Swedish Institute for Social Research (SOFI)

Stockholm University

WORKING PAPER 2/2004

THE EFFECT OF EXTRA FUNDING FOR DISADVANTAGED PUPILS ON ACHIEVEMENT

by

Edwin Leuven, Mikael Lindahl, Hessel Oosterbeek, Dinand Webbink

The effect of extra funding for disadvantaged pupils on achievement[1]

Edwin Leuven[2] Mikael Lindahl[3] Hessel Oosterbeek[4]
Dinand Webbink[5]

[1]We would like to thank the referees for very useful comments. We also thank seminar participants in Amsterdam, Dublin, Sevilla, Stockholm and Uppsala. We thank the Dutch ministry of Education and CITO for supplying the data used in this paper.

[2]Department of Economics, University of Amsterdam, NWO Priority Program Scholar and Tinbergen Institute.

[3]Swedish Institute for Social Research (SOFI), Stockholm University.

[4]Department of Economics, University of Amsterdam, NWO Priority Program Scholar and Tinbergen Institute.

[5]NWO Priority Program Scholar and CPB Netherlands Bureau for Economic Policy Analysis, The Hague.

Abstract

This paper evaluates the effects of two subsidies targeted at disadvantaged pupils in the Netherlands. The first scheme gives primary schools with at least 70 percent minority pupils extra funding for personnel. The second scheme gives primary schools with at least 70 percent pupils from different disadvantaged groups extra funding for computers and software. The cutoffs at 70 percent provide a regression discontinuity design which we exploit in a local difference-in-differences framework. For both subsidies we find negative point estimates. For the personnel subsidy these are in most cases not significantly different from zero. For the computer subsidy we find more evidence of negative effects. We discuss several explanations for these counterintuitive results.
JEL Codes: I21, I28, J24

Keywords: policy evaluation, disadvantaged students, computers, teachers, regression discontinuity

1 Introduction

An ongoing discusion in the economics of education concerns the effect of extra resources on students' achievement. The conflicting views are clearly demonstrated in recent contributions by Krueger (2003) and Hanushek (2002), that debate the effectiveness of class-size reduction as a means to improve student achievement. Hanushek concludes that there is little support for the effectiveness of class-size reduction since the studies that find positive effects do not outnumber the studies that find no such effect. Krueger argues that "there is no substitute for understanding the specifications underlying the literature and conducting well designed experiments", and that better studies should be given more weight in Hanushek's counting exercises.

The class-size debate illustrates the increasing importance that is attached to the proper design of evaluations, and the concomitant emphasis on results based on (quasi-) experimental research. The present study reports the results of two quasi-experiments applicable to primary schools with a large share of disadvantaged pupils. The first quasi-experiment provided an extra payment per teacher of about 10 percent of gross salaries during two consecutive years. Only schools where at least 70 percent of the pupils have an ethnic minority background were eligible for this subsidy. Schools were free to spend the personnel subsidy as they saw fit, as long as it improved working conditions. They could use it (among other things) to hire extra teachers or to give teachers an extra payment. The relative freedom schools had in spending the extra money reflects a current trend in the Netherlands (and elsewhere) towards more decentralized allocation of public spending.

The second quasi-experiment provided a one-time payment of $90 per pupil. This money was earmarked for computers, software and language materials. Only schools where at least 70 percent of the pupils have a disadvantaged background (ethnic minority or low educated parents) were eligible for this subsidy.

For both interventions the 70 percent threshold was maintained almost perfectly thereby creating a regression discontinuity design. The only assumption needed to be fulfilled for this design to produce unbiased estimates of the effect of a program, is that there are no confounding discontinuities at the threshold. We exploit the regression discontinuities in a local difference-in-differences framework to identify the effect of the two programs on pupils' achievement. To this end we combine administrative data with data on the achievement of 8th graders in nationwide exams.

The policy background of both schemes is that, despite a rather generous compensatory funding scheme, ethnic minority pupils and pupils with low educated parents fare worse in school than their non-disadvantaged counterparts. The main funding scheme for primary schools provides schools 25 percent extra funding for pupils with low educated parents and 90 percent extra funding for pupils from an ethnic minority. A school with all of its pupils from an ethnic minority receives therefore almost twice as much funding as a school with all its pupils being non-disadvantaged. In the total population of primary school pupils 18 percent have low educated parents and 13 percent have an ethnic minority background.[1] In 2000 the total amount spent on this compensatory program was $234 million for 450,000 disadvantaged pupils. The two subsidies that we evaluate in this paper were motivated by the belief that the compensation from the main scheme is insufficient, especially for schools with a large share of disadvantaged pupils.[2]

[1] In addition to these two groups, the funding scheme also distinguishes students living in a boarding school or a foster home and whose parents are master of a ship, and students whose parents are transients. Schools receive 40 and 70 percent extra funding for such students respectively. The shares of these groups in the population are, however, negligible.

