**무작위의 마법**

**수습부원 김한나**

**[[1]](#footnote-1)**

1. **빈곤에 대하여**

무엇이 빈곤을 낳는가? 무엇이 특정 사람들에게는 부와 풍요를, 또 다른 사람들에게는 가난과 고난을 쥐여주는가? 빈곤의 원인은 오랜 시간 경제학에서 궁극적인 질문으로 다뤄져 왔다. 학자들이 지역적, 문화적, 종교적 특색으로 설명하려고 노력해왔던 이 질문은 어떠한 명확한 해결책 없이 결국 더욱 구체적으로, 빈곤의 해결에 대한 요구와 욕구로 심화되었다.

그러나 세계의 빈곤은 아직도 해결되지 않았다. 세계에서 13억 명이 하루 1.25 달러 이하로 연명하는 극빈층이며, 30억 명 이상이 하루 2.5 달러 이하로 살아가고 있다. 빈곤으로 고통받는 인구가 전반적으로 줄었음에도 불구하고, 세계 빈곤층의 삼 분의 일 이상을 차지하는 사하라 이남 아프리카의 빈곤층은 오히려 증가해왔다.[[2]](#footnote-2)

빈곤에 대한 세계의 경각심이 줄어들었다고 할 수는 없다. 해외 원조와 각국 정부의 지원은 끊임없이 이어져 왔다. 2018 GHA(Global Humanitarian Assistance) 보고서에 따르면 2017년 한 해에만 정부와 민간에 의한 원조가 273억 달러에 달했다.[[3]](#footnote-3) 그렇다면 재정적 지원을 통해 가난한 사람들을 ‘빈곤의 덫’[[4]](#footnote-4)으로부터 구할 수는 없는 것인지 의문이 든다. 나아가, 몇몇 학자들은 대규모 해외 원조가 오히려 상황을 악화시킨다고 말한다. 가난한 국가를 의존적으로 바꾸며, 발전을 위해 자발적으로 노력할 유인과 성장의 기회를 빼앗는 것에 불과하다는 것이다.

그렇다면 결국 최종적인 질문은 우리가 빈곤의 해결을 위해 지갑을 열 용의가 있다는 가정하에, 우리의 지원이 세계 빈곤 퇴치에 긍정적인 영향을 주는지 여부이다. 나아가, 우리의 지원이 어느 부분을 겨냥해야 하는지 또한 고민해봐야 한다. 오랜 기간 정부와 NGO에서 돈을 쏟아부었음에도 불구하고 해결하지 못한 문제는 결국 돈의 절대적 규모보다는 그 방향에 대한 숙고를 요구한다. 같은 돈으로 교사 한 명을 더 고용하는 것이 아프리카의 한 가족에게 돈을 쥐여주는 것보다 빈곤 퇴치에 더욱 효과적인지 생각해봐야 한다.

그러나 문제는 우리는 언제나 오직 한가지 행동에 관한 결과를 볼 수밖에 없다는 것이다. 동일한 인구 구성, 동일한 문화적 배경, 동일한 환경에 놓인 두 국가나 가족을 비교하지 않는 이상, 아프리카의 한 가족에게 돈을 주는 것이 더 나은 결과를 낳을지, 상황을 악화시킬지는 알 수 없다. 단순하게 말해, 가정은 가정에 불과할 뿐, 시도하지 않은 것의 결과는 알 길이 없다는 것이다.

가정에 대한 결과는 눈으로 확인할 수 없다는 문제에 대한 해답으로 2019년 노벨 경제학상 수상자 바네르지(Banerjee), 뒤플로(Duflo)와 크레이머(Kremer)는 ‘무작위 통제 실험(Randomized Controlled Trial, RCT)’을 제안했다. 그들은 세계 빈곤 퇴치 연구에 RCT를 적용하여 개발경제학의 발전을 이끌었다는 공로를 인정받아 상을 받았다.

