CHAPTER 5

Randomised Controlled Trials and Quasi-Experiments

DO Hoang Van Khanh and Sharon THAM

"Once you have a plan in place, all you have to do is apply a bit of coin flipping, or randomisation. You'll want to compare the difference in results between a 'control' and an 'experimental' situation."

(Uri Gneezy and John List 2013, 242)

What Is a Randomised Controlled Trial?

A randomised controlled trial (RCT) is a study in which the studied sample is randomly split into a control group (those who will receive the standard intervention) and a treatment group (those who will receive the intervention being tested) (see **Figure 1**).

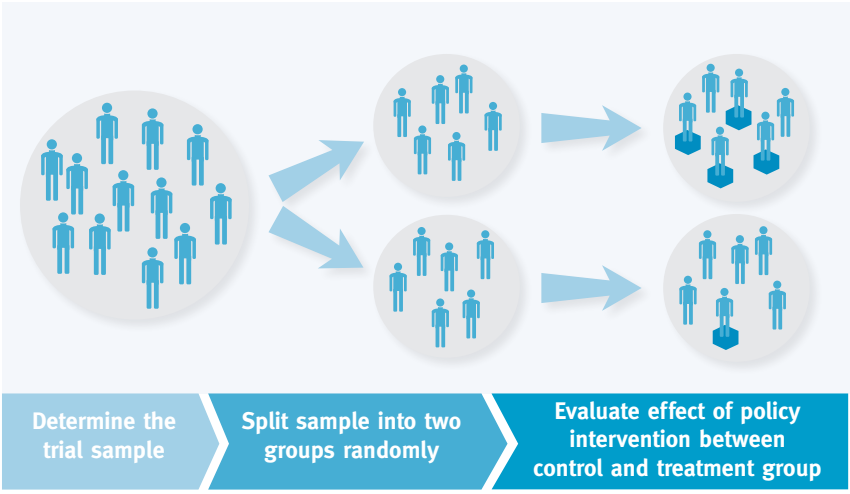

Figure 1: Conducting a Randomised Controlled Trial

Source: Do and Tham 2013

Random allocation means participants would have an equal chance of being assigned to either of these groups. If the sample size is large enough, the control and treatment groups can be expected to have the same characteristics. In addition, procedures are controlled to ensure that all participants in both groups are treated the same except for the new intervention that is being tested. The results of the trial can then be attributed to the new intervention instead of other confounding factors[1].

RCTs can sometimes be confused with simple pilots, since both are associated with testing a new intervention on a selected group of people. The key difference is that pilots do not require a concurrent control group and allocation is not always random (see **Box 1**).

BOX 1

What Is the Difference between an RCT and a Simple Pilot (Pre-Post Test)?

Pilots (pre-post tests) are studies, in which an intervention is introduced to a selected group (treatment group). Changes observed in the *same group* are then measured over time. The results of the intervention are thus measured as the difference in observations before and after the intervention is introduced.

Unlike RCTs, pilots do not have a concurrent control group that could account for possible confounding factors—factors other than the intervention being tested that could influence the outcome. These

1. A confounding factor is an extraneous variable whose presence affects the variables being studied so that the results you get do not reflect the actual relationship between the variables under investigation.

factors could lead to distortions (downwards or upwards) in the observed outcome of the intervention. Hence, an evaluation based on pilot results is less accurate and rigorous compared to an RCT.

In practice, pilots are commonly used by governments to test out policies in many domains, such as education, energy and transport, before they are implemented at a national level. However, in many instances, pilots could have been easily designed as RCTs to obtain more reliable data on the effects of the intervention. This simply entails recruiting and collecting data concurrently on a comparable control group to more accurately ascertain the effects of the intervention.

RCTs are thus considered the golden standard for evaluation because they enable a robust and clean evaluation of how effective a new intervention is. The key advantages of using an RCT design compared to other evaluation methods are that:

i. Random assignment ensures that factors (known and unknown) that could affect the outcome of the trial are evenly distributed across conditions; and

ii. Random assignment and the use of a control group eliminate the potential effects of confounding factors on the results. This establishes a **causal** relationship between the intervention tested and the difference in outcomes observed between the treatment and control groups.

For public policy, RCTs are useful in evaluating the overall effectiveness of programmes. This helps policymakers to determine where resources should be channelled and how they should be used (see **Box 2**). RCTs are also useful in testing which aspects of a programme yield the greatest effect towards desired outcomes, which is valuable information for policy refinements.

