

ランダム化フィールド実験による 教育プロジェクトの費用効果分析 —ケニアを事例に—

小川啓一
(神戸大学)

中室牧子
(コロンビア大学博士課程)

星野絵里
(コロンビア大学博士課程)

1. はじめに

近年、国際援助機関で、教育プロジェクトのインパクトを厳格に測定し、評価する必要があるとの見方が一層強まっている (World Bank 2008)。開発途上国で実施されたプロジェクトが期待された効果を上げることができたか、設定された目標以上の波及効果があったかなどの面で、客観的かつ正確な評価がなされて始めて、援助機関は当該プロジェクトのインパクトを的確に把握することができ、今後そのプロジェクトを拡張するか、縮小するかあるいは取りやめるかを決定することができる。限られた資源でインパクトのあるプロジェクトをつくるうえでもプロジェクト評価はとても重要である。

一方、ケニアでは、2003年に初等教育の無償化の開始により、過去数年において急速に粗就学率を上昇させたが、2006年の初等教育の純就学率は76%と、完全普及までの道のりは未だなお遠い (World Bank 2008)。その一方で教員や教材の供給が就学者数の急激な増加に追いつかず、教員のモチベーションを低下させ、生徒の読解力や計算力といった基礎学力の低さなどが顕在化し、教育の質低下の問題が深刻になっている (Glewwe, Ilias, and Kremer 2003)。

このような問題を解消するため、ケニア教育省は、正確な教育プロジェクト評価のための一手法として、ランダム化フィールド実験 (randomized experiments) に注目し、援助機関の支援を受けて、国内で積極的に同実験を実施している。

本研究の意義は、主に次の二つである。まず第一に、教育プロジェクトの評価手法としては未だあまり馴染みのないランダム化フィールド実験について紹介することである。第二に、これまで個別の政策・プロジェクトに関してランダム化フィールド実験による評価研究はそれなりにみられるようになってきているものの、一国の初等教育関連プロジェクトの評価結果をその費用や効果の面で横断的に比較した分析は未だ見られていないことから、費用対効果分析 (cost-effectiveness analysis) を用いて、実験結果を横断的に比較し、ある目標を達成するためにもっとも費用対効果の高い政策やプロジェクトを特定することを試みることである。後述するが、ランダム化フィールド実験は、セレクション・バイアスをコントロールすることができるという観点から、OLSなどの回帰分析と比較してそれぞれのプロジェクト評価に対する正確性を格段に高めている。それに加え、代替案との比較にもとづき、どのプロジェクトが最も

費用対効果が高いかを分析することは、限られたタイムラインで、限られた資源を効率よく配分するために非常に重要な政策的インプリケーションをもたらすであろうと期待される。従って、まず最初に、プロジェクトのインパクト評価の一手法として近年注目を集めているランダム化フィールド実験について概説し、ケニアの初等教育プロジェクトを対象に行われたランダム化フィールド実験による評価結果をレビューする。更に既存の研究によって得られている処置グループの平均処置効果の推計値を用いて、8つの教育プロジェクトの費用効果分析を行い、ケニアにおいては、費用対効果の観点から、教育の「量」を改善するインパクトに寄与するためには、受益者の機会費用を含めた教育費負担を軽減するような政策・プロジェクトが望ましく、教育の「質」の観点からは生徒や教員のインセンティブに働きかける政策・プロジェクトが有効であり、こうした傾向は他の開発途上国の実験データの費用効果分析からも導かれることを示す。最後に、ランダム化フィールド実験の問題点や課題について指摘する。

2. ランダム化フィールド実験とは何か

世界銀行によると、インパクト評価とは、あるプロジェクトや政策介入の「純粋」な因果効果 (causal impact) を計測するというところにほかならない (World Bank 2009)。例えば、政府が、就学率を上昇させるために、奨学金の給付を行うプロジェクトを立ち上げたと仮定しよう。単純に考えれば、このプロジェクトのインパクトを計測するには、プロジェクトに参加した子どもと参加しなかった子どもを比較すればよいということになる。しかし、この方法では、プロジェクトの効果を正確に計測す

ることはできない。なぜなら、奨学金を得て進学した子どもと、奨学金を得ずに進学しなかった子どもは、「根本的に」違っているので、奨学金を得た子どもと得なかった子どもを比較しても、奨学金を得た子どもが仮に、奨学金を得られなかった環境にあった場合にどうなっていたかという反事実 (counterfactual) との比較にはならないからである。具体的には、奨学金の存在を知っていて、学校の開始時期以前に奨学金に申し込みをする親は、所得が低くても、教育に関心や意欲が高い可能性があり、そうした家庭出身の子どもは、潜在的な学力は、その他の子どもと比較すると高い可能性がある。これを、肯定的選択 (positive selection) という。

