plaider en faveur d'une incitation fiscale à l'emploi; et la contribution des ECR et de leurs partisans à la négligence continue des contributeurs systémiques à de mauvais résultats scolaires. Les deux exemples sont discutés en détail dans des articles séparés.

Incitation fiscale à l'emploi : utilisation d'ECR pour fausser le débat politique

Au milieu des années 2000, le gouvernement sud-africain a invité un groupe d'économistes, connu par la suite sous le nom de « Groupe de Harvard », pour le conseiller sur la politique économique du pays (Center for International Development 2008). Une proposition émanant de cette initiative consistait en une incitation fiscale à l'emploi visant à réduire le taux extraordinairement élevé de chômage national (Levinsohn 2008). La proposition sous-tendait une conceptualisation du chômage comme résultant, au moins dans une large mesure, du coût trop élevé du travail. Ce point de vue a longtemps été contesté par les syndicats, conduisant à une situation polarisée avec des universitaires ayant des intérêts divergents: les uns voyaient les syndicats comme voulant privilégier leurs membres au

détriment des chômeurs, tandis que les autres voyaient les entreprises comme voulant détruire les actions collectives afin de mieux exploiter les travailleurs. Sans surprise, l'incitation fiscale proposée a été rejetée par les syndicats.

L'analyse initiale qui avait proposé l'incitation a reconnu que la question de savoir comment un emploi réactif est lié aux salaires est une question empirique et que, par conséquent, davantage de preuves étaient nécessaires pour étayer toute incitation et déterminer ses caractéristiques. Bien qu'il y ait déjà des preuves que le Trésor national et de nombreux universitaires impliqués pensaient qu'une incitation était souhaitable, deux études ont été menées pour éclairer la décision: l'une était une analyse d'équilibre général calculable (EGC) (Burns et al. 2010) et l'autre, un essai randomisé sur une subvention salariale remise aux demandeurs d'emploi.

La nature des études EGC est telle qu'elles *présument effectivement de la réponse à la question principale* (« une réduction des salaires financée par l'État augmenterait-elle l'emploi de manière causale? »), et modélise la sensibilité des résultats à d'autres hypothèses; en ce sens, ils sont plutôt inintéressants et inutiles pour la principale décision politique, et je ne discuterai pas davantage de ce travail ici.

L'essai randomisé a été mené par des universitaires ayant des liens avec le Trésor national et le soutien financier de 3ie, qui, avec *le Jamaal Abdul Latif Poverty Action Lab (J-PAL) et Poverty Action,* est l'une des principales organisations internationales finançant l'utilisation des ECR dans la recherche sur le développement. Le « plan d'influence sur les politiques » soumis au bailleur de fonds montre que les chercheurs anticipaient les syndicats comme un obstacle à l'impact de leurs conclusions, (car s'attendant clairement à un résultat positif) (Inconnu 2011).

Le document de travail contenant les détails de l'étude et ses conclusions n'a été publié qu'après l'approbation par le Parlement du projet de loi sur l'incitation fiscale à l'emploi (Levinsohn et al. 2014). Cependant, avant la décision, le chercheur local responsable de l'expérimentation a publié nombre d'articles dans la presse populaire plaidant pour l'adoption de l'incitation sur la base des résultats positifs de l'étude (Rankin 2012; 2013). Le Trésor national a également cité l'étude dans sa présentation au Parlement. Pourtant, certaines parties du document de travail complet qui a été publié plus tard sont plus prudentes quant à ce qui peut être revendiqué, et l'examen minutieux des détails de l'étude montre que l'ECR fournit peu, voire pas de perspectives sur la question politique centrale.

Parmi les raisons pour lesquelles l'affirmation selon laquelle les conclusions de l'ECR soutenaient la mise en œuvre de l'incitation nationale était fausse: l'intervention de la subvention qui ressemblait peu à l'incitation prévue; la population expérimentale qui n'était pas représentative au niveau national; des éléments de preuve supplémentaires qui n'appuyaient pas l'affirmation selon laquelle un salaire inférieur était le mécanisme à l'origine du taux d'emploi plus élevé des récipiendaires de bons; et que tout effet positif aurait pu être la conséquence d'un effet de concurrence entre les travailleurs qui disparaîtrait lorsque l'intervention serait intensifiée.