1. **무작위 통제 기법(Randomized Controlled Trial)**

위 세 명이 주요 실험 기법으로 이용한 RCT는 이전부터 임상 연구에 적용되어 왔다. 예를 들면, 신약의 효과를 검증하기 위해, 실험군에는 신약을 제공하고 대조군에는 위약을 제공하여 그 결과를 비교한다. 두 결과값의 차이가 신약의 효과를 그대로 반영한다는 것이 RCT의 핵심이다. 이러한 RCT는 70년대에 들어 미국에서 정책을 평가하기 위한 수단으로 활용되었다. 90년대를 넘어서며 개발 경제학에서 급격하게 인기를 얻었으며, 2000년대에 들어 더욱 그 적용 범위가 넓어져 이러한 추세가 가속화되었다.

RCT의 기본 아이디어는 간단하다. 앞서 예로 든 신약으로 설명하도록 하겠다. 한 제약회사에서 특정 병에 대한 신약을 개발하고 여러 안전성 검사를 진행한 후에 실제 환자를 대상으로 그 효과를 알아보려고 한다면, 어떠한 방식으로 실험군과 대조군을 정해야 할까? 실험군과 대조군의 설정에서부터 오류나 편향에 의해 그 효과에 대한 실험 결과가 오염될 수 있기 때문에, 이러한 오류를 여과하는 것은 매우 중요하다. 대조군과 실험군을 무작위로 설정하고 표본 크기를 키움으로써 표본 추출 상의 오류를 줄이고 편향이 없는 결과를 도출할 수 있는 방식이 바로 무작위 비교 연구(RCT)이다.

모든 환자 i에 대해, 신약을 받은 환자의 경우 Di=1, 위약을 투여 받은 환자의 경우 Di=0으로 표시하고, 신약을 복용한 환자의 결과를 Y1i, 위약을 복용한 환자의 결과를 Y0i로 표시한다고 하자. 그렇다면 신약의 효과는 같은 환자에 대한 신약 투여의 결과와 위약 투여의 결과의 차로 구해질 수 있다.

Y1i - Y0i

그러나 이 값은 한 환자 i에 대해서 그 사람이 신약을 투여 받은 경우와 위약을 투여 받은 경우의 차이이다. 이러한 이상적인 값은 현실에서는 불가능하다. 동일한 환자에 대해 동일한 기간 동안 상반되는 처리를 할 수는 없기 때문이다.

이때, 신약을 투여 받은 환자 i에 대한 결과의 차의 기댓값은 조건부 확률을 이용하여 아래와 같이 나타낼 수 있다.

E(Y1i - Y0i) ---(1)

* E(Y1i - Y0i | Di=1) ---(2)
* E(Y1i | Di=1) - E(Y0i | Di=1) ---(3)

(3)의 첫 항 E(Y1i | Di=1)는 실험을 통해 구할 수 있지만, 두 번째 항은 구할 수 없다. 또한 일반적으로 E(Y0i | Di=1)는 E(Y0i | Di=0)과 다르기 때문에, 이를 실제 존재하는 값으로 대체할 수도 없다. 그렇다면 신약을 받은 사람 i가 반대로 약을 투여 받지 않았다면 어떠한 결과로 이어졌을지 어떻게 예측할 것이며, 어떻게 실험 결과에서 신약의 효과만을 뽑아낼 수 있을까?

|  |  |  |
| --- | --- | --- |
|  | Di=0  i가 위약을 투여 받은 경우 | Di=1  i가 신약을 투여 받은 경우 |
| Y0i  위약을 투여 받은 i의 결과값 | E(Y0i | Di=0)  위약을 투여 받은 i에게 위약을 투여했을 경우의 결과값 – 실험으로 도출 가능 | E(Y0i | Di=1)  신약을 투여 받은 i에게 위약을 투여했을 경우의 결과값 – 실험으로 도출 불가능 |
| Y1i  신약을 투여 받은 i의 결과값 | E(Y1i | Di=1)  위약을 투여 받은 i에게 신약을 투여했을 경우의 결과값 – 실험으로 도출 불가능 | E(Y1i | Di=1)  신약을 투여 받은 i에게 신약을 투여했을 경우의 결과값 – 실험으로 도출 가능 |

일단, 실험으로부터 구할 수 있는 값은 아래와 같다.

E(Y1i | Di=1) - E(Y0i | Di=0)

여기에 E(Y0i | Di=1)를 빼고 더하면 아래의 식으로 정리된다.

