Using Randomised Control Trials to Evaluate Policies

Bansi Malde

EDePo, Institute for Fiscal Studies

December 2013

Outline

- Why Evaluate Policies?

Outline

- Why Evaluate Policies?
- The Evaluation Problem

Outline

- Why Evaluate Policies?
- The Evaluation Problem
- Randomised Control Trials (RCTs)
    - How do these work?
    - Limitations of RCTs

Outline

- Why Evaluate Policies?
- The Evaluation Problem
- Randomised Control Trials (RCTs)
  - How do these work?
  - Limitations of RCTs
- Application:
  - RCTs in Development: Providing Information on Child Nutrition in Rural Malawi

Why Evaluate Policies? (Policy View)

- Policy interventions are very expensive $\rightarrow$ UK govt spending ~45% of GDP

Why Evaluate Policies? (Policy View)

- Policy interventions are very expensive $\rightarrow$ UK govt spending ~45% of GDP
- In the context of development, a lot of debate on the effectiveness of development aid in fostering development
    - Development aid amounts to roughly $128bn a year

Why Evaluate Policies? (Policy View)

- Policy interventions are very expensive $\rightarrow$ UK govt spending ~45% of GDP
- In the context of development, a lot of debate on the effectiveness of development aid in fostering development
    - Development aid amounts to roughly $128bn a year
- Need to know whether money is well spent

Why Evaluate Policies? (Policy View)

- Policy interventions are very expensive $\rightarrow$ UK govt spending ~45% of GDP
- In the context of development, a lot of debate on the effectiveness of development aid in fostering development
  - Development aid amounts to roughly $128bn a year
- Need to know whether money is well spent
- There are many alternative policies possible

Why Evaluate Policies? (Policy View)

- The effects of policies are rarely obvious
    - Unless you have strong beliefs (ideology), policies can have wide-ranging effects

Why Evaluate Policies? (Policy View)

- The effects of policies are rarely obvious
  - Unless you have strong beliefs (ideology), policies can have wide-ranging effects
  - Economic theory is useful but leaves many policy conclusions indeterminate (depends on parameters)

Why Evaluate Policies? (Policy View)

- The effects of policies are rarely obvious
  - Unless you have strong beliefs (ideology), policies can have wide-ranging effects
  - Economic theory is useful but leaves many policy conclusions indeterminate (depends on parameters)
  - Correlation is not causation

Why Evaluate Policies? (Academic View)

- Evaluation is a crucial part of applied economics

Why Evaluate Policies? (Academic View)

- Evaluation is a crucial part of applied economics
- Field experiments offer useful exogenous variation to
    - Estimate parameters
    - Test models and theories

The Evaluation Problems

- Uncovering causal relationships

The Evaluation Problems

- Uncovering causal relationships
- The generic problem: Counterfactuals are missing

The Evaluation Problems

- Uncovering causal relationships
- The generic problem: Counterfactuals are missing
- Specific Obstacles: selection and endogeneity

Uncovering Causal Relationships

Correlation is not causality!

- Looking at what happens after the introduction of a policy is not proper evaluation
  - Long term trends
  - Macroeconomic changes

Uncovering Causal Relationships

Correlation is not causality!

- Looking at what happens after the introduction of a policy is not proper evaluation
  - Long term trends
  - Macroeconomic changes
- Naively comparing outcomes for those who received the policy with those that did not receive the policy may also not be valid evaluation
  - Selection Effects

Uncovering Causal Relationships

- We want to establish whether and how a policy, T affects the outcome of interest Y_i; $i = \{1, ..., N\}$
  - Y_i could be health, income, employment, tax payments, etc

Uncovering Causal Relationships

- We want to establish whether and how a policy, T affects the outcome of interest Y_i; $i = \{1, ..., N\}$
    - Y_i could be health, income, employment, tax payments, etc
- Looking for counterfactuals: What would have happened to this person's behaviour in the absence of the policy or under an alternative policy?
    - Example: Do people earn more when they have more education?

