Economics of Education Review xxx (2011) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

How to improve pupils' literacy? A cost-effectiveness analysis of a French educational project

Sébastien Massoni*, Jean-Christophe Vergnaud

CES, Université Paris 1 Panthéon Sorbonne, Paris, France

ARTICLE INFO

Article history:
Received 29 September 2010
Received in revised form 29 June 2011
Accepted 25 August 2011

JEL classification:
C93
I20

Keywords:
Costs
Educational economics
Efficiency
Resource allocation

ABSTRACT

The *Action Lecture* program is an innovative teaching method run in some nursery and primary schools in Paris and designed to improve pupils' literacy. We report the results of an evaluation of this program. We describe the experimental protocol that was built to estimate the program's impact on several types of indicators. Data were processed following a Differences-in-Differences (DID) method. Then we use the estimation of the impact on academic achievement to conduct a cost-effectiveness analysis and take a reduction of the class size program as a benchmark. The results are positive for the *Action Lecture* program.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

It is well known that pupils who are good at reading and writing at school are also those who practice at home and like books [see PIRLS – *Progress in International Reading Literacy Study* results in 2001 (Mullis, Martin, Gonzalez, & Kennedy, 2003) and 2006 (Mullis, Martin, Kennedy, & Foy, 2007)]. Many teaching specialists consider that motivation for reading is central to acquiring literacy skills. Hence, they call for innovative teaching methods that induce positive attitudes towards reading. In this study we report a cost-effectiveness evaluation of one such method. The program we are interested in is a French educational project called *Action Lecture*, which is run in some nursery and primary schools in Paris. In practice, it takes place in volunteer schools in which pupils do not have any courses for 2 weeks, but work together on a specific topic with different activities (reading, research, museum visits, writing, etc.). The goal of this program is to develop the taste for reading and for discovery in order to promote academic achievement and to increase motivation for attending school. The main idea is to improve pupils' reading by a combination of learning activities and cultural activities in which schoolchildren are pushed to be active and to work collectively.

This evaluation has two aspects. First we perform an estimation of the impact. Second, we run a cost-effectiveness analysis and take a reduction of the class size program as a benchmark. As is often the case with innovative teaching methods, the bold ambitions of the *Action Lecture* program differ from official academic standards and this renders the evaluation problematic. Indeed, no assessments of pupils' achievements are routinely carried out during these programs, which we can rely on. Thus, to estimate the impact of this program we have to design a specific protocol. We focus on two kinds of indicators: academic standards with three different exercises related to different reading skills stemming from the French national evaluation scheme, and measures of attitude to reading

* Corresponding author at: Maison des Sciences Economiques, Office 402, 106-112 Boulevard de l'Hopital, 75647 Paris Cedex 13, France. Tel.: +33 1 44 07 82 06.
E-mail addresses: sebastien.massoni@gmail.com (S. Massoni), Jean-Christophe.Vergnaud@univ-paris1.fr (J.-C. Vergnaud).

0272-7757/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.econedurev.2011.08.013

Please cite this article in press as: Massoni, S., & Vergnaud, J.-C. How to improve pupils' literacy? A cost-effectiveness analysis of a French educational project. *Economics of Education Review* (2011), doi:10.1016/j.econedurev.2011.08.013

following the PIRLS' study. To estimate the impacts, we compare the progression of the pupils from the schools participating in the program with the evolution of pupils from a control group. We compute Differences-in-Differences to estimate the program's effect. As we find that *Action Lecture* has a significant and positive impact, we develop a cost-effectiveness analysis (Levin, 1995).

For the education system, the main costs of this program are the employment costs of the teachers appointed to the program. Therefore, we can relate one teaching job to its impact in terms of marks in the national evaluation scheme. It is useful that we also have data about class-size effects, provided by the study of Piketty and Valdenaire (2006). These are also expressed in terms of marks in France's national evaluation scheme. Thus, we can examine whether the resources devoted to the *Action Lecture* program could be used more efficiently by reassigning the teachers to classrooms. This is a topical subject since the French government intends to cut public spending and is reducing the number of teachers in the public school system, though the favored policy is to eliminate jobs which are not in the classroom, as such cuts are less visible for public opinion.

We find that the project studied here does have a positive impact on literacy. This is true for both types of indicators, i.e., academic standards and attitude scores to reading. The level of progress is quite important and we find that for the skills studied, these 2 weeks of teaching are equivalent to 40% of the average annual progress. Furthermore, compared to a class-size program, our conclusions concerning the efficiency of *Action Lecture* are positive.

The outline of this article is as follows: in Section 2 we present our methodology (data collection, evaluation methods); in Section 3 we show the main characteristics of our sample; in Section 4 we perform an estimation of the impact and a cost-effectiveness analysis; and Section 5 concludes.