E(Y1i | Di=1) - E(Y0i | Di=1) + E(Y0i | Di=1) - E(Y0i | Di=0)

이때 처음 두 항 E(Y1i | Di=1) - E(Y0i | Di=1)이 우리가 목표로 하는 값(Average Treatment Effect)이며, 마지막 두 항 E(Y0i | Di=1) - E(Y0i | Di=0)은 처리를 받은 실험군 A와 받지 않은 대조군 B가 모두 처리를 받지 않은 경우의 결과값을 나타내는 것이다. 이는 선택 편향(Selection Bias)[[5]](#footnote-5)을 의미하는데, 실제 실험에서는 신약을 받은 A가 신약을 받지 않았다고 가정할 때의 결과값의 차이, 즉 두 집단 간 기존에 존재했을 차이를 나타낼 수 있는 것이다.

이때 RCT를 통해서 Di가 Yi와는 독립적인 변수로 설정될 수 있다. 처리에 노출되는 실험군과 처리를 받지 않는 대조군을 무작위로 나누는 것이다. 그 결과로, 마지막 두 항 E(Y0i | Di=1)와 E(Y0i | Di=0)은 모두 Di와 무관해지고, 두 항이 서로 같아지며 선택 편향이 0이 되어 사라지게 된다.

RCT의 핵심은 이곳에 있다. 무작위 배정을 통해서 대조군과 실험군 설정의 선택 편향을 여과할 수 있다는 것이다. 모든 변수나 특성들을 평균적으로 같아지게 하고, 오로지 각 집단에 대한 개입만 달리하면서, 그 순수한 효과를 알아낼 수 있게 된다는 것이다.

1. **작은 디테일의 차이**

선택 편향을 없앰으로써 처리의 실제 영향을 도출할 수 있는 RCT를 개발 경제학에 적용함과 동시에, 바네르지, 뒤플로와 크레이머는 거시적 변화가 아닌 작은 변화가 빈곤 퇴치에 어떠한 영향을 주는 알아보는, 이전의 경제학과는 다른 접근법을 시도했다. 앞서 논의했던 원조와 시장 간의 논쟁이 아닌 새로운 시각에서 빈곤 퇴치에 대한 해결법을 제안한 것이다.

바네르지와 뒤플로가 인도에서 진행한 연구[[6]](#footnote-6)에서는 비금전적인 유인이 1세에서 3세 사이의 아동들의 예방접종률에 어떠한 영향을 주는지 RCT를 통해 살펴보았다. 총 1640명의 아동을 대상으로 진행된 이 실험에서는 134개의 마을을 무작위로 세 가지 집단에 배치[[7]](#footnote-7)했다. 첫 번째 실험군에서는 별도의 유인 없이 접종 캠프를 한달에 한번 설치했고, 두 번째 실험군에서는 같은 방식으로 캠프를 한 달에 한번 설치하되 예방접종을 받는 경우 렌틸콩 1kg[[8]](#footnote-8)을 주고, 필수 예방접종 5가지를 모두 맞는 경우 쇠 접시를 주었다. 마지막으로 대조군에서는 별도의 개입이 없었다.

그 결과로, 대조군에서는 6%가, 첫 번째 실험군(캠프 설치, 유인 없음)에서는 18%가, 두 번째 실험군(캠프 설치, 유인 있음)에서는 39%가 접종을 완료했다. RCT의 개념을 활용하면 이러한 결과는 두 가지 방향으로 해석될 수 있다. 첫째로, 대조군과 첫 번째 실험군의 차이는 캠프 설치 유무이다. 이때 마을들은 무작위로 배정되었으므로 이 두 결과값의 차이인 12%는 예방접종 캠프 설치가 예방접종률에 미치는 영향을 그대로 반영한다고 할 수 있다. 둘째로, 두 실험군 간의 차이는 보상 유무이다. 이전의 논리와 같이, 두 결과값의 차이인 21%는 비금전적인 보상으로 예방접종율이 증가한 정도를 나타낸다. 비록 금전적인 보상이 아니고, 그 가치가 크지 않아도 작은 경제적 보상을 통해서도 가난한 사람들에게 행동의 유인을 부여할 수 있음을 보이는 것이다. 특히, 이러한 유인이 가난한 지역에서 예방접종률을 높일 수 있음을 증명했다는 점에서 이 연구는 가치가 있다.