The Missing Counterfactual

- The causal effect of treatment T on unit i as captured by outcome Y is $\alpha_i = Y_i^1 - Y_i^0$

The Missing Counterfactual

- The causal effect of treatment T on unit i as captured by outcome Y is $\alpha_i = Y_i^1 - Y_i^0$
- Problem: impossible to observe simultaneously both Y_i^1 and Y_i^0 for the same $i \rightarrow$ impossible to observe α_i

The Missing Counterfactual

- The causal effect of treatment T on unit i as captured by outcome Y is $\alpha_i = Y_i^1 - Y_i^0$
- Problem: impossible to observe simultaneously both Y_i^1 and Y_i^0 for the same i $\rightarrow$ impossible to observe α_i
- **Missing data problem**

The Statistical Solution

- Use population to compute the average causal effect of T over all agents, $i = \{1, ..., N\}$

The Statistical Solution

- Use population to compute the average causal effect of T over all agents, $i = \{1, ..., N\}$
- Average outcome of the treated: $E(Y_i^1 | T_i = 1)$

The Statistical Solution

- Use population to compute the average causal effect of T over all agents, $i = \{1, ..., N\}$
- Average outcome of the treated: $E(Y_i^1 | T_i = 1)$
- Average outcome of the control: $E(Y_i^0 | T_i = 0)$

The Statistical Solution

- Use population to compute the average causal effect of T over all agents, $i = \{1, ..., N\}$
- Average outcome of the treated: $E(Y_i^1 | T_i = 1)$
- Average outcome of the control: $E(Y_i^0 | T_i = 0)$
- Difference between averages:
  $D = E(Y_i^1 | T = 1) - E(Y_i^0 | T_i = 0)$

Selection Bias

- Compute the difference between averages

$$D = E(Y_i^1 | T = 1) - E(Y_i^0 | T_i = 0)$$

Selection Bias

- Compute the difference between averages

$$D = E(Y_i^1 | T = 1) - E(Y_i^0 | T_i = 0)$$

$$D = E(Y_i^1 - Y_i^0 | T = 1) + E(Y_i^0 | T = 1) - E(Y_i^0 | T = 0)$$

Selection Bias

- Compute the difference between averages

$$D = E(Y_i^1 | T = 1) - E(Y_i^0 | T_i = 0)$$

$$D = E(Y_i^1 - Y_i^0 | T = 1) + E(Y_i^0 | T = 1) - E(Y_i^0 | T = 0)$$

$$D = \alpha + E(Y_i^0 | T = 1) - E(Y_i^0 | T = 0)$$

Selection Bias

- Compute the difference between averages

$$D = E(Y_i^1 | T = 1) - E(Y_i^0 | T_i = 0)$$

$$D = E(Y_i^1 - Y_i^0 | T = 1) + E(Y_i^0 | T = 1) - E(Y_i^0 | T = 0)$$

$$D = \alpha + E(Y_i^0 | T = 1) - E(Y_i^0 | T = 0)$$

$D =$ average causal effect + selection bias

Selection Bias

- Compute the difference between averages

$$D = E(Y_i^1 | T = 1) - E(Y_i^0 | T_i = 0)$$

$$D = E(Y_i^1 - Y_i^0 | T = 1) + E(Y_i^0 | T = 1) - E(Y_i^0 | T = 0)$$

$$D = \alpha + E(Y_i^0 | T = 1) - E(Y_i^0 | T = 0)$$

$D =$ average causal effect + selection bias

- Examples of selection bias: Evaluating a policy to get the long-term unemployed back to work by comparing the employment outcomes of those who took up policy with those who didn't
  - Those who took up policy may have been more motivated and hence more likely to get a job anyways

Randomised Control Trials (RCTs)

- RCTs offer the best way of constructing a valid counterfactual

Randomised Control Trials (RCTs)

- RCTs offer the best way of constructing a valid counterfactual
- Solve the selection problem by randomly assigning units to the treatment

Randomised Control Trials (RCTs)

- RCTs offer the best way of constructing a valid counterfactual
- Solve the selection problem by randomly assigning units to the treatment
- Assignment to treatment is not based on any criterion related to the characteristics of units $\rightarrow$ it will be independent of the possible outcomes

Randomised Control Trials (RCTs)

- RCTs offer the best way of constructing a valid counterfactual
- Solve the selection problem by randomly assigning units to the treatment
- Assignment to treatment is not based on any criterion related to the characteristics of units $\rightarrow$ it will be independent of the possible outcomes
- $E(Y_i^0|T=1) = E(Y_i^0|T=0)$ now holds

Randomised Control Trials (RCTs)