2. The methodology

In this section we start with a short presentation of the *Action Lecture*, then we describe the experimental protocol that we use and finally we present our methodology.

2.1. A French educational project

The *Action Lecture* project is an educative program focused on reading that is jointly managed by the education system and the City of Paris, for nursery and primary schools.[1] The teaching methods used in this program are non-traditional and belong to problem oriented learning methods. They refer to the pedagogy promoted by Freinet (1896–1966), a French educationalist influencial in some French educational circles (Reuter, 2007). The main principle is to make the pupils active in their training and to leave them some freedom. In the case of the *Action Lecture* program, the idea is to associate culture and academic learning within one project. Reading is seen as a tool both to help pupils to obtain some specific academic skills and to develop cultural tastes. The underlying assumption is that there exists a link between learning and culture.

In practice, this program takes the following form: each volunteer school chooses a topic (for example: *Why are we writing?*, *Art, What is it for?*, etc.) and for 2 weeks the pupils do not have any other courses but work full-time on the project, in small teams (with a maximum of 15 pupils). Teams are heterogeneous with pupils from all grades working together. Presentations by teachers working only for the *Action Lecture* are also scheduled. These 2 weeks end with the production of a book that summarizes what was done. Even if the themes are school-related, the set-up is standardized: research on the topic (books, a museum, etc.) is done in the morning; teachers hold a meeting at lunch time to assess progress; and afternoons are devoted to technical work (writing, oral expression, methodological exercises, etc.).

The aims of the *Action Lecture* are to help pupils to be familiar with many books, to speak with expert readers, to have free time to read, to check their understanding of their readings, to write daily, and to improve their abilities in reading and writing exercises.

2.2. The experimental method and data

Since, the program does not include any evaluation of pupils' achievements, it was necessary to build an *ad hoc* method to estimate the program's impacts. The method includes a control group and consists in surveys administered before and after the project. Since the survey was computerized, only pupils from 2nd to 5th grades were included. The questionnaires used include several indicators as well as questions about individual characteristics.

To measure the impact of the project, several types of indicators were considered: the attitude toward reading (taste of reading, practice of reading, knowledge about books and authors, etc.), the attitude during school life (attitude during class, school life activities, self-evaluation, etc.), and academic abilities. For reading and school attitudes we re-used some questions from PIRLS. We will report two aggregated scores (on 10): the *Student's Attitude Toward Reading* (SATR) and the *Student's Reading Self Concept* (SCRC).[2] Measuring academic abilities is done using exercises issued from French national evaluations that are set at the beginning of 3rd and 6th grade. We use the 3rd grade evaluation exercises for 2nd and 3rd grade pupils and the 6th grade evaluation exercises for 4th and 5th grade pupils. Three types of skills have been studied: *identifying the nature or the type of a text*, *processing information*, and *making inferences*. These three skills represent 10% of the

[1] One specific aspect of the French education system is that local politicians as well as parents are not involved in teaching methods. This program, over which local authorities have some control, is uncommon.

[2] SATR is based on students' agreement with the following statements: I read only if I have to; I like talking about books with other people; I would be happy if someone gave me a book as a present; I think reading is boring; I enjoy reading. The SCRC is based on students' agreement with the following statements: reading is very easy for me; I do not read as well as other students in my class; and reading aloud is very hard for me.

Please cite this article in press as: Massoni, S., & Vergnaud, J.-C. How to improve pupils' literacy? A cost-effectiveness analysis of a French educational project. *Economics of Education Review* (2011), doi:10.1016/j.econedurev.2011.08.013

national evaluation of reading, which is marked out of 100 and thus we use a score out of 10 for these skills.

The collection of the individual characteristics was limited since it was not possible to send a questionnaire to families. A few individual characteristics were gathered directly from pupils: sex, age, month of birth, language spoken at home (the variable *French principally* says that the pupil 'always speaks' or 'almost always speaks' French at home, the variable *African languages* says that the pupil knows a sub-Saharan African language and similarly for *Arabic* and *Asian languages*), housing conditions (the variable *Own bedroom* says that the pupil has his/her own bedroom). Furthermore we have some overall data on the social composition of each school and this indicator is a good measure of pupils' social environment.

This data collection has been done with a set of three questionnaires completed on-line during school time. The timeline was the following: pupils replied to the first questionnaire 1 week before the implementation of the project, to the second in the week following its execution and to the third about 2 months later. Our analysis is focused on schools which followed the project between November 2007 and March 2008. Six schools were concerned and we gathered data on more than 400 pupils with around 100 pupils for each grade. In order to take into account this time gap in the data collection, we have used a variable *time of passage* which indicates the month during which the data was collected: it takes a value of 1 for September and 12 for August; furthermore if the date of passage was t for the first questionnaire, it takes the value $t+1$ for the second and $t+3$ for the third. The same timing has been respected for participating and control groups.

