THE EFFECTS OF A DROPOUT PREVENTION PROGRAM ON SECONDARY STUDENTS' OUTCOMES

Enrico Conti 1, Silvia Duranti1, Alessandra Mattei2, Fabrizia Mealli2 and Nicola Sciclone1

1 Istituto Regionale per la Programmazione Economica della Toscana (IR-PET), (e-mail: enrico.conti@irpet.it, silvia.duranti@irpet.it, nicola.sciclone@irpet.it

2 Dipartimento di Statistica, Informatica, Applicazioni, Università degli Studi di Firenze, (e-mail: mattei@disia.unifi.it, mealli@disia.unifi.it)


1 The Innovare Program: A Cluster-Randomized Experiment

Dropout from high school, and especially from vocational high schools, is a relevant phenomenon in Italy. In order to face the problem of early dropout in vocational high schools, the Regional government in Tuscany (a region in central Italy) promoted Innovare - INsegnare a chi NOOn Vuole imparARE - an Italian expression meaning “Teaching students who do not want to learn.” Innovare is a teacher-based dropout prevention program, aimed at reducing the number of early school leavers in the early grades of vocational high schools through the introduction of innovative teaching methods.

The Innovare study took place in the academic year 2013/2014 and involved a random sample of 53 first-year classes of 12 vocational schools in Tuscany that offer Initial vocational education and training (Istruzione e Formazione Professionale – IeFP).

The Innovare study is a cluster-randomized experiment, where classes are the unit of assignment: Classes are randomly assigned to either a treatment group or a control group. Specifically, classes participating in the study were randomly assigned to either a treatment group (18 classes) or a control group (35 classes). Teachers in classes assigned to the treatment group, that is the Innovare program, participated in 10 meetings with experts in education and epistemology, who helped them to design and implement innovative teaching methods, such as labs. Teachers in classes assigned to the control group used standard teaching methods.

In a cluster-randomized experiment, clusters are assigned to treatment or control, but often individuals are of interest. Thus the unit of assignment may
be different from the unit of analysis. Here we focus on evaluating causal effects of the Innovare program at cluster-level. Conti et al. (2014) also consider an individual-level analysis discussing advantages and drawbacks of both cluster and individual-level analysis.

A cluster-level analysis may provide useful information on the effectiveness of the intervention in reducing high-school dropout. At the cluster level, dropout is viewed as a social problem and focus is on interventions that can limit school dropout as a whole. The class is the natural unit of inference and standard methods for the analyses of randomized experiments can be applied at the cluster level.

2 Cluster-level analysis: A randomization inference approach

The aim of the study is to assess whether the Innovare program may improve class performances, reducing dropout, the probability of failing and absence rates. To this end we adopt a potential outcome approach to causal inference (e.g., Rubin, 2005), using randomization inference at cluster-level.

Let us first introduce some notation. The Innovare study involves $K = 53$ clusters (classes). A fixed number of $M = 18$ classes are randomly assigned to the active treatment and $K - M = 35$ classes are assigned to the control treatment. For each class $k$, let $W_k$ denote the treatment assignment: $W_k = 0$ for classes assigned to the standard treatment, $W_k = 1$ for those assigned the new treatment. Under Stable Unit Treatment Value Assumption (SUTVA, Rubin, 1980), for each class, we can define two potential outcomes for each response variable. Let $Y_k(w)$ be the potential outcome for the response variable $Y$ at cluster-level, given assignment to treatment level $w$. In the study we focus on five outcomes: failure, postponement of the evaluation, dropout, absence rate and failure or dropout.

For each class we observe only one potential outcome for each response variable, depending on the treatment actually assigned. Let $Y_k = Y_k(W_k)$ be the actual outcome observed at cluster-level. A vector of cluster-level pretreatment variables, $X_k$, is observed for each class, which includes both class-specific characteristics and group-average of individual-level pretreatment variables.

We evaluate the causal effect of the Innovare program using randomization inference, which allows us to draw exact inferences using only the random assignment of clusters to treatment or control. In randomization inference focus is on the observed sample, therefore sampling issues do not matter. Also no assumption is made about the underlying model that generated the data and the dependence structure of random cluster effects. Finally, the issue of interfer-
ence between students in the same cluster does not arise, because focus is on cluster-level: we assume that students in different classes do not interfere with each other.