- RCTs offer the best way of constructing a valid counterfactual
- Solve the selection problem by randomly assigning units to the treatment
- Assignment to treatment is not based on any criterion related to the characteristics of units $\rightarrow$ it will be independent of the possible outcomes
- $E(Y_i^0|T=1) = E(Y_i^0|T=0)$ now holds
- $D = E(Y_i^1|T=1) - E(Y_i^0|T=0) = \alpha = E(Y_i^1 - Y_i^0)$

Randomised Control Trials (RCTs)

- RCTs offer the best way of constructing a valid counterfactual
- Solve the selection problem by randomly assigning units to the treatment
- Assignment to treatment is not based on any criterion related to the characteristics of units $\rightarrow$ it will be independent of the possible outcomes
- $E(Y_i^0|T=1) = E(Y_i^0|T=0)$ now holds
- $D = E(Y_i^1|T=1) - E(Y_i^0|T=0) = \alpha = E(Y_i^1 - Y_i^0)$
- Can easily obtain convincing results, BUT only if trial has been well designed and implemented
    - Non-trivial issues

Designing RCTs

Partners

- RCTs involve working with partners, e.g. government, NGOs, who implement the policy/intervention
- Important to work with partners that understand the methodology of RCTs

Define study population

- What is the target population?

Define outcome of interest

- What is the main outcome targeted by the policy?
- Multiple outcomes

Designing RCTs

Choosing sample sizes, or the *power* of the design

- Power = probability that, for a given effect size, α and a defined level of statistical significance, we can reject the null hypothesis of no effect

Designing RCTs

Choosing sample sizes, or the *power* of the design

- Power = probability that, for a given effect size, α and a defined level of statistical significance, we can reject the null hypothesis of no effect
- Affected by sample size, proportion of the sample treated, variance of outcomes, level of randomisation, among other factors

Designing RCTs

Choosing sample sizes, or the *power* of the design

- Power = probability that, for a given effect size, α and a defined level of statistical significance, we can reject the null hypothesis of no effect
- Affected by sample size, proportion of the sample treated, variance of outcomes, level of randomisation, among other factors
- Large sample may be costly, but small sample may only be able to detect very large effects

Designing RCTs

Level of Randomisation:

- Individual or household or group (e.g. village, school, firm) level

Designing RCTs

Level of Randomisation:

- Individual or household or group (e.g. village, school, firm) level
- Choice depends on:
  - The intervention: who does it target?
  - Spillovers: Will the policy affect those not treated?
  - Implementation constraints:
    - Fixed costs of implementation $\rightarrow$ cost-efficient to randomise at the group level
    - Withholding policy from sub-set of group may cause resentment toward implementation organisation
    - Greater scope of mistakes from field staff

Challenges

- Expensive, and time consuming

Challenges

- Expensive, and time consuming
- Implementation issues:
  - Contamination, i.e. control group gets programme
  - Imperfect/Low take-up $\rightarrow$ can still estimate valid effects
  - Anticipation

Challenges

- Expensive, and time consuming
- Implementation issues:
    - Contamination, i.e. control group gets programme
    - Imperfect/Low take-up $\rightarrow$ can still estimate valid effects
    - Anticipation
- Hawthorne and John Henry effects: Units react differently because they know that they are part of an experiment

Limits to RCTs

- RCTs provide us with an answer about whether or not a policy or intervention worked or didn't, *on average*

Limits to RCTs

- RCTs provide us with an answer about whether or not a policy or intervention worked or didn't, *on average*
- They cannot tell us why a policy worked or didn't
    - Sheds light on suitable alternatives if it didn't work

Limits to RCTs

- RCTs provide us with an answer about whether or not a policy or intervention worked or didn't, *on average*
- They cannot tell us why a policy worked or didn't
    - Sheds light on suitable alternatives if it didn't work
- They cannot be used to evaluate policies that cannot be excluded from some individuals or groups, for example, monetary policy

Limits to RCTs

- External Validity
  - Would the policy work in another setting?
    - Understanding why a policy worked can help in this

Limits to RCTs

- External Validity
  - Would the policy work in another setting?
    - Understanding why a policy worked can help in this
  - Would it be as effective if implemented by another provider?