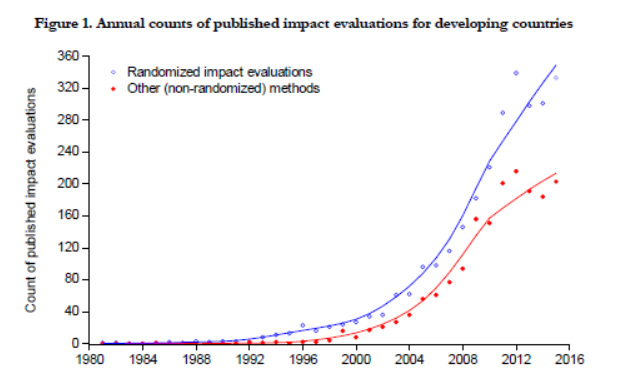
바네르지와 크레머가 케냐에서 진행한 연구[[9]](#footnote-9)는 사하라 이남 아프리카의 여성 청소년들에게 가장 큰 건강상의 위험을 야기하는 성 접촉에 의한 감염(Sexually Transmitted Infections, STIs)과 청소년 임신을 주제로 이뤄졌다. 교육이 이러한 현상에 미치는 영향을 연구하기 위해 첫 번째 실험군에서는 교육 보조금을 제공했다. 앞선 연구와 같이, 금전적인 보상 대신에 초등학교 교복 두벌을 무상으로 제공하였다. 두 번째 실험군에서는 초등학교 교사들이 후천적 면역 결핍증(Acquired Immune Deficiency Syndrome, AIDS) 관련 교육 과정을 이수했다. 추가로, 세 번째 실험군에서는 이 두 가지를 모두 적용하였으며, 이전 연구와 마찬가지로 대조군을 두었다.

지도, 텍스트이(가) 표시된 사진

자동 생성된 설명

실험의 결과는 이전 연구와 같이 각 실험군과 대조군 사이의 차이로 알 수 있다. 결과를 정리하자면, 주황 점선으로 표현된 교육 보조금의 경우, 대조군과 비교했을 때 청소년 임신을 줄였으나 단순 헤르페스 바이러스 2(Herpes Simplex Virus Type 2, HSV2)에는 큰 효과가 없었다. 빨간 점선으로 표현된 HIV 교육의 경우, 대조군과 비교했을 때 큰 효과가 없었다. 반면, 이 두 가지 개입이 모두 이뤄졌을 때의 결과는 독특했다. 초록색 실선을 보면, 대조군과 비교했을 때 청소년 임신에 미치는 영향이 단순 교육 지원금에 의한 영향보다 작았다. 한편 HSV2 유병률에는 더욱 큰 효과를 보였다.

앞서 제기된 빈곤이 여성 청소년에게 야기하는 현상들에 대해 교육의 개선이 중요하게 고려되어왔다. 이러한 일반화된 사고에서 출발한 연구는 구체적으로 어떤 방식을 통한 개입이 가장 효과적인지 보여주었다.

1. **무작위의 환상?**

**[[10]](#footnote-10)**

영향 평가에 있어 무작위 추출 기법은 2000년대 이후 꾸준히 상승세를 보여왔다. 이러한 급속한 증가와 더불어 RCT에 대한 비판 또한 이어져 왔다.