[2] A large number of studies have been conducted to evaluate the effects of the compensatory element of the main funding scheme, many of these commissioned by the Dutch government. These studies do not allow to relate changes in the achievement levels of disadvantaged students to the funding scheme. The reason is that the funding scheme treats all students with the same social background equally. As a result there is no natural control group, nor is there a possibility to construct a suitable comparison group.

For both subsidies we find negative point estimates. For the personnel subsidy these are, however, in most cases not significantly different from zero. For the computer subsidy we find more evidence of negative effects. These outcomes indicate that neither subsidy had a substantial positive effect on pupil achievement. Since the costs of these subsidies were substantial, both schemes therefore perform rather poor in terms of cost-effectiveness. Notice that in both cases the treatments that generate the non-positive effects are well-defined. In the case of the personnel subsidy the treatment is to provide schools with a specific amount of extra funding per teacher to improve working conditions. In the case of the computer subsidy the treatment is to give schools a specific amount of extra funding per pupil for computers and software. We will also present evidence that schools spent the extra money as intended.

An explanation for the poor performance of the personnel subsidy is that due to the main funding scheme targeted schools already have sufficient resources for personnel. Perhaps schools have difficulty spending the extra money in an effective way. The pupil-teacher ratio in these schools is below 14. Although schools spend about half of the extra budget on hiring new personnel, it is unlikely that hiring a new teacher will result in a reduction of average class-size. The other half of the subsidy is spent on improving teachers' remuneration and/or fringe benefits. Since this is a non-permanent increase in teachers' salaries that is not connected to an incentive scheme it may perhaps fail to improve teacher effort or attract better teachers to these schools.

The non-positive effects of the computer subsidy concur with findings of other recent studies relating to other countries, other levels of education and/or other identification methods. The robustness of this result suggests that computer aided instruction may after all be an inferior mode of teaching.

The remainder of this paper is organized as follows. The next section reviews recent studies that look at comparable interventions. Section 3 provides details of the two programs and describes the data. Section 4 outlines the estimation strategy. Section 5 presents and discusses the empirical findings. The final section summarizes and concludes.

2 Review of related studies

2.1 Educational resources and pupils' achievement

For about a decade the consensus amongst economists was that extra resources have no strong or systematic impact on pupils' achievement. Especially the survey articles by Hanushek (1986, 1994, 1996) in which he reviews over 300 empirical studies, have been important in this respect. Most of the studies that Hanushek reviewed, however, use identification methods that by current standards would qualify as inadequate and unconvincing. Recent studies, that arguably use more convincing identification strategies, also find mixed results.

Guryan (2000) uses features of an education finance equalization scheme in Massachusetts to estimate the effect of increased spending on pupils' achievement at schools that are located in historically low-spending districts. For 4th graders (but not for 8th graders) he finds improved test scores, especially for low-scoring students. Papke (2003) exploits a similar equalization scheme in Michigan and uses panel data to identify the effect on 4th grade pass rates and 7th grade math tests. She finds that increases in spending have substantial effects on the math test pass rate. Here, effects are largest for schools with initially poor performance. Card and Payne (2002) analyze the effects of school finance reforms on the distribution of school spending across richer and poorer districts. Unlike the previous two papers, they analyze nationwide data. Card and Payne find that equalization of spending narrows the difference in test score outcomes across family background groups.

A recent study that fails to find an impact of extra resources on achievement of (disadvantaged) pupils is Van der Klaauw (2003) who investigates how Title I affects student achievement. Title I provides financial support for supplementary educational services in mathematics and reading to poor and low achieving students. Van der Klaauw evaluates the effects of Title I in a regression discontinuity framework using data on New York City public schools. He finds that Title I has not proven to be successful in improving student outcomes. He discusses possible explanations for this finding. First there is some evidence that cities and states substitute regular funding away from Title I schools, resulting in a limited increase in total spending in these schools. Another explanation lies in the fact that in practice remedial classes are relatively ineffective because they are often taught by inexperienced teacher aides.

Another study that fails to find effects of extra funding directed at schools with disadvantaged students is Bénabou et al. (2004). For France they investigate the effect of a compensatory funding scheme in a difference-in-differences framework. The extra funding was partly aimed at improving the pay of teachers, and partly at increasing classroom hours of pupils. They do not find evidence that these extra resources improved student achievement as measured by test scores.