Limits to RCTs

- External Validity
  - Would the policy work in another setting?
    - Understanding why a policy worked can help in this
  - Would it be as effective if implemented by another provider?
  - Would impacts be the same if implemented at a different scale?
    - Pilot project vs. Nationwide rollout

Role of Economic Modelling

- Economic theory provides us with tools to model the behaviour of individuals, households and firms
    - Model constraints faced by agents

Role of Economic Modelling

- Economic theory provides us with tools to model the behaviour of individuals, households and firms
  - Model constraints faced by agents
- Uncover channels through which outcomes change
  - These can be tested
  - Detailed example to follow

Role of Economic Modelling

- Economic theory provides us with tools to model the behaviour of individuals, households and firms
    - Model constraints faced by agents
- Uncover channels through which outcomes change
    - These can be tested
    - Detailed example to follow
- Researchers have been designing RCTs in order to test economic theory
    - Example: Ashraf, Karlan and Yin (2006) test the importance of time-inconsistent preferences in explaining saving behaviour

Application: Providing Information on Child Nutrition in Rural Malawi

- Evaluate a cluster randomized control trial in rural Malawi which provided mothers with information on child nutrition via home visits
  - No other resources were provided $\rightarrow$ intervention works by shifting parental knowledge

Application: Providing Information on Child Nutrition in Rural Malawi

- Evaluate a cluster randomized control trial in rural Malawi which provided mothers with information on child nutrition via home visits
  - No other resources were provided $\rightarrow$ intervention works by shifting parental knowledge
- Interested in:
  - Impacts on child health, the key outcome the intervention intended to improve
  - Understanding how the impacts were realised

Setting: Mchinji (Malawi)

- Child health is very poor in Malawi
  - Infant mortality rate of 133 per 1000 births (UK rate: 5 per 1000)
  - 48% of kids aged < 5 years are too short for their weight (i.e. stunted)

Setting: Mchinji (Malawi)

- Child health is very poor in Malawi
    - Infant mortality rate of 133 per 1000 births (UK rate: 5 per 1000)
    - 48% of kids aged < 5 years are too short for their weight (i.e. stunted)
- Factors influencing child health include prenatal maternal behaviour, nutrition, disease environment

Setting: Mchinji (Malawi)

- Child health is very poor in Malawi
  - Infant mortality rate of 133 per 1000 births (UK rate: 5 per 1000)
  - 48% of kids aged < 5 years are too short for their weight (i.e. stunted)
- Factors influencing child health include prenatal maternal behaviour, nutrition, disease environment
- One constraint driving such poor health outcomes is that households don't know how best to feed their infants
  - Common to give porridge with unsterilized water to infants as young as 1 week
  - 1/3 of kids aged 6-54 months do not consume any proteins over a 3 day period

The Intervention

- Set up in 2005 by Mai Mwana, a research and development project that aims to improve maternal and child health, in collaboration with health researchers at UCL

The Intervention

- Set up in 2005 by Mai Mwana, a research and development project that aims to improve maternal and child health, in collaboration with health researchers at UCL
- Trained local women ("peer counsellors") provide information and advice on infant feeding to mothers of babies aged < 6 months
    - 5 home visits: once before birth, 4 times after birth
    - Visit content: primary focus on exclusive breastfeeding and post-breastfeeding nutrition

The Intervention

- Set up in 2005 by Mai Mwana, a research and development project that aims to improve maternal and child health, in collaboration with health researchers at UCL
- Trained local women ("peer counsellors") provide information and advice on infant feeding to mothers of babies aged < 6 months
  - 5 home visits: once before birth, 4 times after birth
  - Visit content: primary focus on exclusive breastfeeding and post-breastfeeding nutrition
- All pregnant women in a cluster are eligible for the intervention:
  - Around 60% are visited $\rightarrow$ compliance was not perfect

The Intervention

- Set up in 2005 by Mai Mwana, a research and development project that aims to improve maternal and child health, in collaboration with health researchers at UCL
- Trained local women ("peer counsellors") provide information and advice on infant feeding to mothers of babies aged < 6 months
  - 5 home visits: once before birth, 4 times after birth
  - Visit content: primary focus on exclusive breastfeeding and post-breastfeeding nutrition
- All pregnant women in a cluster are eligible for the intervention:
  - Around 60% are visited $\rightarrow$ compliance was not perfect
- Intervention began in July 2005 and is still on-going

Experimental Design

- Mai Mwana wanted to evaluate the effectiveness of the intervention $\rightarrow$RCT