우선, 다른 실험 기법으로도 선택 편향을 여과할 수 있다. 예를 들어, 도구 변수(Instrument Variable)의 활용을 통해서도 이러한 편향을 없앨 수 있다.[[11]](#footnote-11) 회귀 불연속 설계(Regression Discontinuity)를 통해서도 편향을 없앨 수 있는데, 이는 실험군에 배정될 확률이 다른 변수의 불연속 함수인 경우에 가능하다. 예를 들어, 학생들이 상위 10% 이상의 점수를 받으면 교수님에게 칭찬을 받고, 그 이하의 점수를 받으면 교수님에게 아무런 평가를 받지 못한다고 가정하자. 이때 독립변수로 점수를, 종속변수를 시험 이후의 학업적 성취로 설정하여 회귀식을 찾고, 상위 9%에 위치한 학생과 상위 11%에 위치한 학생의 시험 이후의 학업적 성취를 비교하면 교수님의 칭찬과 학생의 학업적 성취 사이의 상관관계를 알 수 있다. 이 두 학생 집단은 기존의 학업적 성취에 큰 차이가 없었기 때문에, 유일한 차이인 교수님의 칭찬이 이후의 학업적 성취에 영향을 주는 정도를 확인할 수 있게 된다. 이렇듯 특정 임계값을 중심으로 가까운 샘플에 대해서는 선택 편향이 일시적으로 사라질 수 있다. 이외에도 Differences in Differences 기법을 통해서도 선택 편향을 여과할 수 있다.

RCT는 비용적 측면에서 다른 기법에 비해 열등하다. 선택 편향을 없앨 수 있다는 것은 비단 RCT만의 효력이 아닌 데다가, OLS 기법이나 Differences in Differences 기법처럼 기존의 데이터를 이용해서 결과를 도출하는 방식이 아닌 새롭게 실험을 진행해야 하므로 그 비용이 꽤 크므로, 굳이 RCT를 활용할 이유가 없다는 시각이 있다.

추가로 RCT의 외부 타당도에 대한 논의도 끊이지 않고 있다. 다른 대부분의 실험이 그렇듯, RCT 또한 특정 환경에서만 그 효력과 의미가 있다고 비판받는다. 앞서 말했듯 개발 경제학에서 RCT는 궁극적으로 새로운 정책을 제안하고, 그 효과를 확인하기 위한 목적을 가진다. 이때 RCT가 성립하기 위해서 실험에서 가정하는 내용이 현실과 같다고 단정할 수는 없기 때문에, 이 결과로 새로운 정책을 옹호해서는 안 된다는 주장이다.

또한 무작위로 배정된 두 집단의 성향이 온전히 같다는 보장이 없기 때문에, 두 집단의 결과값의 차인 ATE(Average Treatment Effect)가 그대로 처리의 영향을 반영하고 있다고 말하기 어렵다는 비판이 존재한다.

추가로, 전반적인 RCT 기법에 대한 비판 중 가장 기본적인 것은 윤리적 문제이다. 처리를 받는 집단을 무작위로 선택하는 과정에서 이러한 처리를 진정으로 필요로 하는 사람에게 가지 않을 수도 있다는 것이다. 물론 개발 경제학에서 적용되는 RCT는 의료 환경과 달리 사람의 목숨과 덜 직접적으로 연관되어 있다는 점에서 그 비판에서 조금이나마 자유로울 수 있다.

1. **나가는 글**

RCT는 완벽하게 통제된 상태에서 진행되는 연구실 안의 실험이 아니다. 오히려 우리가 모든 변수를 통제하지 못할 때 혹은 모든 변수에 대해 파악할 수 없을 때, RCT는 요술봉처럼 이러한 장애물을 사라지게 해준다. 무작위로 배정하면서 연구자는 본인이 알지 못하는 변수에 대한 것까지, 이를 찾으려 노력하지 않고도 저절로 사라지게 만드는 것이다.

물론 앞서 밝혔듯 RCT에는 한계가 있다. 그러나 이러한 단점들로 개발 경제학에서 RCT를 배척해서는 안 된다. RCT는 실험 기반의 연구를 가능케 한다. 단순히 거시적 측면에서 ‘원조’의 효과를 관찰하고, ‘시장 메커니즘과 시장 발전’을 주장해온 과거의 연구와는 구분된다. 우리는 인적 자본이 성장에서 중요한 요인임을 배웠고, GDP와 여러 변수의 상관관계를 그린 그래프들을 공부해왔다. 그러나 실질적으로 무엇이 인적 자본을 성장시키는지, 어떤 것이 이러한 환경을 개선하는지 알아보는 노력은 미미했다. 그런 의미에서 RCT를 활용한 연구는 기존의 탁상공론에서 벗어난다는 뛰어난 장점을 지닌다. RCT는 책상에서 벗어난 현실적인 접근을 제공한다. 이를 통해 경제학은 학문을 위한 학문이 아닌, 문제 해결에 대한 현실적인 실마리를 제공한다.