The studies mentioned above all estimate the effect of extra resources on achievement thereby treating spending as a ”black box”. But where most of these studies do not have information on the exact size of the extra resources, we do know how much each school received thereby making the treatment very precisely defined.

2.2 Computers and pupils' achievement

The evidence on the effect of computers in schools on pupils' achievement is limited. In their review Kirkpatrick and Cuban (1998) conclude that the effect of computer use on achievement is questionable. Although many of the reviewed studies report positive outcomes, Kirkpatrick and Cuban conclude that the value of this research is limited because it does not take endogeneity issues into account.

Three recent studies that do address endogeneity find zero or negative effects of extra computers or software on achievement. Angrist and Lavy (2002) evaluate the effects of a program in which the Israeli State Lottery funded new computers in elementary and middle schools in Israel. They use several estimation strategies (OLS and 2SLS) and find “a consistently negative and marginally significant relationship between the program-induced use of computers and 4th grade Maths scores” (p.760). For 8th graders and for scores on Hebrew, the estimated effects are mostly negative although not significantly different from zero.

Goolsbee and Guryan (2002) report the results from a program that subsidized schools’ investment in Internet and communications. The subsidy has a substantial positive impact on the probability of classrooms having an Internet connection. At the same time this increase in Internet connections has had no measurable impact on any measure of pupil achievement.

Finally, Rouse et al. (2004) study the effects of a instructional computer program called Fast ForWord (FFW). The authors find no evidence that the use of FFW results into gains in language acquisition or actual reading skills. Interestingly the time students spent using FFW was in addition to the amount of time they spent in regular reading instruction. Although Rouse et al. do not find negative effects, broader use of computers in instruction is likely to substitute regular instruction. If computer based learning is less effective than more traditional forms of classroom teaching, negative effects cannot be ruled out.

3 Programs and data

3.1 The two programs

In February 2000 the Dutch ministry of Education announced a personnel subsidy for schools with at least 70 percent minority pupils.[3] Eligibility was based on the percentage of minority pupils of a school on October 1 1998 as counted in administrative data. The extra funding amounted to \$2,225 *per teacher* in the school year 1999-2000 and \$2,440 in 2000-2001. These sums were paid in May 2000 and March 2001. In November 2000 it turned out that the available budget for the year 2000 was not exhausted, and in December 2000 the eligible schools received an additional \$585 per teacher. The total payment equaled therefore \$2,625 per teacher per year over a two-year period. This annual amount is roughly equal to 9 percent of the average annual gross salary of Dutch primary school teachers, and 11 percent of the annual gross salary of young teachers. This is a substantial intervention, given that personnel costs are roughly 80 percent of schools' total budget.

Schools were free to spend the budget in ways that matched the schools' needs, as long as they were aiming to improve working conditions. The explanatory memorandum that was circulated following the Ministry's decision listed as examples: a plain financial premium, a bonus to stimulate teachers to work extra hours, compensations for housing costs, traveling costs or childcare facilities, and hiring teaching assistants. Although the memorandum was ambiguous about a possible continuation of the subsidy it emphasized that the extra funding was provided for a limited period and that obligations pertaining after this period had to be paid from the regular budget.

Later that year, in November 2000, the ministry announced another measure,

[3]The formal description of this group is students with parents born in Surinam, the Netherlands Antilles or non-English speaking countries outside Europe or whose parents are refugees, and whose father or mother has at most completed low level vocational education or whose primary earnings parent has a job involving physical labor or has no income from labor.

which stipulated that schools with at least 70 percent of their pupils belonging to any disadvantaged group (ethnic minority or low educated parents) would receive extra funding in the amount of $90 *per pupil*, which is about $1250 per classroom in the eligible schools.[4] For this scheme the percentage of disadvantaged pupils of a school was based on administrative data counted on October 1 1999. Of this sum, $15 were earmarked for renewal of language materials, while the remaining $75 were earmarked for computers or (education) software. The subsidy was paid only once in December 2000.

A common feature of these two interventions is that they specify a minimum percentage of disadvantaged pupils schools need to have to qualify for the extra compensation. The personnel subsidy requires at least 70 percent of ethnic minority pupils, the computer subsidy requires at least 70 percent pupils from any disadvantaged group. All treated schools received the same amount per teacher or per pupil.

3.2 Data construction

The ministry of education provided us with data on the numbers of pupils of different social backgrounds for all primary schools in the Netherlands counted at October 1, 1998 and October 1, 1999. The data also contain information about which schools actually received extra funding. These administrative data were merged with information about pupils' results in nationwide tests. The data also include information on the average social background of the school population ranging from 1 (least disadvantaged) to 7 (most disadvantaged), the degree of urbanization of the school area and the school's denomination.