The first questionnaire was the longest, with 40 questions and 3 exercises. The second was the shortest with only 8 questions and 2 exercises, and the third contained 27 questions and 1 exercise. The exercises were different in each round, and to take into account differences in difficulty, the order of passage was randomized such that half of each class had the first order and the other half the second order.

Let us precise now how we select the treatment schools and the control schools. To benefit from an Action Lecture program, schools apply voluntarily then a selection committee chooses which schools to admit. Application to the program is open to all nursery and primary schools in Paris. The head teacher and his colleagues have to propose a project that is consistent with the Action Lecture guidelines (2 weeks without classes, intervention of external professors). During the year of the evaluation, the number of applicants was very low and all applicant schools were admitted into the program. Thus for the selection of the school, it was not possible to apply a standard randomized process to select the treatment schools and the control schools (see Duflo, Kremer, and Glennerster (2008) for the randomized methodology in evaluation). The control group was constituted by classes in non-concerned schools, from which we had to seek agreement. As this evaluation was quite intrusive for the class, the control group was relatively small. We were limited to three classes (3rd, 4th and 5th grades) that we chose in three different schools that were similar to the treated schools in terms of socio-economic characteristics.

2.3. The econometric model

The evaluation of this program is based on the *Differences-in-Differences* (DID) method which, since its development by Ashenfelter and Card (1985), has been mainly used in empirical economics (see Imbens and Wooldridge (2009) for a presentation of the different econometric models, and Bertrand, Duflo, and Mullainathan (2004) for a critical survey of the DID used in evaluation). The basic principle is to observe the values of outcomes for two groups (the group participating, affected by the program and the control group) between two periods (before and after the program) and to compute a double difference in the evolution of the outcomes: the average improvement of the control group over time is subtracted from the average improvement of the participating group. This double differencing allows correction of a twofold bias: first, the bias in the post-participation period between participating and control groups, which could be due to permanent differences between these two groups; second the bias from comparisons over time in the participating group, which could be due to the effect of time, unrelated to the participation. According to Cameron and Trivedi (2005), the Differences-in-Differences estimation allows to estimate the causal effect of the treatment if the time effects are common across treated and untreated individuals and if the composition of the treated and untreated groups is stable before and after the treatment.

The basic equation of the model is the following:

$$Y_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 A_{it} + \beta_3 A_{it} T_{it} + \epsilon_{it}$$

where Y_{it} is the outcome, T_{it} a dummy with a value of 1 if the subject belongs to the participating group, A_{it} a dummy of 1 if we are in post-participation period, and $A_{it}T_{it}$ the interaction of the two effects which captures the real impact of participation. An OLS regression of β_3 gives us an estimation of the participation effect. Table 1 shows the principle of the DID estimation.

In order to estimate the effect of the participation, we use different models with the integration of multiple groups (different schools), multiple levels (2nd and 3rd versus 4th and 5th), and multiple periods (different times of execution of the program). We improve our analysis step-by-step by estimating the following equations with fixed effects:

$$Y_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 A_{it} + \beta_3 A_{it} T_{it} + \alpha o_i + \epsilon_{it} \quad (1)$$

$$Y_{it} = \beta_0 + \beta_1 T_{it} + \beta_3 A_{it} T_{it} + \alpha o_i + \gamma v_t + \lambda l_i + \epsilon_{it} \quad (2)$$

$$Y_{it} = \beta_0 + \beta_3 A_{it} T_{it} + \alpha o_i + \gamma v_t + \lambda l_i + \tau u_i + \epsilon_{it} \quad (3)$$

$$Y_{it} = \beta_0 + \beta_3 A_{it} T_{it} + \alpha o_i + \gamma v_t + \lambda l_i + \tau u_i + \alpha X_{it} + \epsilon_{it} \quad (4)$$

$$Y_{it} = \beta_0 + \beta_3 A_{it} T_{it} l_1 + \beta'_3 A_{it} T_{it} l_2 + \alpha o_i + \gamma v_t l_1 + \gamma' v_t l_2 + \lambda l_i + \tau u_i + \alpha X_{it} + \epsilon_{it} \quad (5)$$

Model (1) is the basic estimation of the impact of participation, taking into account the order of exercises (o_i is a dummy variable); in Model (2) we add the time effects (v_t

Please cite this article in press as: Massoni, S., & Vergnaud, J.-C. How to improve pupils' literacy? A cost-effectiveness analysis of a French educational project. *Economics of Education Review* (2011), doi:10.1016/j.econedurev.2011.08.013

Table 1
Differences-in-Differences methodology.