当然、逆のケースも考えられる。奨学金を得たいと考える親は、自らの収入を教育費に充てることはないが、奨学金が得られるならば学校へ行かせてもよいという、教育に対する支出優先順位の低い親である可能性もあり、これを否定的選択 (negative selection) という。理論的には、こうした自己選択が肯定的か否定的かは、事前には判断できない¹。奨学金プロジェクトに参加するか否かという自己選択が、親の教育熱心さや子どもの潜在的な学力にかかる観察不能な要因によって影響を受けている場合、モデルの自己選択変数と誤差項が相関するという内生小生問題が生じる。このように、あるプロジェクトに参加した特定の個人が、自己選択 (self-select) していることによって正確なプロジェクトのインパクト評価が困難になる問題を、セレクション・バイアス (selection bias) といい、データを用いた実証研究における最も重要な課題となっている (Glazerman, Levy, and Meyers 2003)。

この問題を解決するのに、最も有効な手法の一つがランダム化フィールド実験である (Kremer 2003)²。先の奨学金の例で

は、まずプロジェクトの前に、対象となる子供の属性を綿密に調査し、その後、奨学金を供与される処置グループ (treatment group) と供与されない対照グループ (control group) を (抽選などにより) ランダムに振り分け、自己選択の可能性を排除する。そして、プロジェクト後に計測された二つのグループの平均的な教育指標の差が、プロジェクトの純粋な効果であると考えられる。これを平均処置効果 (Average Treatment Effect: ATE) と呼ぶ。奨学金を得た子供が仮に、奨学金を得られなかった環境にあった場合にどうなっていたかという仮定の反事実を、属性が似通っている対照グループを観察することによって得るのである。

ランダム化フィールド実験は、元来、医薬品の効果実験などでよく使われる手法であったが、近年、教育プロジェクトのインパクト評価の試金石ともいわれている (Angrist and Pischke 2008)。セレクション・バイアスは特に、教育にかかる問題において顕著であることから、同バイアスを取り除いた上で、プロジェクトの効果を厳密に測定することが可能であるランダム化フィールド実験の結果には、高い信頼性が寄せられている。こうした認識が広がって以降は、国際援助機関、特に世界銀行 (以下、世銀) がランダム化フィールド実験に積極的な支援を行っており、数年以内に、更に多くの実験結果にアクセスすることが可能になると考えられる (Kremer 2003)。

3. ケニアにおけるランダム化フィールド実験

ここでは、ケニアで実施された8つの初等教育プロジェクトにおけるランダム化フィールド実験の結果をレビューし、それぞれのプロジェクトが教育の質・量をどのように改善したか (あるいは改善しなかつ

たか) について分析する (実験の詳細は、図表1を参照)。

(1) 奨学金の給付

先にセレクション・バイアスの事例として取り上げた奨学金の例は、実際に2001年から2003年にかけてケニアで行われた教育プロジェクトの一つである。ケニアでは、近年、初等教育における習熟度向上のための教育資材の直接的投入の効果などを計測する一方で、例えば奨学金の給付を通じ生徒自身の勉学への意欲的参加を促し、学習への動機付けをおこなうことの効果が注目されている。奨学金の給付プロジェクトにおけるランダム化フィールド実験では、奨学金が授与された処置小学校 (treatment school) の生徒の平均テストスコアを対照小学校 (control school) と比較し、処理グループの平均処置効果 (Average Treatment Effect on the Treated: ATT) を推計している。結果、全体では、処置校の平均テストスコアは対照校より0.12標準偏差高く、特に女子は0.22標準偏差高いことが明らかになった。こうしたプラスの効果は奨学金のための選抜が終了した後も持続し、翌年以降の成績の向上にも寄与したことが明らかになっている。また、プシア学区においては、奨学金給付の対象とならない男子生徒の間でも成績の向上が見られたり、生徒や教師の欠席率を下げるなどの波及効果も認められた (Kremer, Miguel, and Thornton 2008)。

(2) 虫下し薬の配布

ケニアでは、不衛生な学校保健管理状況により、生徒の寄生虫病感染率が非常に高く、このための体調不良が高い欠席率の温床となっている (Miguel and Kremer 2004)。ケニアのプシア学区で行われた虫下し薬配布プロジェクトでは、実験の結果、処置校の子供は対照校の子供と比較して