More than 80 percent of primary schools participate in a nationwide testing

[4]Formally a pupil's parents are low educated if one parent has at most an education at the lowest level of vocational education.

round.[5] All pupils who are in the highest (8th) grade make a standardized test that covers four areas:

- Language: spelling, writing, reading and vocabulary;
- Arithmetic: understanding of numbers, mental arithmetic, percentages, fractions, dealing with measures, weights, money and time;
- Information processing: use of texts and other information sources, reading and understanding of tables, graphs and maps;
- World orientation (optional): applying knowledge in the fields of geography, history, biology, science and form of government.

Testing takes place during three days in February. The complete test consists of over 200 multiple-choice questions. Pupils' scores on this test are used for the assignment of pupils to different levels of secondary schools. Many secondary schools apply strict thresholds to admit pupils to the more advanced types of secondary education. This gives pupils an incentive to perform well on this test. Furthermore, the average scores of schools' pupils are currently used as information to judge the quality of primary schools. These average scores are public information and parents use it in their choice of primary school. This gives schools an incentive to prepare their pupils well for the test. To illustrate the importance of the test, every year all national newspapers as well as national television pay special attention to it. The impression often is, that preparing and making this test is the main activity of pupils in their last two years in primary school (7th and 8th grade).

For our analysis we use data of the test scores from pre-intervention years 1999 and 2000 and from post-intervention years 2002 and 2003.[6] In the empirical anal-

[5]For the samples we use in the analysis we do not find any statistically significant differences in test participation between schools above and below the thresholds.

[6]We do not use data from 2001 because it is unclear whether this is a pre- or a post-intervention year.

Table 1: Timing of events

October 1 1998	Reference date for personnel subsidy
February 1999	Nationwide test 1999
October 1 1999	Reference date for ICT subsidy
February 2000	Nationwide test 2000
February 2000	Decision and announcement personnel subsidy
May 2000	Payment of $2,225 per teacher as personnel subsidy
November 2000	Decision and announcement ICT subsidy
November 2000	Decision and announcement of extra payment personnel subsidy
December 2000	Payment of $90 per pupil as ICT subsidy
December 2000	Extra payment of $585 per teacher
March 2001	Payment of $2,440 per teacher as personnel subsidy
February 2002	Nationwide test 2002
February 2003	Nationwide test 2003

ysis, the scores of schools' pupils on the language, arithmetic and information processing parts serve as the outcome variables. To standardize the estimated effects, the scores are divided by their standard deviations and normalized to mean zero relative to the whole population.

Table 1 gives an overview of the timing of the relevant events. It is clear from the table that the tests of February 1999 and February 2000 took place before schools received extra funding. The 2003 (2002) test took place almost three (two) years after the first payment of the personnel subsidy, more than two (one) years after the extra payment of the personnel subsidy, and the payment of the computer subsidy, and almost two (one) years after the payment of the last tranche of the personnel subsidy. We use the 2002 and 2003 test scores as relevant outcome measures for both subsidies.

In 1998 there were 7,045 primary schools in the Netherlands. Of these schools, in total 270 (4%) had at least 70 percent of their pupils belonging to an ethnic minority group thereby qualifying for the personnel subsidy. Out of these 270 there were 267 schools that actually received the personnel subsidy.[7] Seven schools with less than 70 percent of their pupils belonging to the ethnic minority cate-

[7] The 3 schools not receiving the personnel subsidy had shares of ethnic minority students equal to 1, 0.84 and 0.73.

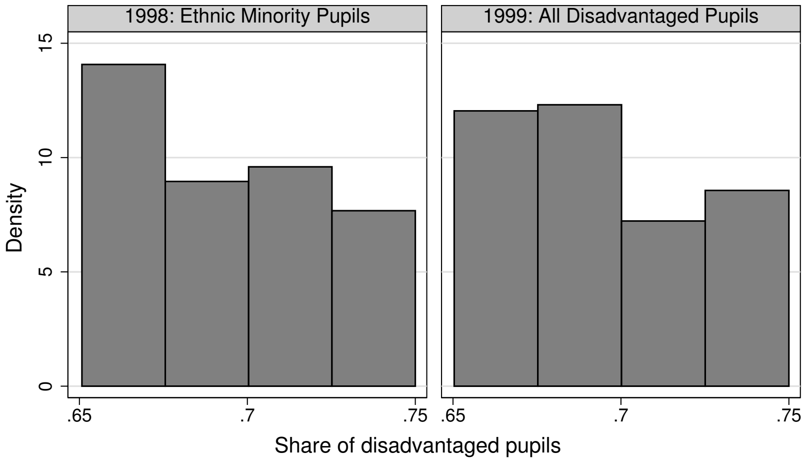

Figure 1: Distribution of schools

gory (mistakingly) received this subsidy.[8] Considering the computer subsidy there were 7,028 primary schools in 1999, of which 564 (8%) had at least 70 percent of their pupils belonging to any disadvantaged group in 1999. 551 of these 564 schools received the computer subsidy.[9] Sixteen schools with less than 70 percent of their pupils belonging to any disadvantaged group (mistakingly) received this subsidy.[10] We do not know the reasons for the misclassifications. Section 4 describes how schools that received the extra funding while not eligible (and vice versa), are treated in the analysis.