	Before participation	After participation	Differences
Participating group	Y_{t1} $(\beta_0 + \beta_1)$	Y_{t2} $(\beta_0 + \beta_1 + \beta_2 + \beta_3)$	$\Delta Y_t = Y_{t2} - Y_{t1}$ $(\beta_2 + \beta_3)$
Control group	Y_{c1} (β_0)	Y_{c2} $(\beta_0 + \beta_2)$	$\Delta Y_c = Y_{c2} - Y_{c1}$ (β_2)
Differences			$\Delta\Delta Y = \Delta Y_t - Y_c$ (β_3)

is a time variable which takes the value of the month of the program's execution plus 12 months for the 3rd and the 5th grades: this supplementary information allows the variable of period A_{it} to be suppressed), along with level effects (l_i is a dummy of 1 if the pupils are in 4th and 5th grades). Model (3) takes into account school effects (u_i is a set of dummies for each school: likewise this supplementary information leads to the deletion of the treatment variable T_{it}). Model (4) puts the individual characteristics into the regression (X_{it} contains the following variables: sex, progression in school years, lagging in school years, languages spoken at home, having an own room). Finally, Model (5) differentiates the impact of the program according to the level with some cross variables.

On the basis of the estimated impact, we try to find which groups of pupils have obtained the most benefits from this program with the help of some cross variables. Model (6) is thus an extension of the Model (4):

$$Y_{it} = \beta_0 + \sum_{k=1}^{n} \beta_3^k A_{it} T_{it} G_k + \gamma \nu_t + \lambda l_i + \tau u_i + \alpha X_{it} + \epsilon_{it} \quad (6)$$

where G_n is a dichotomous variable with k modalities (e.g. sex, languages, levels, etc.) and β_3^n gives the estimated effect for each type of pupil.

2.4. The cost-effectiveness method

Cost-effectiveness analysis is an evaluation tool used to examine different alternatives in which costs and efficiencies are taken into account, and to determine which alternatives are the most appropriate with respect to the goals of a project. This methodology is little used in the field of education (Behrman, 1996; Levin, 2001). As we are able to rely on results from two French studies (Piketty, 2004; Piketty & Valdenaire, 2006), which estimate the class-size effect on marks scored in the French national evaluation scheme, we design our evaluation so as to obtain results that permit a cost-effectiveness comparison between the *Action Lecture* program and a class-size reduction program. Class-size reduction is one of the most discussed educational programs. Many empirical studies find that diminishing class size leads to an increase pupils' results (Akerhielm, 1995; Angrist & Lavy, 1999). The methodology used by Piketty and Valdenaire (2006) is similar to Angrist and Lavy and their results are robust and pertinent. They used data from a French panel – the 1997 primary panel – which follows a national sample of around 9600 pupils who started their 1st grade in 1997. Their main result is that each additional pupil in a 2nd grade class leads to a 0.339 point fall in the evaluation rating for reading, at the beginning of 3rd grade. These evaluations are based on a score of 100 points and the three skills studied in the *Action Lecture* represent 10% of the overall score. Therefore, the impact on skills that we measure with a score out of 10 is directly comparable to this class-size effect. The costs of the *Action Lecture* program stems from the teaching jobs it requires. If the teachers who work in this program were reallocated to classroom teaching, this would permit the opening of new classes and a reduction of class sizes in general. Furthermore, we can compute a cost-effectiveness ratio respectively for the *Action Lecture* program and for a class-size reduction program, because all the measurement units are marks per teaching job.

This comparison is only possible under the assumption that the results of Piketty and Valdenaire based on 2nd grade are also valid for the other levels: 3rd, 4th and 5th grades. Two reasons justify this hypothesis: first Piketty and Valdenaire also estimate the class size effect for 6th to 9th grades and find a value of 0.2, which is not too different from the class-size effect for 2nd grade (0.339 points).[3] Furthermore the observed standard errors for the results in 3rd, 6th and 9th grades are quite similar with values between 15 and 20 and our results have standard errors between 1.8 and 2, similar to the previous standard errors if we take into account the factor 10 in the scores' gap. Reading marks in national evaluation are relatively homogeneous for all grades.

3. Overview

We will first present the main individual characteristics of our sample and pupils' initial results in terms of academic results. Then we will control the quality of our control group.

3.1. Descriptive statistics

Table 2 shows the individual characteristics of the pupils. The first thing to note is that the schools present an important degree of social heterogeneity. Concerning the language spoken at home, only 64% use only French and the three main other languages are African languages, Arabic and Asian languages. The percentage of socially privileged schools is equally distributed across the participating schools; for the control group we have a bias of underprivileged pupils, but the effect should be

[3] A lower class-size effect in higher grades is to be expected. By using the 0.339 point estimation, we take a conservative and unfavorable point of view about *Action Lecture* program.