図表 1：ケニアにおけるランダム化フェイルド実験の概要と費用効果分析（単位はすべて、米ドル）

プロジェクト名 融資先、期間	ターゲット 結果	場所・期間 サンプル数	手法・概要	処置グループの 平均処置効果	
				(教育の量)	(教育の質)
3-1. 奨学金の給付 世銀、その他財団 2001-2003年	テストスコア (Δ) 出席率 (Δ)	ケニア西部州、ブ シアとテソ学区、 127の小学校(処置： 64、対照：63)	処置校において、学力テストの上位15%の女子 生徒に奨学金が給付される。奨学金が授与され た学生に対しては2年間の授業料が直接学校に 支払われると同時に、制服・教科書代などの教 育費用（約12.8ドル）も、奨学生の両親へ給付 される。	90.21	3.53
3-2. 虫下し薬の配布 世銀 児童発育パート ナーシップ等 1998-2002年	テストスコア (Δ) 出席率 (ー)	ケニア西部州、ブ シア学区、75の小 学校（処置：25、 対照：25）	処置校において、虫下し薬を配布する。75校を 3グループに分け、1年ごとに処置校を変えて、 効果を計測する。最初の処置校郡（グループ1） を対象にした虫下しの配布が1998年に行われ、 その後99年に、次の処置校郡（グループ2）に、 2000年に最後の処置校郡（グループ3）に配布 された。	3.50	効果なし
3-3. 視覚教材の提供 世銀、国立科学財団 1997-1998年	テストスコア (▼) 出席率 (ー)	ケニア西部州、ブ シアとテソ学区、 178の小学校(処置： 89、対照：89)	1997年度中、処置校の第6学年から第8学年に 在籍する生徒に、科学や数学の授業において視 覚教材を用いる。	効果なし	効果なし
3-4. 教師のインセン ティブ導入 世銀、その他財団 1997年	テストスコア (Δ) 生徒の出席率 (ー) 教師の欠勤率 (ー)	ケニア西部州、ブ シアとテソ学区、 100の小学校(処置： 50、対照：50)	処置校では、第4学年から第8学年のクラス を教える教師を対象に、生徒の成績（学校ご との平均）に応じて最高得点（top-scoring school）、及び最も改善を示したものの（most- improved school）にボーナスが与えられる。	効果なし	3.70

<p>3-5. 経常的な教育費用の無償化 世銀、1994年</p>	<p>就学率 (Δ) 退学率 (Δ) 就学期間 (Δ)</p>	<p>ケニア西部州、ブシアおよびテソ学区、14小学校(処置: 7, 対照: 7)</p>	<p>14の小学校のうち、7の処置校には、教科書や制服代としておおよそ7.94ドルが支給される。</p>	<p>12.51</p>	<p>効果なし</p>
<p>3-6. 教科書の支給 世銀、1995-2000年</p>	<p>テストスコア (Δ/ー) 留年率 (ー) 退学率 (ー)</p>	<p>ケニア西部州、ブシアおよびテソ学区、100小学校</p>	<p>処置校において教科書を無料支給する。100校を4つのグループに分類し、実験開始年(1996年)には、第1グループ(25校)を処置校群とし、残り(75校)を対照校とした。1997年には、第1・2グループを処置校群(50校)、第3・4グループを対照校群(50校)とし、1998年には第1・2・3グループを処置校群(75校)、第4グループを対照校群とした。</p>	<p>効果なし (5.61)</p>	<p>(5.61)</p>
<p>3-7. 学校給食の提供 財団等、2000-2002年</p>	<p>テストスコア (Δ) 出席率 (Δ)</p>	<p>ケニア西部州、ブシアおよびテソ学区、50幼稚園(処置: 25, 対照: 25)</p>	<p>50の幼稚園のうち、25の処置園では、出席者全員に、無料の朝食が提供される。</p>	<p>36.00</p>	<p>効果なし</p>
<p>3-8. ピアエフエクト 世銀、オランダ政府 2005-2007年</p>	<p>テストスコア (Δ)</p>	<p>ケニア西部州、ブシゴマ、ブテレ、ムミアス学区、121小学校(処置: 60, 対照: 61)</p>	<p>処置校では、成績順にクラスわけ (tracking) を行い、対照校ではランダムにクラスわけを行った。</p>	<p>N./A. N./A.</p>	<p>N./A. N./A.</p>

(注) テストスコアに対する教科書支給の効果は、1997年に行われた児童支援プログラム(3-5の経常的な教育費用の無償化プロジェクト)の一貫として行われた教科書の無料配布によるもの。教科書の無料配布プロジェクトでは、テストスコアに対する効果は観察されなかった。
(出所) 各実験結果を元に、筆者ら計算。

25%程度虫感染率の低いこと、身長が0.5cm高いことが認められるなど著しく健康状態が改善した。この結果、出席日数が第1～第4学年で15日、第5学年以上では、10日間の増加が観察された。更に、プロジェクトの対象となった小学校の周辺地域でも感染率が低下し、出席日数も増加（平均4～5日）するなど、地域レベルでの波及効果も認められた。ただし、処置校と対照校の間で、テストスコアに対する差は観察されなかった (Miguel and Kremer 2004)。

(3) 視覚教材の提供

非実験系データを用いた過去の研究では、教科書や視覚教材など、教育資材の投入は、習熟度の向上に寄与することが明らかになっている (Shepard 1967 等)。1997年からケニア西部州で行われたランダム化フィールド実験では、科学や数学などの科目に、説明用の図表が書かれたフリップチャートを用いることが、学習向上につながるかに焦点をあて、処置校と対照校の平均テストスコアを2年後に比較した。しかし、この比較では、両校の平均テストスコアには殆ど差がない (4.8%の標準誤差で、処置校のほうが、0.06標準偏差低い) という結果が得られた。また両校の間で、出席率についても有意な差が観察されなかった (Glewwe, Kremer, Moulin and Zitzewitz 2004)。

(4) 教師のインセンティブ導入

多くの開発途上国の教育現場同様、ケニアでも、教師の欠勤率は非常に高く、教師のモラルの欠如が問題となっている。教員インセンティブプロジェクトが実施されたケニアのブシア・テソ学区では、教師の欠勤率は20%にも及んでおり、教師をどのように動機付けするかは目下、喫緊の政策課題であった。こういった状況下、同プロジェクトは、4から8学年を指導する教員を対