Experimental Design

- Mai Mwana wanted to evaluate the effectiveness of the intervention $\rightarrow$RCT
- Considerations taken into account when designing the experiment include:
  - Spillovers, esp. cross-village spillovers
  - Implementation costs

Experimental Design

- Mchinji District divided into 48 clusters with ~ 8000 individuals each

Experimental Design

- Mchinji District divided into 48 clusters with ~ 8000 individuals each
- Within each cluster, the villages closest to the geographical centre chosen to be part of the study area (~ 3000 individuals)

Experimental Design

- Mchinji District divided into 48 clusters with ~ 8000 individuals each
- Within each cluster, the villages closest to the geographical centre chosen to be part of the study area (~ 3000 individuals)
- Natural buffer area, limiting contamination between neighbouring clusters

Experimental Design

- Mchinji District divided into 48 clusters with ~ 8000 individuals each
- Within each cluster, the villages closest to the geographical centre chosen to be part of the study area (~ 3000 individuals)
- Natural buffer area, limiting contamination between neighbouring clusters
- 12 clusters randomly assigned the intervention, 12 clusters serve as controls

Experimental Design

- Mchinji District divided into 48 clusters with ~ 8000 individuals each
- Within each cluster, the villages closest to the geographical centre chosen to be part of the study area (~ 3000 individuals)
- Natural buffer area, limiting contamination between neighbouring clusters
- 12 clusters randomly assigned the intervention, 12 clusters serve as controls
- Remaining 24 received another intervention focused on maternal health

Experimental Design

Theory

- Households with an adult decision-maker, who cares about his own consumption, A and leisure, L and the child's health, H

Theory

- Households with an adult decision-maker, who cares about his own consumption, A and leisure, L and the child's health, H
- Child health, H is a function of two things:
  - The amount a child eats, C
  - How efficiently the household transforms the consumption to child health, θ, which captures things such as knowledge

Theory

- Households with an adult decision-maker, who cares about his own consumption, A and leisure, L and the child's health, H
- Child health, H is a function of two things:
    - The amount a child eats, C
    - How efficiently the household transforms the consumption to child health, θ, which captures things such as knowledge
- Households have limited resources $\rightarrow$ budget constraint

Theory

- Providing mothers with information on child nutrition increases θ

Theory

- Providing mothers with information on child nutrition increases θ
- Parents realise that the child's consumption is more important than they had thought in shaping their health
    - Increase in child consumption, C

Theory

- Providing mothers with information on child nutrition increases θ
- Parents realise that the child's consumption is more important than they had thought in shaping their health
    - Increase in child consumption, C
- But additional child consumption costs money and no additional resources were received

Theory

- Providing mothers with information on child nutrition increases θ
- Parents realise that the child's consumption is more important than they had thought in shaping their health
  - Increase in child consumption, C
- But additional child consumption costs money and no additional resources were received
- To fund this extra child consumption, adults can reduce their leisure, L and/or reduce their own consumption, A
  - Depends on degree of substitutability of A and L

Theory

- Providing mothers with information on child nutrition increases θ
- Parents realise that the child's consumption is more important than they had thought in shaping their health
  - Increase in child consumption, C
- But additional child consumption costs money and no additional resources were received
- To fund this extra child consumption, adults can reduce their leisure, L and/or reduce their own consumption, A
  - Depends on degree of substitutability of A and L
- Increase in C is more than decrease in A $\rightarrow$ total household consumption will increase

Data

- Baseline census of all women aged 10-49 years in the study areas conducted by Mai Mwana in 2004

Data

- Baseline census of all women aged 10-49 years in the study areas conducted by Mai Mwana in 2004
- Follow-up data collected in 2008-09 and 2009-10
  - Sample of 1660 women and their households
  - Significant attrition between 2004 and 2008, but results robust to encountered attrition

Data

- Baseline census of all women aged 10-49 years in the study areas conducted by Mai Mwana in 2004
- Follow-up data collected in 2008-09 and 2009-10
    - Sample of 1660 women and their households
    - Significant attrition between 2004 and 2008, but results robust to encountered attrition
- The sample is balanced on a broad set of woman and household socio-economic characteristics, suggesting randomisation worked