BOX 2

The Effects of Exposure to Better Neighbourhoods on Children: The Moving to Opportunity Experiment

The Moving to Opportunity (MTO) experiment offered randomly selected families living in high-poverty housing projects in the U.S. housing vouchers to move to lower-poverty neighbourhoods. The experiment used administrative data from tax returns to access the impacts of MTO on children's long-term outcomes.

The study found that moving to a lower-poverty neighbourhood had significantly positive impacts on children who were young (below the age of 13) when their families moved. In their mid-twenties, children from the treatment group were likely to have an annual income that is $3,477 (31%) higher, on average, relative to their counterparts in the control group. Children in the treatment group also lived in better neighbourhoods themselves as adults, were less likely to become single parents, and had higher college attendance rates.

On the other hand, for children who were older than 13 years when their parents moved, the treatment had negative long-term impacts, perhaps because of disruption effects.

The findings imply that resources should be channelled to programmes that help families with young children move from high-poverty neighbourhoods to lower-poverty neighbourhoods. These programmes would not only reduce intergenerational persistence of poverty but might ultimately generate positive returns for taxpayers.

Source: Chetty et al. 2015

When Can RCTs Be Used?

RCTs can be a powerful tool to test and evaluate policy, especially if:

i. The intervention is well-defined and/or simple and easy to standardise, e.g., sending reminder text messages, altering payment structure;
ii. The outcome of the treatment can be easily and consistently measured;
iii. The intervention can be implemented consistently; and
iv. The intervention is likely to have large impact.

However, RCTs would not work well in certain situations. For example, it is impossible to randomise the treatment when a policy has already been rolled out to a selective group, potentially resulting in self-selection or sampling biases. RCTs are also not appropriate if there are likely to be interferences between treatment and control groups; and when it is not possible to ensure minimal attrition and good compliance of the treatment. To illustrate, a hospital planning to test a new house visit programme to improve health outcomes would not be able to conduct a robust RCT if large numbers of patients are likely to drop out or refuse to continue with the treatment halfway.

Under such circumstances, the next best alternative would be to explore the use of quasi-experiments to evaluate the interventions. The reliability and rigour of quasi-experiments can range from a very simple pre-post analysis to more sophisticated techniques like statistical matching (or propensity-score matching) and regression discontinuity design, which are much closer to the standard of RCTs.[2]

To ensure rigour, quasi-experiments may require extensive knowledge, context and a literature review to identify all possible external factors, other than the treatment, that can affect the outcome. Researchers and policymakers may also need to carry out further data analyses (e.g., statistical regression) to control for these external factors.

2. See Listing of Experimental Methods at: http://www.povertyactionlab.org/sites/default/files/documents/Experimental%20Methodology%20Table.pdf.

Challenges and Limitations

Ethical Concerns

While RCTs are considered the golden standard for evaluations, there may be concerns that they are unethical or unfair because a new intervention is being withheld from people who could benefit from it. This concern is heightened when additional money is being spent on programmes which might improve the health, wealth, or educational experience of the treatment group.

In this respect, policies that are planned to be rolled out slowly and on a staggered basis—due to financial or other resource constraints—would present natural opportunities for experimentation in the public policy sphere (see **Box 3**). In such cases, even without RCT, only a selected group of citizens would have received the intervention via pilots. By selecting a group to receive the treatment through randomisation and by tracking the data of a comparable control group, policymakers can assess the effectiveness of the intervention in a more rigorous manner.

BOX 3

A Natural Opportunity for RCT: PROGRESA (Conditional Cash Transfer Programme in Mexico)

PROGRESA combines a traditional cash transfer programme with financial incentives for recipient families to invest in human capital (health, education, nutrition) of their children. In order to receive the cash transfers, families must accept preventive healthcare and participate in growth monitoring and nutrition supplements programmes.

Due to resource constraints, approximately 10% (506) of the 50,000 PROGRESA eligible communities were chosen to receive the programme immediately, while the rest would receive it two years later. This resulted in a natural opportunity to conduct an RCT for the programme. The communities were randomly assigned to the treatment group that received PROGRESA first, while the rest became part of the control group that received the benefits later. The trial showed that the utilisation of public health clinics increased faster in PROGRESA villages than in control areas, with significant improvements in the health of both adult and child beneficiaries.