**참고 문헌**

Banerjee et al., 「Improving Immunization Coverage in Rural India: Clustered Randomized Controlled Evaluation of Immunization Campaigns With and Without Incentives」, BMJ 2010;340:c2220.

Duflo et al., 「Education, HIV, and Early Fertility: Experimental Evidence from Kenya」, American Economic Review 2015, 105(9): 2757–2797.

Ravallion, Martin, 「Should the Randomistas Continue to Rule」, CGD Working Paper 492, 2018 p. 4.

‘Extreme Poverty is Getting Rarer-But It Is Also Getting Harder To Root Out’, The Economist, 20180922 <https://www.economist.com/finance-and-economics/2018/09/22/extreme-poverty-is-growing-rarer>

‘What’s it like to be really poor?’ <https://www.theworldcounts.com/stories/Poverty-in-the-World-Today>

‘Global humanitarian assistance report 2018’ <http://devinit.org/post/global-humanitarian-assistance-report-2018/>

1. ‘Extreme Poverty is Getting Rarer-But It Is Also Getting Harder To Root Out’, The Economist, 20180922 <https://www.economist.com/finance-and-economics/2018/09/22/extreme-poverty-is-growing-rarer> [↑](#footnote-ref-1)
2. ‘What’s it like to be really poor?’ <https://www.theworldcounts.com/stories/Poverty-in-the-World-Today> [↑](#footnote-ref-2)
3. ‘Global humanitarian assistance report 2018’ <http://devinit.org/post/global-humanitarian-assistance-report-2018/> [↑](#footnote-ref-3)
4. 컬럼비아대 교수 제프리 삭스(Jeffrey Sachs)는 대규모 원조를 통해서 가난한 나라를 ‘빈곤의 덫(Poverty Trap)’에서 꺼낼 수 있다고 주장했다. 이러한 그의 주장은 빅 푸쉬(Big Push) 이론과 연관 지어져 빈곤의 해결에 있어 해외로부터의 원조와 투자의 중요성을 뒷받침하고 있다. [↑](#footnote-ref-4)
5. 선택 편향(Selection Bias)은 표본 편향(Sampling Bias)라고도 불리며, 일반적으로 통계 결과의 객관성을 저해하는 오류를 일컫는다. 특히, 표본을 선택하는 과정에서 일어나는 표본 선정의 오류로, 결론에 유의미한 영향을 준다. 결과 도출에 있어 왜곡을 일으킬 수 있다. 예를 들어서, 한국에서 영어로 된 설문을 진행하는 경우 교육 수준에 따라 응답률이 다를 것이므로 그 결과에도 영향을 주게 된다. [↑](#footnote-ref-5)
6. Banerjee et al., 「Improving Immunization Coverage in Rural India: Clustered Randomized Controlled Evaluation of Immunization Campaigns With and Without Incentives」, BMJ 2010;340:c2220. [↑](#footnote-ref-6)
7. 층화추출법으로 지역을 중심으로 층별로 집단을 배분한 뒤 Stata를 통해서 무작위 추출이 이루어졌다. [↑](#footnote-ref-7)
8. 렌틸콩 1kg의 가치는 실험 지역의 평균 하루 임금 3/4에 상응한다. [↑](#footnote-ref-8)
9. Duflo et al., 「Education, HIV, and Early Fertility: Experimental Evidence from Kenya」, American Economic Review 2015, 105(9): 2757–2797. [↑](#footnote-ref-9)
10. Ravallion, Martin, 「Should the Randomistas Continue to Rule」, CGD Working Paper 492, 2018 p. 4. [↑](#footnote-ref-10)
11. 처리 변수 X와 상관관계에 있고 결과 변수 Y와는 상관관계에 놓여있지 않은 추가적인 도구 변수 Z를 선택하여, Z를 추가했을 때 X, Y의 변화를 통해 X가 Y에 미치는 진정한 효과를 알 수 있다. [↑](#footnote-ref-11)