象に、毎年実施する学区試験の結果をもとに、一定のボーナスを提供することで、教員のモチベーションを向上させることを試みた。実験の結果、2年後には、処置校の生徒のテストスコアは、対照校の生徒と比較して、0.14標準偏差高いことがわかった。しかし、これは、最大の目標であった出勤率（および生徒の退学率や留年率）の改善にはつながらず、また教授法や宿題提出率などの質的な改善を引き出すこともできなかった。その代わりに、教師は、より直接的に給与の上昇に繋がるテストには注力し、通常授業の枠外でも集中的なテスト対策を行うなどしたため、短期的にテストスコアが上昇するという結果となったものの、その1年後にはテストスコアへの影響も消滅した (Glewwe, Ilias, and Kremer 2003)。

(5) 経常的な教育費用の無償化

ケニアの教育行政は、学校建設の費用は地方が負担し、教員の配置と給与の支払いを中央政府が担当する仕組みになっているが、この地方と中央の役割分担が誤った資源配分の温床になっているとの指摘がある。即ち、一旦学校を建設すれば、中央政府から教員とその給与が継続的に手当てされるため、地方政府は、地域の雇用確保のためもあって小規模校を乱立し、中央政府の財政を圧迫しているという現実がある³。その結果、中央政府は教員給与以外の教育支出に手が回らず、経常的な教育費用、例えば教科書などは受益者負担となり家計を圧迫し、子供の就学を妨げる要因となっている。ケニア西部州で行われたランダム化フィールド実験は、仮に中央と地方の分担の非効率化が是正され、経常的な教育費用が中央政府から補填された場合を想定し、家計の教育費用の負担（教科書、制服などおよそ7.94ドル）がない処置校と、受益者負担となる対照校の間で就学期間、就学率、退学率を比較した。5年後には、被験校の生徒

の就学期間が0.5年、テストスコアを下げるこなしに退学率は大幅に低くなることがわかった (Kremer, Moulin and Namuny 2003)。

(6) 教科書の支給

ケニアにおける教科書の供給不足は深刻で、特に地方では、複数の生徒が教科書を共有することは日常的な光景となっている。開発途上国を対象にした先行研究の多くが、教科書支給が、テストスコアに、プラスの影響を与えることを明らかにしていることもあって、教科書の無償支給プロジェクトが実施された。結果は、プロジェクトの対象となった処置校と対照校の生徒の平均テストスコア、留年率、退学率の間には、有意な差がみられないことが明らかになった。しかし、平均では差がみられなかったものの、プログラム開始前に受けた試験の結果が、上位40%の生徒のテストスコアは有意に上昇しているほか、彼らの中等学校への進学率も上昇するという結果も明らかになった。これは、ケニアの教科書が、大方の生徒にとっての第三言語である英語で書かれている上、英語で指導を行う教師の不在率が高いことから、もともと学習能力の低い生徒にとっては読みこなすことが難しく、学習成果の向上には繋がらないことを意味しているものと考えられる (Glewwe, Kremer, and Moulin 2007)。

(7) 給食の支給

学校出席率の上昇を促すための政策介入として、インド、バングラディシュ、パキスタンなどで、政府の補助金で賄われる学校給食が提供されている (World Food Program 2002)。これらのプロジェクトは、学校給食の提供が、両親が子供を学校に就学させる動機に繋がるとの考え方に基づいている。ケニアにおける低学年児童の出席率の低さを解消するために、幼稚園におけ

る給食無償化が、児童の出席率を高めるかどうかを評価するため、ランダム化フィールド実験が行われた。結果は、処置校の子供の出席率は、対照校の子供と比較すると30%も高いことが明らかになった。但し、テストスコアは、NGOによって同時期に行われた教師のトレーニングを受けた新人教師がいる学校のみで、処置校の子供は、対照校の子供よりも0.4標準偏差高くなっており、全体としての効果は観察されなかった。また、給食を提供し始めた処置校は出席率の増加に伴い、授業料を引き上げ、それに対して対照校が授業料を引き下げるといった行動反応が見られた (Vermeersch and Kremer 2004)。

(8) ピアエフェクト

近年ケニアでは、MDGの達成にむけて急速に就学率が上昇しているのに対して、教員の供給が追いつかず、教育の質低下が懸念されている。ケニア西州部で行われたランダム化フィールド実験では、まず、契約教員を採用する予算が与えられ、1学年に複数のクラスを作り、クラス分けを行う。処置校ではクラスわけはテストの成績に基づいて行われ、契約教員と元々の教員がどちらを教えるかはランダムに決定される。対照校では、テストの成績に関係なくランダムにクラスわけがおこなわれ、どちらの教員が教えるかもまたランダムに決定される。この実験は、成績にもとづくクラスわけ (tracking) により、同じクラスの生徒の学習レベルを均質化することに伴う指導のしやすさが教育の質改善に繋がるかどうかを評価するために行われたものである。18ヵ月後のテスト結果をみると、処置校は、対照校よりも0.14標準偏差高くなっている。更にその1年後には処置校の平均テストスコアは、対照校よりもさらに0.16標準偏差高くなっていることが明らかになった (Duflo, Dupas, Kremer 2008)。