Please cite this article in press as: Massoni, S., & Vergnaud, J.-C. How to improve pupils' literacy? A cost-effectiveness analysis of a French educational project. *Economics of Education Review* (2011), doi:10.1016/j.econedurev.2011.08.013

Table 2
Individual characteristics.

	Participating schools							Control schools			
Localization in Paris[a]	Tot.	10	11	13	14	18	19	Tot.	2	13	20
Number of pupils	477	54	103	97	78	24	121	75	27	21	27
Level (a = 2nd, . . .)		a, d	a, d	a, b, c, d	b, c, d	a	b, c, d		d	b	c
Privileged (%)	50	41	46	74	61	14	40	37	57	30	23
Own bedroom (%)	44	40	41	51	40	36	42	44	64	35	44
French spoken (%)	64	50	62	72	77	41	61	56	63	57	48
African language (%)	8	19	1	1	3	30	12	8	4	5	15
Arabic language (%)	14	17	12	9	12	26	19	12	8	0	26
Asian language (%)	6	10	12	1	3	9	4	19	4	52	7
Backward (%)	10	27	7	6	3	30	10	11	4	10	19
In advance (%)	4	4	4	2	3	4	4	4	4	5	4

[a] Districts 10, 18, 19 20 are the least privileged in Paris, while Districts 2 and 14 are more wealthy.

Table 3
Initial results depending on individual characteristics.

	Total (S.E.)	Girl	2nd–3rd	Own room	Other languages			Lagging
					Afric.	Arab.	Asiat.	
Reading[a]	6.0 (2.3)	6.1	4.8	6.4	5.0	5.5	5.7	4.5
SATR	7.7 (1.7)	8.1	7.7	7.8	7.3	7.5	7.4	7.2
SCRC	6.9 (2.4)	7.1	6.8	7.1	6.0	6.5	6.2	5.3

[a] Reading refers to the aggregate score of the three exercises.

Table 4
Mean of the initial reading results according to the level of the attitude toward reading (SATR) and self-evaluation (SCRC) (% corresponds to the share of each level).

	SATR			SCRC		
Level[a]	High	Middle	Low	High	Middle	Low
Reading (%)	6.74 (61%)	5.85 (38%)	5.36 (1%)	6.95 (62%)	5.72 (35%)	5.04 (3%)

[a] We follow the PIRLS' classification. Compared to the results of the French sample in PIRLS, we observe higher SATR and SCRC.

compensated by the importance of the part of the Chinese community which is known to have good academic results.

In Table 3 we give some statistics concerning the three indicators' initial results (all noted out of 10), depending on different individual characteristics.

The reading results are quite as expected: better results for girls, and worse results for lagging pupils and for pupils of immigrant origin (except for pupils from the Chinese community). In Table 4we report the initial reading results according to the level of the attitude toward reading (SATR), and self-evaluation (SCRC).

These results were also expected.

3.2. The quality of the control group

As the procedure of selection of the control group is not optimal, we check if the two groups are not too different in terms of initial results. We first carry out a simple OLS regression of the following model (a):

$$Y_{i1} = \beta'_0 + \beta'_1 T_{i1} + \alpha' o_i + \epsilon_{it}$$

then we introduce the effect of time and academic levels in Model (b) and the individual characteristics in Model (c). If the coefficient (β'_1) of the participation variable is not significantly different to 0, we can consider that the control schools are similar to the schools participating. Table 5 reports the coefficients of the different regressions with the standard errors and the R^2.

For all specifications of the model and for the three variables of interest (aggregated results to exercises; SATR and SCRC scores) we find no significant differences between the control schools and the participating schools in terms of initial levels. Furthermore, the composition of the two groups is similar in terms of individual characteristics. In tables not reported here we have no significant differences in terms of percentage between the two groups for gender, academic progress, lag in school years, to speak only French

Table 5
OLS regression for the reliability of the control group.

	(a)	(b)	(c)
Reading	−0.241	+0.074	+0.006
(S.E.)	(0.171)	(0.159)	(0.158)
R^2	0.1317	0.3153	0.4468
SATR	−0.030	−0.036	+0.017
(S.E.)	(0.046)	(0.045)	(0.048)
R^2	0.0031	0.0061	0.0643
SCRC	−0.039	−0.035	+0.024
(S.E.)	(0.050)	(0.049)	(0.053)
R^2	0.0016	0.0040	0.0521
Control variables			
Time	No	Yes	Yes
Level	No	Yes	Yes
Individual characteristics	No	No	Yes

Please cite this article in press as: Massoni, S., & Vergnaud, J.-C. How to improve pupils' literacy? A cost-effectiveness analysis of a French educational project. *Economics of Education Review* (2011), doi:10.1016/j.econedurev.2011.08.013

Table 6
Estimation with DID of the program's impact on academic abilities.