In our identification setup one might be concerned that schools anticipated the subsidies and accordingly manipulated their relevant shares of disadvantaged pupils to become eligible. This seems unlikely since they would have needed to

[8]The shares of ethnic minority students at these 7 schools are: 0.69, 0.69, 0.68, 0.67, 0.64, 0.58, and 0.34.

[9]The 13 schools not receiving the personnel subsidy had shares of disadvantaged students equal to: 0.71, 0.75, 0.76, 0.89, 0.92, 0.93, 0.93, 0.94, 0.95, 0.96, 0.97 and 1 (twice).

[10]The shares of disadvantaged students at these 16 schools are: 0.39, 0.43, 0.48, 0.56, 0.56, 0.57, 0.58, 0.61, 0.65, 0.67, 0.67, 0.67, 0.68, 0.69, 0.69 and 0.69.

anticipate the personnel subsidy by one-and-a-half year and the computer subsidy by one year. Nevertheless, one check of such manipulation is to compare the distribution of schools around the cutoff level. Manipulation would lead to a drop below the 70 percent cutoff and a rise just above. Figure 1 shows the frequency distributions of schools in the range of 10 percent around the cutoff levels of 70 percent. These distributions give no indication of such manipulation thereby confirming that schools did not anticipate the implementation of the two programs.

Schools that have at least 70 percent minority pupils in 1998 are also very likely to have at least 70 percent disadvantaged pupils in 1999. In other words, schools that qualify for the personnel subsidy are also very likely to qualify for the computer subsidy. In the empirical analysis we focus on schools with their shares of minority pupils or disadvantaged pupils at most 10 percentage points away from the 70 percent thresholds. Within these subsamples nearly all schools that qualify for the personnel subsidy also qualify for the computer subsidy and almost no school that qualifies for the computer subsidy qualifies for the personnel subsidy.

4 Empirical strategy

This section discusses the empirical strategy used to identify the effect of the two subsidies. The discussion is phrased in terms of the personnel subsidy. The approach for identification of the effect of the computer subsidy is identical. We first briefly describe the standard (sharp) regression discontinuity design, and then describe how this is exploited in the analysis.

4.1 Regression Discontinuity Design

The eligibility rule of the personnel subsidy specifies that all schools with at least 70 percent minority pupils receive the subsidy and all schools with less than 70 percent minority pupils do not receive the subsidy. Without exceptions to this rule we would have a so-called sharp regression discontinuity design in which treatment depends in a deterministic way on the share of minority pupils.[11]

To estimate the effect of the treatment we can compare the average outcome of the group just above the threshold with the average outcome of the group just below the threshold. This gives an unbiased estimate of the average treatment effect for schools with 70 percent of disadvantaged pupils if there are no confounding discontinuities at the threshold.

Denote the share of minority pupils in school j in 1998 by s_j^{98}. With a sharp regression discontinuity design the variable denoting treatment, d_j^{98}, is defined as follows

$$d_j^{98} = \begin{cases} 1 & \text{if } s_j^{98} \geq 0.7 \\ 0 & \text{if } s_j^{98} < 0.7 \end{cases}$$

The outcome can be written as

$$E[y_j] = \alpha + \delta d_j^{98}$$

where $\alpha \equiv E[y_{0j}]$ is the (average) test score without the subsidy, and $\delta \equiv E[y_{1j}] - E[y_{0j}]$ is the change in test scores due to the subsidy. Under the assumption of a common treatment effect, it can be shown that δ can be identified by (cf. Hahn et al., 2001):

$$\delta = y^+ - y^-$$

where $y^+ \equiv \lim_{s \downarrow 0.7} E[y|s]$ and $y^- \equiv \lim_{s \uparrow 0.7} E[y|s]$. The major identifying assumption is that there are no other discontinuities around 0.7. This is an exclusion restriction with respect to the discontinuity.

[11] Leuven and Oosterbeek (2004) provide a recent application of the sharp regression discontinuity design.