(9) 非実験系データを用いた分析との比較

非実験系データを用いた従来の回帰分析では、セレクション・バイアスにより正確なインパクト評価が困難であることは既に述べた。これを確認するため、ランダム化フィールド実験の処置校のデータだけを用いた、最小二乗法 (OLS) による遡及評価 (retrospective evaluation)⁴の結果をみてみよう。例えば、視覚教材提供プロジェクトの遡及評価は、フリップチャートの使用は、テストスコアを0.2標準偏差上昇させることを示している (Glewwe et al. 2004)。しかし既知のように、ランダム化フィールド実験は処置校と対照校のテストスコアには有意な差はないことを明らかにしている。教科書支給プロジェクトの遡及評価 (Glewwe, et al. 2007) もまた、教科書配布は、3科目平均で0.12標準偏差、テストスコアを上昇させることを示している。しかし、これについても、ランダム化フィールド実験は、教科書配布によるテストスコアの上昇は観察されないことが明らかになっており、非実験系のデータを用いた従来のインパクト評価には上方バイアスがあることが窺われる⁵。また、非実験系データを用いた分析は、俗に言う出版バイアス (publication bias) によって、多分に政策担当者をミスリードしている可能性がある点にも注意が必要である。この点、仮にそのプロジェクトに効果がなかった場合もその研究価値を失わないランダム化フィールド実験は、適切な教育資源の配分を行うための政策判断に不可欠であると考えられる。

4. 費用効果分析

ここまでのランダム化フィールド実験が既存の評価手法と比較して正確にプロジェクトのインパクトを評価できる点について強調してきたが、こうした結果を単独でも、どのプロジェクトを優先すべきかと

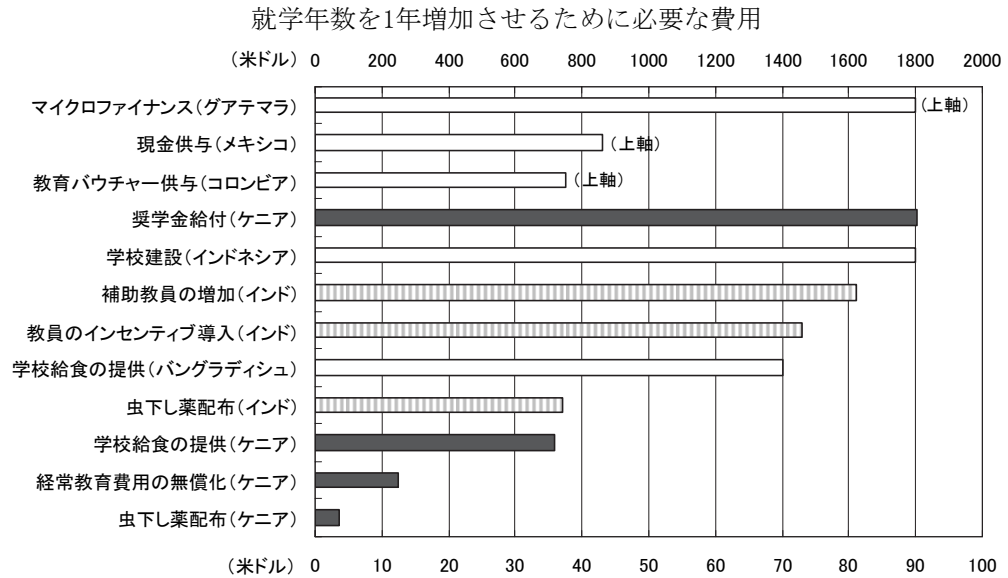
いう観点から有効な政策的インプリケーションは得られない。そこで、就学率あるいはテストスコアの向上という目標に対して、複数の手段がある場合、ある目標を達成するためには、どれが最も費用対効果の高いプロジェクトであったかを費用効果分析を用いて把握することを試みる。政策担当者にとっては、なるべく費用が低く、効果があげられるプロジェクトに投資することが望ましいことはいまでもなく、費用効果面から、プロジェクトのインパクトを比較することは国際援助機関の援助の効率性の観点からも重要である。ここでは、ケニアの教育にかかる状況を鑑みて特に重要だと考えられる教育指標、教育の量という面から就学率、教育の質という面からテストスコアという2点に対する費用効果分析を行う。ランダム化フィールド実験の推計パラメータを用いて、費用効果分析を行うことの利点は二つある。まず第一に、共通の政策目標 (ここでは、就学率とテストスコアの改善) があれば、政策ツールが何であろうと (例えば、制服、虫下し、視覚教材など) 横断的に比較することができる点である。次に、セレクション・バイアスをコントロールできることによって、正確な評価が可能になる点である。