	(1)	(2)	(3)	(4)
To identify the nature or type of a text	+0.117	+0.133	+0.316**	+0.353**
(S.E.)[a]	(0.157)	(0.144)	(0.177)	(0.153)
R^2	0.1279	0.2838	0.3417	0.3928
To process information	+0.061	+0.019	+0.006	−0.053
(S.E.)	(0.191)	(0.188)	(0.203)	(0.210)
R^2	0.0107	0.0540	0.0852	0.1222
To make inferences	+0.642***	+0.567***	+0.445**	+0.457**
(S.E.)	(0.184)	(0.168)	(0.177)	(0.180)
R^2	0.1531	0.3293	0.3598	0.4149
Overall results	+0.316**	+0.322**	+0.369***	+0.383***
(S.E.)	(0.139)	(0.126)	(0.131)	(0.128)
R^2	0.1348	0.3199	0.3711	0.4547
Control variables				
Time	No	Yes	Yes	Yes
Level	No	Yes	Yes	Yes
School	No	No	Yes	Yes
Individual characteristics	No	No	No	Yes

** Significant at 5%, *** Significant at 1%.
[a] All standard errors have been clustered at the school level for this table and the following ones.

at home, and the fact of having his/her own room. Another crucial point is to see if control and treatment schools have not been affected by specific reforms during the evaluation time. We run some interviews in each school to control for potential shocks during the time of the evaluation and we found no significant changes (no head teacher change, no teacher substitution for long term sick leave, etc.). We also find that there was no specific reasons that would have induced the treated schools to apply to the Action Lecture: the staff is stable and running project is part of the normal life of these schools. In the control group, we observed the same staff stability before, during and after the evaluation.

We can conclude that the treatment and the control groups have mainly the same initial trends and that our estimation by Differences-in-Difference will not be affected by a selection bias.

4. The results

We first show the project's impact, estimated by the method of Differences-in-Differences (DID). Then we present the cost-effectiveness analysis with a comparison of this project and a policy of decreasing class size.

4.1. The impact on reading skills

In Table 6 we detail the impact estimation for the three academic reading skills (with marks also out of 10) and the aggregate reading score, according to the DID method and for the four models that we have specified previously:

Models (3) and (4) are the most robust, and we observe a statistically significant impact for two skills out of three: *To identify the nature or type of a text* and *To make inferences*. For all models, the positive impact is significant for the aggregate result. We can also observe that the coefficients seem to not be very sensitive to different modelizations. For the two skills which presented a significant positive impact we test Model (5), in order to differentiate by levels the impact of the project.

Table 7 shows that, even if the 4th and 5th grades seem to make greater progress in terms of the first skill and inversely for the second skill, the overall difference between levels is not significant. Likewise, it is not pertinent to distinguish the effect of time, according to level, as we found no statistically significant differences.

It is known that some individual characteristics affect the pupils' abilities (see the impact of gender on reading abilities in Brown (1991)); thus it would be interesting to see if this project is more beneficial for some types of pupils. As most of the subgroups that we can define according to individual characteristics are too small, we focus only on two aspects: the gender and the language spoken at home. To estimate this we introduce some cross variables in the Model (4) concerning individual characteristics and the variable estimating impact. The results presented in Table 8 show that there are no significant differences.

To conclude, we can note that the impact of the program is large. Indeed, we may compare it to the average estimated progress of the reading score given by the coefficients' value of the time variable in the OLS regression of

Table 7
Estimation of the program impact and the effect of time on academic abilities, according to the level.

	Class 2nd–3rd	Class 4th–5th	Test of inequality	Time 2nd–3rd	Time 4th–5th	Test of inequality
Nature of a text	+0.262	+0.416*	NS	+0.112***	0.053**	NS
(S.E.)	(0.198)	(0.225)		(0.026)	(0.021)	
Inferences	+0.612***	+0.327	NS	+0.092***	0.070***	NS
(S.E.)	(0.245)	(0.266)		(0.023)	(0.017)	

* Significant at 10%, ** Significant at 5%, *** Significant at 1%.

Please cite this article in press as: Massoni, S., & Vergnaud, J.-C. How to improve pupils' literacy? A cost-effectiveness analysis of a French educational project. *Economics of Education Review* (2011), doi:10.1016/j.econedurev.2011.08.013

Table 8
Estimation of the program's impact on academic abilities according to some individual characteristics.

	Coeff.	(S.E.)	Test of inequality
Gender			
Girls	+0.347	(0.181)	NS
Boys	+0.418	(0.171)	
Languages			
African/Arabic	+0.401	(0.283)	NS
Others	+0.378***	(0.140)	

*** Significant at 1%.

Model (4). This mean progress is equal to +0.949 and thus the *Action Lecture* represents 40% of the annual increase in these three skills.