Cet exemple illustre les points soulevés dans les sections précédentes. Les allégations de principe fondées sur l'ECR n'étaient pas appropriées compte tenu des limites de l'étude. De plus, les chercheurs ont montré un parti pris clair en faveur de la politique. L'un d'eux a même travaillé pour une société de conseil qui fournissait des services à des courtiers en main-d'œuvre qui bénéficieraient directement de l'incitation. Pourtant, tirant parti du scientisme douteux et de l'autorité épistémique associés au projet randomista, un ECR a été utilisé pour approuver une politique qui engageait le gouvernement à des milliards de rands d'incitations fiscales pour le secteur privé à un moment où il mettait en œuvre l'assainissement de son budget.

ECR et déni sélectif des contributeurs systémiques aux mauvais résultats scolaires

De nombreuses études citées dans le prix Nobel 2019 concernent des expérimentations pédagogiques. La grande majorité d'entre elles se préoccupent d'interventions qui, soit n'accroissent pas sensiblement les ressources disponibles pour les écoles, soit, comme dans le cas de l'embauche d'enseignants à bas salaire, d'une manière qui mine le salaire ou le pouvoir des enseignants titulaires. Cela suit une longue tradition dans l'économie de l'éducation consistant à nier ou à minimiser la pertinence des ressources fiscales (sur la base, incidemment, de résultats économétriques bas, selon les normes randomistas).

Puisque que l'économie universitaire sud-africaine est en grande partie une entreprise imitative (Muller 2017), il n'est pas surprenant que ces deux positions aient été reproduites localement. Les chercheurs qui dominent actuellement cet espace politique

en Afrique du Sud ont produit un rapport sur les « contraintes dans l'éducation » qui ne mentionnait pas les ressources comme une contrainte (van der Berg et al. 2016). Dans un pays régulièrement classé comme le plus inégalitaire au monde, avec des taux de chômage, de pauvreté et de violence élevés, et un système éducatif pour les Sud-Africains noirs qui, jusqu'en 1994, était tristement célèbre pour produire des « porteurs d'eau ».

La justification de cette position remarquablement extrême est double. Premièrement, l'un des auteurs a précédemment affirmé que les dépenses d'éducation en Afrique du Sud étaient élevées, par rapport à d'autres pays (van der Berg 2007) et que les ressources ne pouvaient donc pas être la cause de mauvais résultats. Deuxièmement, dans les études réalisées à l'aide de méthodes économétriques non expérimentales, les auteurs et leurs collaborateurs n'ont apparemment pas trouvé de preuves que les ressources avaient un impact significatif sur les résultats. L'opinion selon laquelle les ressources sont sans importance concorde avec une opinion négative des enseignants, de la direction des écoles et des syndicats: ce ne sont pas les inégalités léguées par l'apartheid qui provoquent de mauvais résultats scolaires, mais simplement une gestion inefficace de ressources adéquates.

Les ECR s'inscrivent parfaitement dans cette position car, comme cela a été le cas au niveau international, ils concentrent leur attention sur les questions non structurelles, l'optimisation des ressources et les modèles de déficit des fonctionnaires des pays en développement. Il n'est donc pas surprenant que ces chercheurs et leurs collaborateurs partageant les mêmes idées au sein du Département de l'éducation de base aient, avec enthousiasme, adopté ces méthodes (même si l'acceptation de la prétention randomista sur la crédibilité devait rendre une grande partie de leurs travaux antérieurs non crédibles). Les exemples incluent des interventions mal pensées telles que l'envoi aléatoire de guides d'étude dans les écoles d'une province, puis aller à la pêche à des fins statistiques (Taylor et Watson 2015), ainsi que des projets à plus grande échelle un peu plus soigneusement étudiés pour tester la formation des enseignants, ou des interventions de lecture en début de scolarité.

Alors que le petit groupe de chercheurs menant ces études affirme que le problème d'extrapolation n'est « [pas] assez grave pour remettre en cause la méthode » (Fleisch et al. 2017:10), il est évident qu'ils ne comprennent pas le problème fondamental décrit ci-dessus. Il convient peut-être de mentionner qu'étant donné l'engouement actuel pour les ECR, l'adoption de cette méthode de recherche répond à la fois aux aspirations des chercheurs en matière de publication académique et au désir d'influence sur les politiques, qu'elles servent ou non l'intérêt public.