これまでみてきた8つのプロジェクトの結果を、それぞれ、(1) 就学率の場合は、1年就学率を増加させるために必要な費用に換算し、(2) テストスコアの場合は、標準偏差を0.1増加させるために必要な費用に換算する。これまでに紹介してきたランダム化フィールド実験の既存研究において、処置グループと比較グループの平均的な就学率の差や、平均的なテストスコアの差が明らかになっていることから、それぞれの政策やプロジェクトにかかった費用をこれらの推計パラメータで割ることによって、費用対効果比が求められる。一例を挙げると、奨学金給付プロジェクトにかかった費

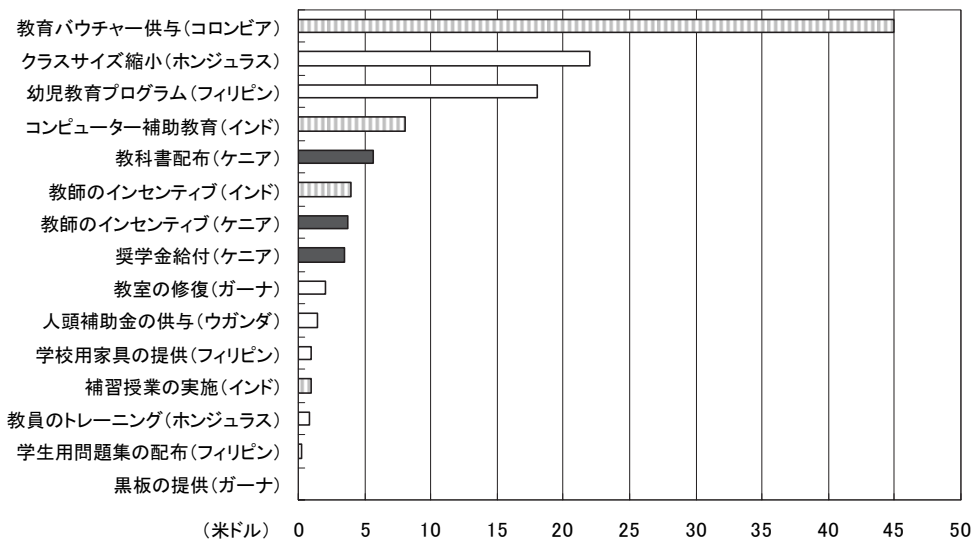
用は、生徒1人当たり4.24ドルであり、ランダム化フィールド実験によって推計された処置グループの平均処置効果(就学率ベ

ス)は0.047であることから、費用対効果は $4.24/0.047$ で90.21ドルと計算される。また、同プロジェクトにおける処置グルー

図表2：実験データに基づく教育プロジェクトの費用効果分析の国際比較



テストスコアを0.1標準偏差増加させるために必要な費用



(注) 黒はケニアで行われたランダム化フィールド実験、灰色縞線はケニア以外の開発途上国で行われたランダム化フィールド実験、白は擬似実験から得られたデータに基づく費用効果分析の結果を示す。

(出所) 各実験結果をもとに、筆者ら計算。

プの平均処置効果（テストスコアベース）は0.12であることから、この費用対効果は、 $4.24/0.12 \times 0.1$ で3.53ドルと計算される。

以上のように、各プロジェクトの1年あたりの費用対効果比を計算し、比較したものが図表1である。図表から明らかなように、奨学金給付と視覚教材提供プロジェクトの例外を除けば、就学率を改善するプロジェクトは、テストスコアの上昇には効果を発揮しない（あるいは逆もまた然り）ことが示唆されており、それぞれの教育指標改善のためには、それぞれ異なる政策が必要であることが窺われる。就学率の改善のためには、虫下し薬の配布がもっとも費用対効果が高く（3.50ドル）、次いで経常的な教育費用の無償化（12.51ドル）や学校給食の提供（36.00ドル）が高い。一方、テストスコアの上昇には、奨学金の給付や、教師のインセンティブ導入（3.70ドル）の費用対効果が高いことがわかる。総じて見れば、費用対効果の観点から、ケニアにおいては、親が子どもを就学させることの意味決定が、教育にかかる費用に対して弾力的であることが窺われ、機会費用を含む受益者負担を引き下げようとする政策・プロジェクトが有効である可能性が高い。一方、生徒の学習達成度の向上のためには、費用の引き下げよりもむしろ、生徒や教員に対するモチベーションを高める政策・プロジェクトが有効であることが示唆されている。

次に、ケニア以外の国で行われた実験データのうち、就学率とテストスコアの改善を目標にしたプロジェクトとケニアの8プロジェクトを比較し、就学率とテストスコアを改善するために費用対効果の高い初等教育プロジェクトが、国際的にみて共通の傾向があるかどうかについて検証する。より多くの事例との国際比較を可能にするために、ランダム化フィールド実験だけでなく擬似実験⁶の事例も扱うこととし、費用は全て実験当時の為替レートとPPPを用いて、

米ドルに換算した。図表2をみると、就学率の改善については、ケニア同様、虫下し配布（インド）や給食提供（バングラディッシュ）など、教育の機会費用を引き下げるプロジェクトの費用対効果が高くなっている。テストスコアの改善については、ケニアでみられたようなインセンティブに働きかけるプロジェクト以外にも、学生用問題集の配布（フィリピン）、教員のトレーニング（ホンジュラス）、補習授業の実施（インド）など、学習改善に直接的な効果が見込めるプロジェクトの費用対効果が高いことがわかる。ただし、ガーナにおいては黒板の提供や教室の修復など教育資本に対する投資のリターンが高いことが窺われる。就学率の結果と比較すると、テストスコアの結果のほうが、国ごとのばらつきが大きく、教育の質改善に関するプロジェクトは、社会経済的あるいは文化的環境に配慮して、その国独自のものが効果的であると考えるのが自然であろう。