4.2. The impacts on Student's Attitude Toward Reading (SATR) and Student's Reading Self-Concept (SCRC)

These impacts are also positive for the two scores that capture the pupils' reading attitude. We do not detail the results here for all the models, but just report the DID's results for Model (4) in Table 9.

The *Action Lecture* program's effect is equivalent to 21% of the variance of initial score for SATR and 13% for SCRC (see Table 3).

4.3. Cost-effectiveness analysis

The previous results show that the project *Action Lecture* has a positive impact on academic results and reading attitudes. But it may be asked if it is enough to justify that resources should be used for it. Reassigning the teachers involved in the *Action Lecture* program in classrooms could be an alternative option that permits new classes to be opened and so reductions in class size. On the one hand, the *Action Lecture* program increases reading marks by +0.383 per pupil. On the other hand, a one-pupil reduction of the class size increases reading marks by 0.339 points, according to Piketty and Valdenaire (2006). To make the comparison effective, we estimate the potential class-size reduction generated by reassigning *Action Lecture* teachers to classrooms. We can compare this impact with the effect of the assigning teachers working in the program to new classes. We compute this effect in the following way:

- in the 6 schools we considered, a total of 36 classes followed the *Action Lecture* program,
- for 5 of these projects, 2 teachers participated and for the last project there was only 1 teacher out of a total of 11 teachers interventions,
- each teacher intervenes 9 times per year,
- the cost in terms of a teaching position is equal to $11/(9 \times 36)$ per class,
- the mean class size in our sample was 24.8 pupils,
- with a reassignment of the teachers, the mean class size in these classes would fall to $24.8/(1 + 11/(9 \times 36)) \approx 24$, i.e. a decrease of 0.8 pupils per class.

Table 9
Estimation with DID of the program's impact on the attitude toward reading (SATR) and self-evaluation (SCRC).

	(4)
SATR	+0.362***
(S.E.)	(0.118)
R^2	0.1251
SCRC	+0.311**
(S.E.)	(0.178)
R^2	0.1156

** Significant at 5%, *** Significant at 1%.

In Table 10 we summarize the different impacts estimated:

With the reassignment of the teachers to different classes, the impact will only be of $0.8 \times 0.339 = 0.271$ points at best, out of an overall 100 points, in terms of the results found in the reading evaluation. The impact of the *Action Lecture* program is estimated at 0.383 points out of 10% for the reading evaluation. Thus, this cost-effectiveness analysis does favor *Action Lecture*. The cost-effectiveness ratio expressed in terms of reading marks per teacher job can be computed.

- The cost-effectiveness ratio for a one point class-size reduction in a class that has X pupils is equal to $(X \times 0.339)/(1/X)$, since the total effect will be $X \times 0.339$ in terms of reading marks and $1/X$ in terms of teacher cost. For $X = 24.8$, the cost-effectiveness ratio for a class size reduction program is 208.5 points per teaching position.
- The cost for each class involved in the *Action Lecture* program is equal to $11/(9 \times 36)$ job positions and it produces 24.8×0.383 points in reading marks. Thus the cost-effectiveness ratio for an *Action Lecture* program is $(24.8 \times 0.383)/(11/(9 \times 36)) \approx 280$ points per teaching position.

It should be noted that the cost-effectiveness ratio for a class size reduction program is a quadratic function of the initial class size and it will equalize the *Action Lecture* cost-effectiveness ratio for a mean class size of 28.8.

This result holds under the assumption that the *Action Lecture* program does not have any negative impacts on other academic skills (mathematics, etc.). We may hope that even if other subject matters were dropped for 2 weeks, there are no negative effects on other skills. Indirect benefits may come from the progress in pupils' motivation observed for SATR and SCRC.

One limit of this cost-effectiveness analysis is that we use staff time instead of full costs. Since in the French education system, wages depends mainly on teachers' seniority, it would have been misleading to use the real wage costs. It was not possible to gather data on administrative costs but we guess that the administrative costs linked to a teacher position does not differ a lot whether the teacher is involved in the Action Lecture program or is

Table 10
Impact over evaluation results.

Effect of teachers reassignment	*Action Lecture* impact
+0.271	+0.383

Please cite this article in press as: Massoni, S., & Vergnaud, J.-C. How to improve pupils' literacy? A cost-effectiveness analysis of a French educational project. *Economics of Education Review* (2011), doi:10.1016/j.econedurev.2011.08.013

in a classroom. If this difference exists, it is certainly small compared to a teacher cost.