Ce qui a été particulièrement frappant dans le passage récent aux ECR de la politique sud-africaine d'éducation de base, c'est que les éducateurs et les militants de la société civile qui ont défendu les interventions en lecture pendant des décennies ont été ignorés par le gouvernement. Mais les économistes avec peu, voire aucune, connaissance directe du système éducatif qui préconisaient des interventions de lecture précoce dans le cadre d'expérimentations «scientifiques» sont rapidement arrivés à faire que leurs recommandations soient reproduites mot à mot dans le discours du chef de l'État à la nation[7].

Lié à cela, les chercheurs qui améliorent leur statut épistémique à l'aide d'ECR ont également plus de pouvoir pour éclairer les politiques en utilisant différentes méthodes pour répondre à des questions de politique distinctes sur lesquelles d'autres chercheurs peuvent avoir des connaissances et une expertise de plus longue date. Cela aggrave encore une hiérarchie épistémique sans doute déformée dans laquelle les universitaires et d'autres qui tirent leur expertise de la recherche récoltent presque tout le poids attribué aux contributions non politiques tandis que les «experts de la pratique», tels que les enseignants, sont ignorés, sauf lorsque leur expertise est filtré à travers la recherche, l'analyse et les anecdotes des universitaires. Et tout cela, comme dans le cadre de l'ECR d'incitation à l'impôt sur l'emploi qui a induit le public et les décideurs en erreur, s'inscrit dans un contexte plus large de « politique factuelle ».

L'impasse ECR

Pour ceux qui sont préoccupés par le développement, le principal problème du projet randomista n'est donc pas son intolérance méthodologique (Harrison 2013) en soi. C'est qu'un accent indu sur les ECR à des fins politiques n'est pas étayé sur le plan méthodologique, introduit clandestinement des préjugés idéologiques et épistémiques, détourne l'attention de questions importantes et, ce faisant, détourne des ressources intellectuelles et une volonté politique limitées vers des projets qui ne respecteront que rarement leurs promesses étriquées. Les randomistas semblent, sur leurs conclusions, être animés par un « zèle de missionnaire » (Bardhan 2013), avec une « incroyable certitude » qu'ils sont les élus qui sauveront les habitants des pays en développement de la pauvreté

(Manski 2011), ce qui est loin d'être justifié. Et la combinaison présente un réel danger pour les pays en développement qui ont des ressources limitées mais doivent empêcher un projet bien doté en ressources de déterminer leurs politiques.

De plus, comme l'ont noté de nombreux critiques, le projet randomista concentre à la fois la recherche et la politique sur des questions et des interventions limitées qui se prêtent aux ECR, plutôt que sur les plus importantes pour les pays en développement. La poursuite délibérée d'un changement structurel à moyen et long terme à travers un processus d'apprentissage qui a caractérisé les voies de développement de la plupart des nations désormais riches est en dehors du champ d'application du projet randomista (voir Chelwa, dans ce numéro). Ainsi, lorsque que les randomistas ont tenté de présenter leur position comme une position d'espoir plutôt que de pessimisme sur les perspectives d'améliorations majeures dans les pays en développement, cela s'est avéré faux. Le projet randomista se fonde, principalement, implicitement mais parfois explicitement, sur un pessimisme fondamental pour des pays en développement qui réalisent les améliorations économiques de leurs prédécesseurs.

Les connexions doivent être développées plus en détail, mais le projet randomista peut être considéré comme une manifestation extrême des tendances impérialistes des économistes, fondées sur des affirmations douteuses sur l'économie en tant qu'activité scientifique. À cet égard, il convient de noter que même ceux qui ont approuvé la proclamation d'une « révolution de la crédibilité » n'ont rien dit sur ce que cela implique pour tous les conseils politiques passés donnés par des économistes. S'il est effectivement vrai que des ECR sont nécessaires pour des estimations causales crédibles, alors toutes les déclarations de principe passées d'économistes utilisant les hypothèses invraisemblables d'autres méthodes doivent avoir été au mieux inappropriées, ou au pire nuisibles. Le zèle du projet randomista se manifeste dans l'argument selon lequel l'orgueil des économistes qui l'ont précédé était fondamentalement défectueux, et que « cette fois, c'est différent ». Un examen attentif du projet suggère le contraire: « cette fois, c'est pire ».