既述のように、費用効果分析は、同じ目標を持つプロジェクトであれば、異なる介入ツールのインパクトを費用と効果の面から比較可能であるが、例えば、就学率の上昇か、テストスコアの改善か、というように異なる目標を持つプロジェクトを比較することはできない。例えば、ケニア教育省が、農村部での就学率上昇か、都市部でのテストスコアの改善のどちらを優先すべきかを判断すべき場面では、これまで用いてきたような費用効果分析のフレームワークは有用ではない点には留意が必要である。

5. ランダム化フィールド実験の問題点

これまでは、新しい教育プロジェクトの評価ツールとして、ランダム化フィールド実験のメリットに焦点を当ててきたが、当然のことながらランダム化フィールド実験

は、万能ではない。澤田 (2008) は、ランダム化フィールド実験は、費用がかかりすぎるため、学校建設の大型インフラストラクチャーの効果測定にはなじまない手法であると述べている。更に、Kremer (2003) は、実験データを評価するにあたり、(1) サンプルセレクション (2) 欠損値バイアス (attrition bias)、(3) 行動反応による波及効果、の3点について、実験データや実験の行われた環境を精査するよう促している。サンプルセレクションは、仮にランダム化フィールド実験によって、処置グループと対照グループがランダムに振り分けられたとしても、実験開始後に転入・転出を試みるなどして実験の恩恵に与ろうとする個人の行動によってバイアスが生じる可能性があることである。欠損値バイアスは、例えば給食や制服の支給を目的に、学校に通学することを決めた新規入学者が、これまでの学習蓄積のなさから処置グループの平均点を引き下げるなどする可能性があることである。波及効果は、上記の給食支給プロジェクトでも触れたように、処置グループに給食が支給されると、対照グループのほうで、新規入学者の獲得のため、学費を引き下げるなどの競争原理が働くことがあり、こうした副次的な効果が、正確なインパクト評価の障害になっている可能性についてである。こうした技術的な問題に加え、ランダム化フィールド実験の実施そのものを困難にする理由として、(子供を実験台にするという) 倫理上の問題や、教育現場に統計に基づく判断を持ち込むことへの嫌悪感から敬遠されることが指摘されている。従って、ランダム化フィールド実験は、実験参加者や開発途上国の政府関係者と綿密に合意形成を図りつつ実施されることが望ましく、更に、評価を行う際には潜在的なバイアスによる影響を最小限にとどめるため、実験データを念入りに点検することが肝要であるといえよう。

6. まとめ

本研究では、近年、開発援助機関で注目を集めているインパクト評価の一手法であるランダム化フィールド実験が、ケニアの教育プロジェクトをどのように評価しているか、8つの事例を考察した。この事例から得られた処置グループの平均処置効果を用いて、8つのプロジェクトの費用対効果を比較したところ、就学率の改善には、受益者の機会費用を含めた教育費負担を軽減するような政策・プロジェクトが有効であり、テストスコアの改善には教師や生徒への動機付けを可能にするような政策・プロジェクトが有効であることがわかった。この傾向は、他の国のプロジェクトの費用効果比と比較しても同様のことが言えることが明らかになった。ただし、ランダム化フィールド実験は、万能ではなく、評価においては、実験データを精査し、可能な限りバイアスを取り除くことが必要である。

注

- ¹ 技術的には、回帰分析によって奨学金が就学率に与えたインパクトを推計する際に、奨学金を得たかどうかという (通常は、奨学金を得れば1、得なければ0で表されるダミー変数である) 説明変数と、誤差項に含まれる自己選択にかかると観察不可能な要因が相関することを通じて、最小二乗法による推定量が一致性を持たなくなるという問題をひきおこす。
- ² セレクション・バイアスをコントロールしつつ、プロジェクトや政策介入のインパクトを評価する推計手法は、無作為抽出実験だけではない。例えば、回帰不連続モデル (Regression Discontinuity Design) や、操作変数法 (Instrumental Variable Methods)、傾向スコアマッチング (Propensity Score Matching)、差分の差分法 (Difference in Difference)、固定効果モデル (Fixed Effects Model) など

がある。

- ³ 例えば、ケニアの一人当たり GDP が 340 ドル程度に対し、教員給与は、年間約 2,000 ドルと破格に高い。このため、クラスサイズを縮小し、教員支出を抑えれば、1 年間の一人当たり教科書代 3.44 ドル、2 年間の一人当たり制服代 5.63 ドル、教室の改築・補修の平均費用 130 ドルを賄うことが可能であると考えられる。
- ⁴ 遡及評価とは後ろ向き評価ともいい、因果関係を検討するため、過去のデータが、現時点の結果を予測できているかどうかを確認するための手法。
- ⁵ 非実験系データを用いたインパクト評価が下振れではなく、上振れする理由について、澤田 (2008) は、プロジェクトが、行政上の容易さや受け入れ側のキャパシティを手がかりに内部要因で決定されているためであると分析している。
- ⁶ 擬似実験とは、既存のデータから、処置グループと対照グループの括りだしを行い、比較するという、いわば、無作為抽出実験に近い状態を、研究者が統計上再現した分析手法。

