

Assessing Causality: A Case for Randomized Controlled Trials

Adrien Bouguen [J-PAL Europe, PSE – abouguen@povertyactionlab.org]

Marc Gurgand [J-PAL Europe, Paris School of Economics – gurgand@pse.ens.fr]

Randomized Controlled Trial (RCT) is an experimental method used to rigorously evaluate the effects of education policies. Experience shows that running RCTs is taking place in Europe, and that findings from one European country can be relevant for others. The European Union could therefore explore opportunities to experiment, evaluate, and disseminate findings.

The European Union has long been committed to achieving ambitious objectives in term of education, as the long-term effects of educational skills on various aspects of success in life are firmly established. However, given the current public spending crisis and the already relatively high share of GDP devoted to education budgets in Europe, how can Member States improve overall education without unduly increasing the public deficit?

One obvious response would be to carefully choose the most efficient policies. To do so, precise measures and rigorous methodology need to be implemented to evaluate educational programs such as class-size reduction, boarding schools, remedial education programs, or new pedagogies. Such programs are typically very costly, and often hard to terminate once they have been scaled up. A rigorous evaluation of their effect beforehand is thus desirable.

A TRANSPARENT METHODOLOGY

Simply comparing a group of beneficiaries with a group of non-beneficiaries is unfortunately unable to yield useful results as the two groups tend to be intrinsically different. Some non-experimental methods (matching or multivariate regression), although valuable, hardly deal with unobserved characteristics (motivation, intelligence, readiness, ...) that might differ between the groups of participants and non-participants. Such non-observed characteristics are nevertheless crucial to account for

in the education context.

One alternative way to evaluate the effect of an education program (class-size reduction, remedial education, ...) is to establish an experimental setting – a Randomized Controlled Trial (RCT) – in which the outcome (test scores, drop-out rates, etc.) obtained by the beneficiary students (the so-called “treatment group”) will be compared to the ones of similar non-beneficiaries (the “control group”). To ensure that both groups are statistically identical before the beginning of the program, experimental groups of eligible students (or schools) are formed randomly.

When assigned randomly, control and treatment groups are expected to be statistically identical before the program is implemented. Consequently, at the end of the evaluation, any difference in the outcome can be interpreted as the impact of the program – the treatment effect.

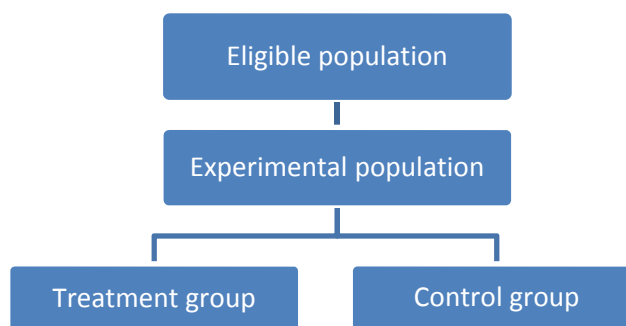
RECENT EUROPEAN EXAMPLES

In recent years, several large-scale experiments have been run in European countries in the education sector. For instance, in 2007, the French Government evaluated a parental involvement program targeted at sixth graders that was randomly allocated to a subset of 200 classes. This simple program was shown to strongly improve students’ behavior, in particular truancy. In the wake of these positive results, the French government has decided to scale up the program.

External and internal validity in randomized controlled trials (RCTs)

Out of the eligible population, a subset of individuals or schools are chosen.

Random selection ensures that both experimental groups have the same characteristics.



Since September 2012, the Danish government has been evaluating the impact of providing extra teachers to schools in order to help teachers with Danish and math. Out of 105 schools, 35 are to receive an extra teacher, 35 will hire an extra resource person who does not need to be a teacher, and 35 schools will form the control group. This is a less costly alternative to class-size reduction, but one whose impact is unknown.

MANY OPPORTUNITIES TO RANDOMIZE

Although RCTs are used in many contexts, this method is particularly well suited for evaluating new education policies. Indeed, conversely to non-experimental methods that can evaluate policies after they started, RCTs can only be used for new education programs that are not yet scaled-up. In addition, RCTs require that the groups of participants and non-participants in the program are formed randomly before the program starts.

This latter constraint is often considered as ethically or politically delicate. In principle, excluding candidates using random assignment does not deprive them from any benefit as the program’s impact is unknown. Yet in many cases, candidates may value the participation in the program per se without any consideration of impact (think of a cash transfer program). But several opportunities allow this random allocation. For instance, when a program is introduced in few schools before being scaled-up, it is relatively convenient to randomly assign which schools will first benefit from

the program (phase-in method). Likewise, when the demand for a service provided by the program exceeds the supply (restricted number of dorms in a boarding school, for instance), it is rather natural to randomly select a number of students out of the pool of eligible candidates. Random assignment is often very well accepted and can even be considered fairer than a selection based upon *ad hoc* criteria.

EXTERNAL VALIDITY AND REPLICATION AT THE EUROPEAN LEVEL

When properly designed, RCTs generate reliable, robust, and precise estimates of the impact of a program: in other words, they are said to have “internal validity” (see figure). However, since RCTs are implemented locally, some evaluations may lack “external validity”, i.e., it is not clear whether a result from one specific evaluation would carry over to other economic, social, and institutional contexts. This issue is in no way specific to RCTs and does not have one simple solution.

One possible way to verify a result found in one specific context is to replicate and evaluate the same program elsewhere. In line with the open method of coordination, the European Union would be a perfect venue to replicate programs that have proved successful in one Member State in other countries. European countries face many similar education issues and could benefit from cooperating on experiments, disseminating results, and advocating for efficient policies.

For more details see: Adrien Bouguen, Marc Gurgand, *Randomized Controlled Experiments in Education*. EENEE Analytical Report No. 11, February 2012, http://www.eenee.de/dms/EENEE/Analytical_Reports/EENEE_AR11.pdf.

European Expert Network on Economics of Education (EENEE)
Funded by the European Commission, DG Education and Culture



EENEE publications and further information at: <http://www.eenee.org>. – Coordination: Ifo Institute – Leibniz Institute for Economic Research at the University of Munich e.V., Poschingerstr. 5, 81679 Munich, Germany, Phone: +49-89-9224-1388, E-mail: eenee@ifo.de.

© EENEE 2012. All rights reserved. Opinions expressed in this brief are those of the author(s) alone and not of the European Commission.