5. Discussion

In this study, we conclude that the impacts of the *Action Lecture* program were positive and that it is more efficient than a class-size reduction program. One obvious question that this study raises concerns the scope of such results and their robustness. First, our cost-effectiveness analysis is expressed in terms of the marks used in France's national evaluation scheme. The results may be different according to other ways of measuring the achievement of academic standards. Indeed, compared to international evaluations such as PIRLS or PISA, French national evaluation proceeds quite differently: the definition of skills is broken down more and the skills tested are less sophisticated. It is therefore reasonable to think that the *Action Lecture* program will perform even better according to international forms of evaluation. A second reason that may limit the scope of our results is the particular class-size impact estimation we use. Wößmann and West (2006) show that there is an important heterogeneity in the level of this impact regarding school systems (their results are based on the equivalent of PIRLS for mathematical skills: i.e. TIMMS – *Trends in International Mathematics and Science Study*). In France, a common view is that class-size reductions are ineffective unless teachers also change their teaching methods. But as the results by Piketty and Valdenaire (2006) show, this is not true. Nevertheless, their results are probably peculiar to the French education system. Hence, our results only make sense for education systems similar to France's one. One common criticism made of French teaching methods is that they are too directive and do not promote pupils' self-development. This may explain the observed efficiency of an innovative teaching program. In other countries, the *Action Lecture* program may appear to be less innovative and more similar to usual teaching methods.

This discussion shows that the economics of education is far from being able to provide cost-effectiveness ratios of a large scope, such as found in health economics. Indeed, education practices are far from being standardized, as is the case for medicine. It is also hard to find a common indicator of efficiency, such as *Quality Adjusted Life Years*, for instance. In this evaluation, we consider several indicators, and the fact that the impacts are positive for various indicators shows that it was not by chance that the impact was positive for the main indicator we chose.

Acknowledgements

The authors are grateful to Robert Caron, Fabian Gouret, Marion Hainaut, Victor Lavy, Thierry Rousse, Antoine Terracol and an anonymous referee for useful comments.

References

Akerhielm, K. (1995). Does class size matter? *Economics of Education Review*, *14*(3), 229–241.
Angrist, J. D. & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, *114*(2), 533–575.
Ashenfelter, O. & Card, D. (1985). Using the longitudinal structure of earnings to estimate the effect of training programs. *The Review of Economics and Statistics*, *67*(4), 648–660.
Behrman, J. R. (1996). Measuring the effectiveness of schooling policies in developing countries: Revisiting issues of methodology. *Economics of Education Review*, *15*(4), 345–364.
Bertrand, M., Duflo, E. & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, *119*(1), 249–275.
Brown, B. W. (1991). How gender and socioeconomic status affect reading and mathematics achievement. *Economics of Education Review*, *10*(4), 343–357.
Cameron, A. C. & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. Cambridge University Press.
Duflo, E., Kremer, M. & Glennerster, R. (2008). Using randomization in development economics research: A toolkit. In J. Strauss, & P. Schultz (Eds.), *Handbook of development economics, vol. 4* (pp. 3895–3962). Elsevier-North Holland.
Imbens, G. W. & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*, 5–86.
Levin, H. M. (1995). Cost-effectiveness analysis. In M. Carnoy (Ed.), *International encyclopedia of economics of education* (pp. 381–386). Oxford: Pergamon.
Levin, H. M. (2001). Waiting for Godot: Cost-effectiveness analysis in education. *New Direction for Evaluation*, *90*, 55–68.
Mullis, I. V. S., Martin, M. O., Gonzalez, E. J. & Kennedy, A. M. (2003). *PIRLS 2001 international report: IEAs study of reading literacy achievement in primary schools*. Chestnut Hill, MA: Boston College.
Mullis, I. V. S., Martin, M. O., Kennedy, A. M. & Foy, P. (2007). *PIRLS 2006 international report: IEAs study of reading literacy achievement in primary schools*. Chestnut Hill, MA: Boston College.
Piketty, T. (2004). L'impact de la taille des classes et de la ségrégation sociale sur la réussite scolaire dans les écoles françaises: Une estimation à partir du panel primaire 1997. *Working Paper. Presentation in English: "Should We Reduce Class Size or School Segregation? Theory and Evidence from France"*, http://www.jourdan.ens.fr/piketty/fichiers/public/PikettySlides.pdf.
Piketty, T., Valdenaire, M. (2006). *L'impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français: Estimation à partir du panel primaire 1997 et du panel secondaire 1995* (pp. 1–153). Paris: Ministère de l'éducation nationale. Les Dossiers n°173.
Reuter, Y. (2007). *Une école Freinet: Fonctionnements et effets d'une pédagogie alternative en milieu populaire*. L'Harmattan.
Wößmann, L. & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review*, *50*(3), 695–736.

Please cite this article in press as: Massoni, S., & Vergnaud, J.-C. How to improve pupils' literacy? A cost-effectiveness analysis of a French educational project. *Economics of Education Review* (2011), doi:10.1016/j.econedurev.2011.08.013