参考文献

- 澤田 康幸 (2008) 「インフラは貧困削減に寄与するか？インドにおける大型ダムの効果」 [http://www.rieti.go.jp/]，独立行政法人経済産業研究所 RIETI。
- Angrist, J.D. and Pischke, J.S. (2008). *Mostly Harmless Econometrics: An Empiricist's Comparison*. New Jersey: Princeton University Press.
- Duflo, E., Dupas, P. and Kremer, M. (2008). "Peer Effects and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *NBER Working Paper*, 14475.
- Glazerman, S., Levy, D.M. and Myers, D. (2003). "Nonexperimental Versus Experimental Estimates of Earnings Impacts." *The ANNALS of the American Academy of Political and Social Science*, 589(1), p.63-93.
- Glewwe, P., Kremer, M. and Moulin, S. (1997). *Textbooks and Test scores: Evidence from a Prospective Evaluation in Kenya*. Department of Economics, Harvard University, Cambridge, MA.
- Glewwe, P., Kremer, M. and Moulin, S. (2007). "Many Children Left Behind? Textbook and Test Scores in Kenya." *Poverty Action Lab Paper No. 44*.
- Glewwe, P., Ilias, N. and Kremer, M. (2003). "Teacher Incentives." *NBER Working Paper*, p.9671.
- Glewwe, P., Kremer, M., Moulin, S. and Zitzewitz, E. (2004). "Retrospective v. Prospective Analysis of School Inputs: The Case of Flip Charts in Kenya." *NBER Working Paper*, p.8018.
- Kremer, M., Miguel, E. and Thornton, R. (2004). "Incentive to Learn." *NBER Working Paper*, p.10971.
- Kremer, M. (2003). "Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons." *American Economic Review: Papers and Proceedings*, 93 (2), p.102-106.
- Kremer, M., Moulin, S. and Namunyu, R. (2003). *Decentralization: A Cautionary Tale*, Department of Economics, Harvard University, Cambridge, MA.
- Lockheed, M.E. and Hanushek, E. (1988). "Improving Educational Efficiency in Developing Countries: What Do We Know?" *Compare* 18(1), p.21-38.
- Miguel, E. and Kremer, M. (2004). "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72(1), p.59-121.
- Pritchett, L. (2002). "It Pays to Be Ignorant: A Simple Political Economy of Rigorous Program Evaluation." *Journal of Economic Policy Reform* 5(4), p.251-269.
- Shepard, R. (1967). "Recognition Memory for Words, Sentences, and Pictures." *Journal of Verbal Learning and Verbal Behavior* 6, p.156-163.
- Vermeersch, C. and Kremer, M. (2004). "School Meals, Educational Achievement and School Competition: Evidence from a Randomized

- Evaluation.” *Policy Research Working Paper* 3523, World Bank. [worldbank.org](http://www.worldbank.org)
- World Bank. (2008). *World Development Indicators* [<http://www.worldbank.org>]

Cost-Effectiveness Analysis of Education Project Evaluation by Randomized Experiments: The Case of Kenya

Keiichi Ogawa
Kobe University

Makiko Nakamuro
Columbia University

Eri Hoshino
Columbia University

Are there strategies to measure the unbiased impact of a particular project, program or policy in such a way as to prioritize public action and maximize the outcome? The decision to expand, reduce or maintain the size of a particular project is based on impact evaluation. How an evaluation can eliminate potential bias and assess the impact of intervention in an accurate fashion is extremely important for the implementation of project, program and policy by development communities and government officials.

Randomized experiments are generally thought to be the "gold standard" in policy research. Experimental designs that involve randomly assigning participants to one or more treatment conditions and a no-treatment control group for comparison allow researchers to more confidently answer research questions about whether or not a certain policy or intervention has an influence on observed changes in the population of interest. Selection, or the possibility that an individual's unobserved characteristics are actually explaining the observed association between the intervention and the outcome, is largely ruled out in a randomized experiment.

This paper introduces recent randomized experiments of eight educational projects in the primary education sub-sector in Kenya to demonstrate how this experimental design works to reduce potential bias. Further, using the average treatment effect on the treated (ATT) estimated by the experiments, this paper runs a cost-effectiveness analysis in terms of (i) quantitative measures expressed as gains in school attendance and (ii) qualitative measures expressed as gains in test scores.

The analysis reveals that education projects diminishing the burden on parents of educational costs are more cost-effective than alternatives to improve quantitative measures at primary education in Kenya. On the other hand, those providing appropriate incentives for both students and teachers are more cost-effective to improving qualitative measures. We confirm the same result from the cost-effectiveness analyses using data obtained by randomized and quasi-experiments, in other developing countries.