4.2 Estimation

Although we could exploit the regression discontinuity in the standard way, and compare schools around the discontinuity, we follow a difference-in-differences strategy to increase the power and obtain more precise estimates.[12] In terms of implementation we estimate fixed effect regressions of the following form

$$y_{ijt} = \alpha + \delta \cdot (D_{ij}^{98} \times m_t) + \gamma \cdot D_{ij}^{98} + \tau \cdot m_t + \eta_j + \phi_{jt} + \varepsilon_{ijt} \tag{1}$$

where y_{ijt} is the test score of pupil i in school j in year t, and D_{ij}^{98} is a zero-one indicator variable which equals one if school j received the subsidy, η_j is a school fixed effect, ϕ_{jt} is a school-cohort (class) random effect, m_t are time effects (dummies) and ε_{ijt} is an i.i.d. error term. Note that the estimate of δ in (1) recuperates the standard difference-in-differences estimate when restricting the sample to one post-intervention and one pre-intervention year. Below we estimate (1) on a sample of two pre-intervention and two post-intervention years.

In a standard regression discontinuity design one compares observations just below the cutoff to observations just above it. Although we calculate difference-in-differences estimates we will estimate them locally and exploit the discontinuity to add to the credibility of the common trend assumption that is necessary for difference-in-differences estimates.

For this purpose we construct so-called discontinuity samples. The x percent Discontinuity Sample (DS$\pm x$) consists of the eligible group of schools with their percentage of minority pupils at most x percent above the cutoff of 70 percent, and the non-eligible group of schools with their percentage of minority pupils at most x percent below the cutoff of 70 percent. Widening the bandwidths around the discontinuity increases the number of observations but at the same time increases

[12] Estimation in levels does not lead to different conclusions, but gives much less precise estimates.

the risk that the common trend assumption is violated. In the analysis, we will work with DS$\pm$5 and DS$\pm$10. These samples are relatively close to the discontinuity and include sufficient schools to obtain meaningful results.

An additional advantage of the difference-in-differences estimation is that it also addresses imperfect compliance. As mentioned in the previous section, a few schools did receive the personnel subsidy although they had less than 70 percent minority pupils. Because the rule behind these exceptions is unknown (at least to us), this breaks down the sharp regression discontinuity design. There is no longer a deterministic relation between treatment and the share of minority pupils.

Shadish et al. (2002) suggest to either retain the misclassified cases in the analysis and classify them according to their eligibility status rather than by the their treatment status, or to eliminate misassigned observations from the analysis. The first solution does not give an estimate of the treatment effect of interest while the second solution appears to be somewhat ad hoc and does not give an estimate of a well defined effect.

The advantage of the difference-in-differences strategy is that it eliminates the potential bias arising from imperfect compliance if this is captured by time invariant unobserved school effects. As a check, we also estimated 2SLS equations where the dependent variable is the change in test scores over time, while instrumenting treatment with eligibility. This gave very similar point estimates, and we could not reject equality with the difference-in-differences estimates. Since the latter procedure gives more precise estimates we will present these below.

Table 2: Sample means for population and estimation samples, 2002

		Personnel		Computer	
	Population	DS±5	DS±10	DS±5	DS±10
	(1)	(2)	(3)	(4)	(5)
Language	0.000	-0.538	-0.587	-0.428	-0.442
(s.d.)	(1.000)	(1.058)	(1.048)	(1.062)	(1.065)
Arithmetics	0.000	-0.343	-0.412	-0.318	-0.320
(s.d.)	(1.000)	(1.074)	(1.081)	(1.057)	(1.063)
Information	0.000	-0.549	-0.602	-0.441	-0.423
(s.d.)	(1.000)	(1.107)	(1.083)	(1.086)	(1.089)
Share minority 1998 (s^{98})	0.125	0.693	0.696	0.362	0.362
Share disadvantaged 1999 (s^{99})	0.294	0.835	0.852	0.695	0.688
Socio-economic index					
-1 (least disadvantaged)	0.102	0.000	0.000	0.000	0.000
-2	0.336	0.000	0.000	0.000	0.000
-3	0.311	0.000	0.000	0.024	0.017
-4	0.070	0.000	0.000	0.155	0.204
-5	0.093	0.010	0.008	0.458	0.421
-6	0.038	0.526	0.482	0.363	0.305
-7 (most disadvantaged)	0.048	0.464	0.509	0.000	0.052
Urbanization school area					
-Very High	0.150	0.738	0.696	0.520	0.422
-High	0.227	0.184	0.197	0.241	0.286
-Median	0.210	0.051	0.067	0.111	0.131
-Modest	0.249	0.023	0.031	0.094	0.100
-Low/None	0.164	0.004	0.009	0.034	0.060
School denomination					
-Public	0.316	0.568	0.512	0.364	0.460
-Catholic	0.355	0.206	0.252	0.343	0.303
-Protestant	0.267	0.167	0.195	0.263	0.203
-Montessori/Daltonian	0.053	0.023	0.022	0.031	0.026
-Other	0.010	0.036	0.019	0.000	0.008
Number of pupils	150821	1817	3392	3954	8263
Number of schools	5938	63	124	150	328

5 Results

5.1 Data description

Before turning to the main results of this paper this section first describes the data. Table 2 shows the sample means in 2002 for the estimation samples, and how they compare to the whole population of pupils. Since the effects we estimate are local effects it is important to know how these samples compare to the population as a whole.