Pour toutes les raisons ci-dessus, j'ai une position aussi forte que celle de nombreux détracteurs des ECR : non seulement l'adoption généralisée et le recours à cette méthode ne produiront pas les avantages promis par le projet randomista, mais ils entraveront probablement la réalisation de projets à long terme, l'amélioration des perspectives et du bien-être des résidents des pays en développement. Bien placés méthodologiquement et épistémiquement, les ECR joueraient tout au plus un petit rôle d'éclairage des décisions politiques des pays en développement. Si on donne l'autorité et le pouvoir recherchés par les randomistes, les ECR seront une impasse pour le développement de l'Afrique. Quels que soient les facteurs qui ont entravé la réalisation de progrès plus importants dans les pays africains depuis l'indépendance, il n'y a aucune raison de croire que les ECR permettront de les résoudre ou de les contourner. Le défi pour les pays africains reste de définir, comme d'autres pays l'ont fait historiquement, une voie alternative au nouveau complexe du missionnaire qui s'est cristallisé autour du projet randomista.

Notes

1. Par souci de concision et d'exposition, je garde les références au minimum; des références plus détaillées peuvent être trouvées dans d'autres travaux sur lesquels le présent article est basé (Muller 2014b; 2014a; 2015; 2020) et des versions plus courtes de certains de ces arguments peuvent être trouvées dans Chelwa et Muller (2019) et Chelwa, Hoffmann et Muller (2019a; 2019b).
2. Ce problème est désigné comme le problème de la « validité externe », selon Cook et Campbell (1979) qui l'a opposé au problème de l'identification d'un effet causal (« validité interne »). Il est également appelé problème de « généralisabilité » ou de « transportabilité ».
3. Dans un chapitre de livre à paraître (Muller à paraître), je discute d'une série d'efforts pour résoudre le problème, dont la réplication et l'apprentissage automatique, et j'explique pourquoi ils sont inadéquats.
4. Il existe quelques critiques convaincantes des conceptualisations populaires de la notion de « développement », mais j'utilise le terme d'une manière large et relativement irréprochable ici pour faire référence à l'amélioration du bien-être et des perspectives des personnes à l'intérieur d'un pays, sans aucunement présumer de ce que pourrait signifier une amélioration.
5. Si les randomistas devaient proposer une autre approche, ils devraient probablement aussi suggérer que quelqu'un d'autre, qui serait en fait un expert de cette approche, soit consulté.
6. Dans certains endroits, les randomistas ont fait une grande partie de leur consultation avec les partenaires locaux pour décider de l'intervention à tester. Mis à part le fait qu'il existe peu de preuves indépendantes de cela, il est tout à fait clair, même dans de tels cas, que les chercheurs n'acceptent pas de mener des interventions qui, selon eux, sont susceptibles d'être inefficaces. De plus, la nature et la justification de la grande majorité des interventions sont évidemment de nature économique.
7. *Et ce n'est peut-être pas une coïncidence* si les groupes démographiques dominants dans le premier groupe étaient des femmes noires alors que dans le dernier groupe, il y avait majo-

ritairement des hommes blancs formés à l'université qui était le cœur intellectuel de l'apartheid.

Références

Angrist, J.D. et J.-S., Pischke, 2009, *Mostly Harmless Econometrics,* Princeton NJ: Princeton University Press.

Angrist, J.D. et J.-S., Pischke, 2010, '*The credibility* revolution in empirical economics: how better research design is taking the con out of econometrics', *Journal of Economic Perspectives* 24 (2): 3–30.

Banerjee, A. and Duflo, E., 2011, Poor *Economics: A Radical Rethinking of the Way to Fight Global Poverty,* Public Affairs.

Banerjee, A.V., 2007, *Making Aid Work,* Cambridge MA: MIT Press.

Banerjee, A.V. et Duflo, E., 2009, 'The experimental approach to development economics', *Annual Review of Economics 1: 151–78.*

Bardhan, P., 2013, 'Little, big: two ideas about fighting global poverty', Boston Review, 20 May.

Burns, J., Edwards, L. et Pauw, K., 2010, 'Revisiting wage subsidies: how pro-poor is a South African wage subsidy likely to be?'disponible à: http://www.opensaldru.uct.ac.za/handle/11090/73.