Source: Gertler and Boyce 2001

The "wait list" approach is an alternative way to address ethical concerns, especially if the intervention is anticipated to be popular or favoured by a large number of people. In this instance, policymakers can solicit sign-ups for limited slots in a programme. Once the slots are filled, the participants are randomly assigned to the treatment group (those who would receive the intervention immediately) while the rest would be assigned to the wait list or control group (those who would receive the intervention at a later stage or after the trial has been completed). To qualify for the wait list, participants would need to agree to be tracked during the trial period. With this approach, everyone who signed up would have a chance to benefit from the intervention, albeit with some time difference, hence minimising ethical concerns.

Interestingly, from a citizen's perspective, *pilots are not too different from RCTs*. Both include a small sample receiving the treatment or intervention prior to the rest of the population. Hence, they share common ethical concerns, centred on the "unfairness" of only a selected group benefiting from an intervention. This means that whenever pilots are considered, RCTs can easily substitute pilots to yield more rigorous and reliable evaluation without aggravating ethical concerns.

Administrative Challenges (Cost and Time)

In reality, complete randomisation is often extremely costly and may take a long time to administer. This is one of the reasons why RCTs are often substituted by simple experiments or pilots, even though they lack rigour and reliability and are prone to selection bias.

However, high costs can be overcome in certain scenarios. For instance, experiments could be carried out on current policy when outcome data is already being collected from routine monitoring systems (whether administrative or survey data). In such cases, the cost of the experiment can be narrowed down to the time taken to design and set up the trial. In addition, if randomisation at the individual level is impossible or too costly, policymakers could randomise at larger units of measurement (i.e., at cluster level)—such as by classes in schools or housing blocks.

In the case of timelines, while RCTs may require a longer planning time, the actual experiment itself should not take longer than a simple pilot.

Political Considerations

Sometimes, experimentations in policy might not be possible due to political reasons, i.e., some policies may have to be implemented within a very short time frame without requiring evidence on their effectiveness. In these instances, RCTs can still be used to provide insights on how certain aspects of a policy can be tweaked for greater effectiveness. For example, as researchers could not test

the overall effectiveness of the Singapore Work Support Programme since the policy had already been implemented on a national level in 2011, they sought to test different versions of the policy instead, varying the amount and duration of assistance (Do and Tham 2013). This was to find out how best to help recipients attain financial independence and sustained employment over time.

Conclusion

Conducting a pure RCT might not always be possible in public policy as carrying out complete randomisation and having an ideal control group can be challenging. However, this should not discourage policymakers from using RCTs or setting some RCT standards when designing and evaluating policies. To ensure a high level of rigour in the analysis, the following three key principles should be kept in mind:

i. ***Randomise*** the selection of people to be in the control and treatment groups as far as possible;

ii. ***Track*** data for the most comparable control group possible, before the data is lost or difficult to retrieve thereafter; and

iii. ***Be aware*** and clearly state the possible confounding factors in the trial.

References

Cabinet Office and Behavioural Insights Team. "Test, Learn, Adapt: Developing Public Policy with Randomized Controlled Trials". 2012. Accessed January 30, 2015. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/62529/TLA-1906126.pdf.

Chetty R. N. Hendren, and L. F. Katz. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment". *American Economic Review* 106, no. 4 (April 2016): 855–902. Accessed May 31, 2016. http://www.aeaweb.org/articles?id=10.1257/aer.20150572.

Do, H. V. K., and S. Tham. "Randomized Controlled Trials in Policymaking". *Ethos* 12 (June 2013): 36–44. Accessed January 30, 2015. https://www.cscollege.gov.sg/Knowledge/Ethos/Ethos%20Issue%2012%20June%202013/Pages/Randomised%20Controlled%20Trials%20in%20Policymaking.aspx.

Gertler, P. J., and S. Boyce. "An Experiment in Incentive-Based Welfare: The Impact of PROGRESA on Health in Mexico". Working paper, April 3, 2001. Accessed January 30, 2015. http://faculty.haas.berkeley.edu/gertler/working_papers/PROGRESA%204-01.pdf.

Gneezy, U., and J. List. The *Why Axis: Hidden Motives and the Undiscovered Economics of Everyday Life*. USA: Public Affairs, 2013.