As seen in the first three rows of column (1), we standardized the test scores to have mean zero and standard deviation one in the population. Compared to the average student, the pupils in the schools around the personnel discontinuity score on average more than half a standard deviation lower on both the language and information processing test. Performance in arithmetics is about one third of a standard deviation lower in these schools compared to the population average.

For the local samples around the computer eligibility discontinuity, test scores for language and information processing are more than 0.4 of a standard deviation below the population average. For arithmetics the difference is somewhat above 0.3 of a standard deviation. The fact that only the language and information processing scores and not the arithmetics scores of the personnel discontinuity samples are worse than those of the computer discontinuity samples, reflects that minority pupils do worse on language and information processing than Dutch disadvantaged students but not on arithmetics.

The schools that are (almost) eligible for the personnel subsidy are in the two most disadvantaged groups of the socio-economic classification index of the school population, whereas the schools in the computer subsidy sample have on average less disadvantaged students. To compare, in the whole population the vast majority of the schools have students from the three least disadvantaged categories.

Table 2 also shows that the more disadvantaged the student population the more likely the school is situated in one of the major cities. About 70% of the students in schools around the personnel discontinuity live in one of the major cities, compared to 50% for the computer subsidy sample, and only 15% of the total population.

Finally, minority pupils are more likely to attend public schools. The bottom panel in columns (2) and (3) shows that more than half of these pupils are in public schools compared to 32% in the population. In contrast, the denomination of the schools that find themselves around the qualifying discontinuity for the computer subsidy is quite similar to those in the population.[13]

To see how test scores vary with individual and school characteristics, table 3 presents the results from an OLS regression. Column (1) shows that girls score on average one fifth of a standard deviation higher on the language test, whereas on the arithmetics test boys do better (column 2). The difference between boys and girls is larger on the arithmetics test than on the language test. There are no gender differences on the information processing test.

As expected, socio-economic background strongly correlates with achievement. Pupils at schools with the most disadvantaged backgrounds score on average a full standard deviation lower than pupils at non-disadvantaged schools. This difference is smaller on the arithmetics test where the gap amounts to two-thirds of a standard deviation.

Interestingly, students in Catholic schools do better than students in public or protestant schools. This is a well known finding for the United States. However, table 3 shows that pupils in schools based on an educational principle such as Montessori, or Daltonian do even better, while students in the residual category score even a quarter of a standard deviation higher than students in public schools.

[13] We tested whether treatment and control schools were different with respect to denomination and urbanisation. We did this for all four of our estimation samples. Out of the 8 tests we only rejected equality for urbanisation in the DS$\pm$10 for the computer subsidy.

Table 3: Descriptive regression of 2002 test scores on student and school characteristics

	Language (1)	Arithmetics (2)	Information (3)
Girl	0.198	−0.334	−0.041
	(0.040)	(0.042)	(0.039)
Socio-economic index			
-1 (least disadvantaged)	reference	reference	reference
-2	−0.090	−0.066	−0.078
	(0.016)	(0.016)	(0.015)
-3	−0.236	−0.176	−0.224
	(0.016)	(0.017)	(0.016)
-4	−0.431	−0.340	−0.422
	(0.022)	(0.022)	(0.021)
-5	−0.508	−0.374	−0.492
	(0.021)	(0.021)	(0.020)
-6	−0.762	−0.564	−0.750
	(0.028)	(0.028)	(0.027)
-7 (most disadvantaged)	−0.998	−0.656	−1.000
	(0.027)	(0.028)	(0.026)
School denomination			
-Public	reference	reference	reference
-Catholic	0.077	0.092	0.072
	(0.011)	(0.011)	(0.011)
-Protestant	0.008	−0.012	−0.025
	(0.012)	(0.012)	(0.011)
-Montessori/Daltonian	0.145	0.098	0.123
	(0.021)	(0.022)	(0.020)
-Other	0.249	0.302	0.227
	(0.046)	(0.047)	(0.044)
Urbanization school area			
-Very high	0.056	0.013	−0.014
	(0.018)	(0.019)	(0.017)
-High	0.017	−0.001	−0.033
	(0.015)	(0.015)	(0.014)
-Median	0.026	−0.000	−0.011
	(0.015)	(0.015)	(0.014)
-Modest	0.007	0.000	−0.007
	(0.014)	(0.014)	(0.013)
-Low/None	reference	reference	reference
Constant	0.101	0.315	0.254
	(0.027)	(0.028)	(0.026)
R-squared	0.35	0.22	0.38
Number of pupils	150061	150061	150061
Number of schools	5896	5896	5896

Note: Standard errors (in parentheses) are heteroscedasticity robust and take into account clustering at the school level.