Center for International Develop-ment, 2008, 'CID South Africa Growth Initiative', available at: http://www.cid.harvard.edu/south africa/, consulté 13 septembre 2014.

Chelwa, G., Hoffmann, N. et Muller, S., 2019a, 'How randomised trials became big in development economics', The Conversation, 9 décembre disponible à: https://theconversation.com/how-randomised-trials-became-big-in-development-economics-128398.

Chelwa, G., Hoffmann, N. et Muller, S., 2019b, 'Randomised trials in economics: what the critics have to say', The Conversation, 10 décembre, disponible à: https://theconversation.com/randomised-trials-in-economics-what-the-critics-have-to-say-128538.

Chelwa, G. et Muller, S., 2019, 'The poverty of poor economics', Africa is a Country, 17 octobre, disponible à: https://africasacountry.com/2019/10/the-poverty-of-poor-economics.

Cook, T.D. et Campbell, D.T., 1979, *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Wadsworth.

Deaton, A., 2010, 'Instruments, randomization, and learning about development', *Journal of Econo-mic Literature* 48 (2): 424–55.

Fleisch, B., Taylor, S., Schoer, V. et Mabogoane, T., 2017, 'The value of large-scale randomised control trials in system-wide improvement: the case of the reading catch-up pro gramme', *South African Journal of Educa-tion* 37 (1): 1–13.

Harrison, G.W., 2013, 'Field experiments and methodological intolerance', *Journal of Economic Methodology* 20 (2): 103–17.

Inconnu, 2011, 'Policy influence plan – grant number OW2.044 – University of the Witwatersrand', disponible à http://policyimpact-toolkit.squarespace.com/storage/pip-examples/PIP OW2-044.pdf, consulté le 17 juin 2018.

Levinsohn, J., 2008, 'Two policies to alleviate unemployment in South Africa', Center for International Development Working Paper 166.

Levinsohn, J., Rankin, N., Roberts, G. et d Schoer, V., 2014, 'Wage subsidies and youth employment in South Africa: evidence from a randomised control trial', Stellenbosch Economic Working Papers 02/14, disponible à http://www.ekon.sun.ac.za/wpapers/ 2014/wp022014.

Manski, C., 2011, 'Policy analysis with incredible certitude', *Economic Journal* 121 (554): F261–F289.

Muller, S.M., 2014a, 'The external validity of treatment effects: an investigation of educational production', thèse de doctorat, University of Cape Town.

Muller, S.M., 2014b, 'Randomised trials for policy: a review of the external validity of treatment effects', working paper 127, SALDRU.

Muller, S.M., 2015, 'Causal interaction and external validity: obstacles to the policy relevance of randomized evaluations', *World Bank Economic Review* 29 (suppl. 1): S217–S225.

Muller, S.M., 2017, 'What does an (south) African economics look like?', in *Proceedings of the Scholarship of Teaching and Learning (SoTL) in the South Conference.*

Muller, S.M., 2020, 'The implications of a fundamental contradiction in advocating randomized trials for policy', *World Development* 127.

Muller, S.M., à paraître, 'Randomised trials in economics', in H. Kincaid et D. Ross eds. The Modern Guide to Philosophy of Economics, Edward Elgar.

Nobel Media, 2019, 'Advanced information', available at https://www.nobelprize.org/prizes/economic-sciences/2019/advanced-information/.

Rankin, N., 2012, 'A common approach', *Financial Mail,* 24 mai.

Rankin, N., 2013, 'Give youth wage subsidy a chance', *Mail and Guardian,* 1 novembre.

Ravallion, M., 2009, 'Should the Randomistas rule?', *Economists' Voice.*

Ravallion, M., 2018, 'Should the randomistas (continue to) rule?', Center for Global Development Working Paper 492.

Taylor, S. et Watson, P., 2015, 'The impact of study guides on "matric" performance: evidence from a randomised experiment', *Stellenbosch Economic Working Papers* 13 (15).

van der Berg, S., 2007, 'Apartheid's enduring legacy: inequalities in education', *Journal of African Economies* 16 (5): 849–80.

van der Berg, S., Spaull, N., Wills, G., Gustafsson, M. et Kotzé J., 2016, 'Identifying binding constraints in education', RESEP report for the Presidency, 24 mai.