Empirical Model

- With an RCT, we can simply compare means

$$Y_{ict} = \alpha + \beta_1 T_c + \mathbf{X}'_{ict}\beta_2 + \mathbf{Z}'_{c0}\beta_3 + \mu_t + u_{ict}$$

Empirical Model

- With an RCT, we can simply compare means
- Regression framework allows us to include covariates to improve estimation efficiency

$$Y_{ict} = \alpha + \beta_1 T_c + \mathbf{X}'_{ict}\beta_2 + \mathbf{Z}'_{c0}\beta_3 + \mu_t + u_{ict}$$

Empirical Model

- With an RCT, we can simply compare means
- Regression framework allows us to include covariates to improve estimation efficiency

$$Y_{ict} = \alpha + \beta_1 T_c + \mathbf{X}'_{ict}\beta_2 + \mathbf{Z}'_{c0}\beta_3 + \mu_t + u_{ict}$$

- $T_c = 1$ if main respondent in the follow-up survey resided in a treated cluster in 2004

Empirical Model

- With an RCT, we can simply compare means
- Regression framework allows us to include covariates to improve estimation efficiency

$$Y_{ict} = \alpha + \beta_1 T_c + \mathbf{X}'_{ict}\beta_2 + \mathbf{Z}'_{c0}\beta_3 + \mu_t + u_{ict}$$

- $T_c = 1$ if main respondent in the follow-up survey resided in a treated cluster in 2004
- Pool data from both follow-up surveys in our estimation

Empirical Model

- With an RCT, we can simply compare means
- Regression framework allows us to include covariates to improve estimation efficiency

$$Y_{ict} = \alpha + \beta_1 T_c + \mathbf{X}'_{ict}\beta_2 + \mathbf{Z}'_{c0}\beta_3 + \mu_t + u_{ict}$$

- $T_c = 1$ if main respondent in the follow-up survey resided in a treated cluster in 2004
- Pool data from both follow-up surveys in our estimation
- Inference:
    - Wild cluster bootstrap-t (Cameron, Gelbach and Miller 2008)
    - Randomization Inference (Fisher 1935; Rosenbaum 2002)

Empirical Framework

- Study impact of the following outcomes along the causal chain and suggested by theory to uncover how intervention worked:
    - Maternal nutritional knowledge
    - Child consumption
    - Household consumption
    - Labour supply
    - Child health

Nutritional Knowledge

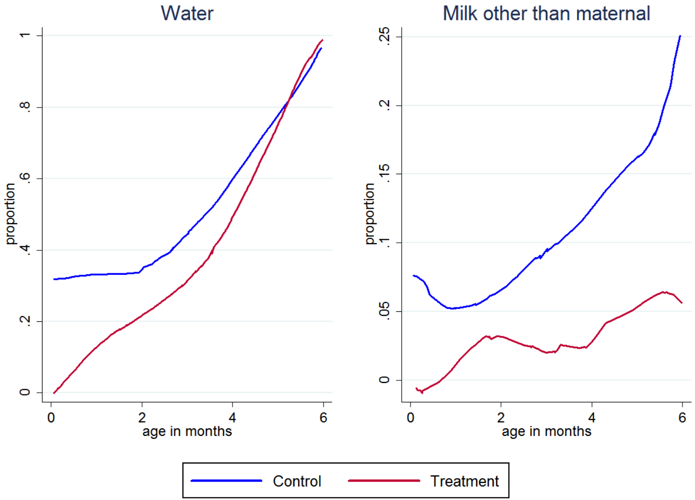
- Index computed from responses to 7 questions on child nutrition

	Summary Index	Breastfeeding when infant has diarrhoea	Best way of cooking fish with porridge for infant aged > 6 months?
T_c	0.169+	0.253+	0.067**
Standard Error	[0.086]	[0.115]	[0.019]
Wild Cluster Bootstrap p-value	{0.058}	{0.084}	{0.002}
Randomization Inference p-value	{0.065}	{0.028}	{0.008}
Observations	1512	1512	1512
IntraCluster Correlation	0.169	0.277	0.057
Mean, Control	-0.04	0.217	0.026