In principle, randomization of the treatment implies the pretreatment variables being closely balanced in the two subsamples defined by treatment. However in our sample there exist some differences in background pretreatment variables between the treatment group and the control group. Therefore it seems sensible to account for differences occurring in classes background characteristics, viewing the Innovare study as a stratified cluster-randomized experiment (Small et al. 2008). A stratified randomized experiment is defined by a strongly ignorable assignment mechanism (Rosembaum & Rubin, 1983). Formally the treatment assignment mechanism is strongly ignorable if, for each

\[ k = 1, \ldots, K, \]

(i) \( Pr(W_k = 1|Y_k(0), Y_k(1), X_k) = Pr(W_k = 1|X_k) \); and (ii) \( 0 < Pr(W_k = 1|X_k) < 1 \), where \( Pr(W_k = 1|X_k) \) is called the propensity score. Let \( e(X_k) = Pr(W_k = 1|X_k) \). Rosenbaum & Rubin (1983) show that (i) the propensity score is a balancing score, i.e., \( Pr(X_k|W_k, e(X_k)) = Pr(X_k|e(X_k)) \), and (ii) if the treatment is randomly assigned within cells defined by the pretreatment variables, it is also randomly assigned within cells defined by propensity score:

\[ Pr(W_k = 1|Y_k(0), Y_k(1), e(X_k)) = Pr(W_k = 1|e(X_k)) \]

\[ k = 1, \ldots, K. \]

In our analysis, the propensity score is estimated using a logistic regression model. Based on the estimated propensity score, we restrict the analysis to the subsample of classes that satisfies an overlap condition, discarding four control classes and one treated class. Then, we re-estimated the propensity score using the selected subsample of 48 classes and use it to adjust treatment comparisons for differences in background covariates using sub-classification into \( H = 4 \) strata.

Under the assumption that data come from a stratified cluster randomized experiment, our cluster-level analysis uses randomization inference to draw exact inferences for our finite population (sample) of size \( K = 48 \). We adopt the Fisher Exact \( P \)−values approach (Fisher, 1925), focusing on the sharp null hypothesis of no effect of the treatment: \( H_0 : Y_k(0) = Y_k(1) \) for each \( k = 1, \ldots, K. \) Evidences against the null hypothesis will suggest that the Innovare program has some effect on classes’ performances.

We consider two test statistics: the difference in average outcomes by treatment status, \( S_{ave} \), and the difference in average ranks for treated and control units, \( S_{rank} \). Table 1 presents the observed values of the test statistics and the \( p \)−values for one-sided tests of significance. The observed values of the test statistics show some evidence that the new teaching method reduces the percentage of dropouts and the percentage of failures and the absence rate, and
Table 1. Observed values of the test statistics and corresponding p-values for the sharp null hypothesis $H_0: Y_k(1) - Y_k(0) = 0$ for all $k^a$.

<table>
<thead>
<tr>
<th>Outcome variable</th>
<th>$S_{obs}^{ave}$</th>
<th>$S_{obs}^{rank}$</th>
<th>$S_{ave}$</th>
<th>$S_{rank}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of failures</td>
<td>-2.78</td>
<td>-1.78</td>
<td>0.302</td>
<td>0.378</td>
</tr>
<tr>
<td>Percentage of dropouts</td>
<td>-2.41</td>
<td>-3.12</td>
<td>0.324</td>
<td>0.287</td>
</tr>
<tr>
<td>Absence rate (per cent)</td>
<td>-0.15</td>
<td>-1.79</td>
<td>0.467</td>
<td>0.362</td>
</tr>
<tr>
<td>Percentage of failures +</td>
<td>-5.19</td>
<td>-5.25</td>
<td>0.174</td>
<td>0.172</td>
</tr>
<tr>
<td>Percentage of dropouts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of postponements of the evaluation</td>
<td>5.87</td>
<td>5.77</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

$^a$ The p-values are estimated using 10,000 draws from the randomization distribution.

increases the percentage of postponements of the evaluation. The $p-$values for the one-sided tests of significance are quite large implying no significant results at least at the standard levels of significance. However they still suggest that the Innovare program may have some positive effect, reducing the percentage of negative events (failures and dropouts) and increasing the percentage of postponements of the evaluation.

References