Notes: ** Significant at 1% level, * at 5% level, + at 10% level

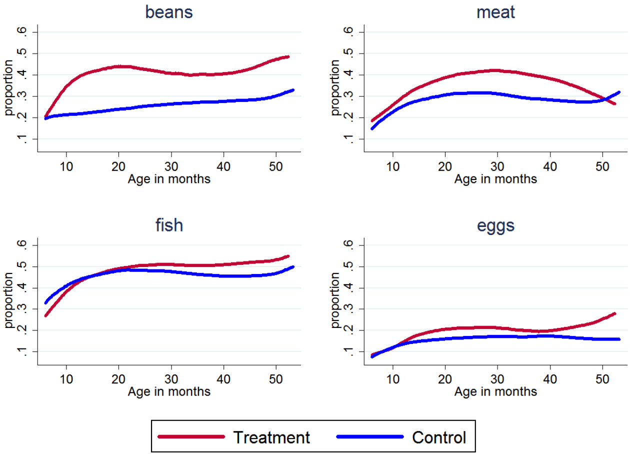
Child Consumption

- Significant improvements in consumption among children aged < 6 months:



Child Consumption II

- Significant improvements in diets of children >6 months and born after July 2005



Household Consumption

	[1]	[2]	[3]	[4]	[5]
		Per Capita Monthly Food Consumption for:			
	Summary Index	**Cereals**	**Proteins**	**Fruit and Vegetables**	**Other Foods**
T_z	0.218*	-9.878	128.359*	269.819+	60.453
Standard Error	[0.082]	[52.450]	[54.798]	[108.600]	[33.561]
Wild Cluster Bootstrap p-value	{0.018}	{0.931}	{0.022}	{0.060}	{0.150}
Randomization Inference p-value	{0.037}	{0.952}	{0.016}	{0.042}	{0.020}
Observations	3200	3205	3202	3204	3204
R-squared	0.063	0.118	0.02	0.195	0.024
IntraCluster Correlation	0.087	0.074	0.042	0.172	0.053
Mean Control Areas	-0.10	606.00	349.80	679.70	149.70

Notes: ** Significant at 1% level, * at 5% level, + at 10% level

How is the increased consumption funded?

	Male Adults			
	[1]	[2]	[3]	[4]
	Summary Index	Works	Has at least 2 jobs	Weekly Hours Worked
T_z	0.262+	0.096	0.072*	4.31
Standard Error	[0.131]	[0.078]	[0.028]	[2.918]
Wild Cluster Bootstrap p-value	{0.074}	{0.303}	{0.020}	{0.230}
Randomization Inference p-value	{0.062}	{0.251}	{0.057}	{0.202}
Observations	3642	3961	3958	3642
R-squared	0.183	0.17	0.05	0.16
IntraCluster Correlation	0.146	0.208	0.036	0.100
Mean, Control	-0.135	0.836	0.122	25.740

Notes: ** Significant at 1% level, * at 5% level, + at 10% level

- No impacts on female labor supply

Child Health

	[1]	[2]	[3]	[4]
	Summary Index	**Height for Age**	**Healthy weight for age**	**Healthy weight for height**
T_z	0.102*	0.271*	0.030	0.048
Standard Error	[0.036]	[0.102]	[0.019]	[0.027]
Wild Cluster Bootstrap p-value	{0.022}	{0.022}	{0.150}	{0.132}
Randomization Inference p-value	{0.035}	{0.055}	{0.312}	{0.147}
Observations	2175	2192	2265	2217
R-squared	0.026	0.046	0.024	0.029
IntraCluster Correlation	0.021	0.022	0.018	0.017
Average, Control	0.266	-2.338	0.817	0.845

Notes: ** Significant at 1% level, * at 5% level, + at 10% level

- Positive but statistically insignificant impacts on physical growth for children aged < 6 months

Concluding Remarks

- RCTs can provide a credible counterfactual group for policy evaluation
- Gold standard evaluation method
- Careful design needed, taking into account implementation constraints
- Limitations to what they can tell us
- Economic modelling has a role to play

References

- "Applying Behavioural Insights to reduce fraud, error and debt", Cabinet Office Report
- Banerjee, A. and E. Duflo, "Poor Economics"
- Duflo, E., Kremer, M., and R. Glennester, "Using Randomization in Development Economics Research: A Toolkit", Handbook of Development Economics.
- Fitzsimons, E., Malde, B., Mesnard, A., and M. Vera-Hernandez, (2013), "Household Responses to Information on Child Nutrition: Experimental Evidence from Rural Malawi", IFS WP.
- EDePo website: www.ifs.org.uk/centres/EDePo