It should be noted that these numbers are correlations and that the regression only limitedly controls for parental background through the school population index. It seems therefore likely that these coefficients pick up unobserved background characteristics.

5.2 Effects of personnel subsidy

Table 4 reports the findings for the personnel subsidy on the three outcome variables language, arithmetic and information processing for the $\pm 5\%$ and $\pm 10\%$ samples around the discontinuity. We report the effects for the post intervention years 2002 and 2003 separately since, strictly speaking, these are different outcomes. We also report a pooled estimate for 2002 and 2003 which is more precise, and the statistic of the test for equality of the effects for the separate years.

First consider the results on the language test for the 5% discontinuity sample. All estimated effects are negative and of comparable size. Equality between 2002 and 2003 cannot be rejected and the pooled estimate of the effect of the personnel subsidy on language scores is -0.069. The effects for the subsample with a wider bandwidth around the discontinuity, DS$\pm$10, are very similar. They are slightly more negative in 2003 than in 2002. The pooled estimate is -0.055 with a standard error of 0.043. We can therefore rule out effects on language scores in excess of 3% of a standard deviation with a 95% probability.

For the arithmetics scores a very similar pattern emerges. All point estimates have a negative sign, although none of the effects is significantly different from zero. Close around the discontinuity we obtain identical point estimates, for 2002 and 2003. Increasing the sample to DS$\pm$10, the point estimate is basically zero in 2002 while for 2003 it is very close to the initial estimate of -0.05. Using this latter estimate we can rule out effects larger than 8% of a standard deviation with 0.95 likelihood.

Table 4: Difference-in-differences estimates of the effect of the Personnel Subsidy on test scores

	2002		2003		Pooled		Test $\hat{\delta}_{2002} = \hat{\delta}_{2003}$	
	$\hat{\delta}$	s.e.	$\hat{\delta}$	s.e.	$\hat{\delta}$	s.e.	χ^2	p-val
Language								
DS±5	-0.087	(0.066)	-0.051	(0.071)	-0.069	(0.057)	0.221	0.639
DS±10	-0.049	(0.053)	-0.061	(0.052)	-0.055	(0.043)	0.044	0.833
Arithmetics								
DS±5	-0.050	(0.071)	-0.050	(0.083)	-0.050	(0.066)	0.000	0.996
DS±10	-0.004	(0.053)	-0.043	(0.059)	-0.023	(0.047)	0.391	0.532
Information								
DS±5	-0.115	(0.062)	-0.150	(0.071)	-0.133	(0.056)	0.244	0.637
DS±10	0.001	(0.052)	-0.073	(0.053)	-0.035	(0.043)	1.564	0.212

Note: Standard errors (in parentheses) take into account clustering at the school*year (class) level and are heteroscedasticity robust.

For the scores on the information processing items, we find quite large negative effects for DS±5 which are all significantly different from zero. Increasing the bandwidth around the cutoff to 10 percent reduces the size of estimated effects considerably. For 2002 the effect disappears while for 2003, although the point estimate is reduced by a factor two, it still is -0.073.

Results are fairly robust to changes in the outcome measure and the exact discontinuity sample. It should be noted that different effect estimates for different outcome variables and different years cannot be ruled out. An extra teaching assistant, for example, may affect language skills differently than arithmetic proficiency. Similarly, effects may vary over time following the hiring of extra personnel.

Although never very different, effect estimates for different discontinuity samples vary somewhat in a few cases. It should be noted that increasing the bandwidth around the discontinuity makes observations less comparable. However, in all cases the estimates obtained from DS±10 fall within the 95 percent confidence interval of the DS±5 estimates.

Summarizing, all (but one) point estimates of the effects of the personnel subsidy are negative. In addition, comparing the estimates between years, there is some evidence that the negative effects are not short-term effects. If anything, they seem to be more negative in 2003 than in the previous year. These results show that it is quite unlikely that the personnel subsidy had a substantial positive impact on pupils' achievement measured on any of the three domains covered by the tests.

Our results contrast with those reported by others. Guryan for example estimates that a 10 percent increase in resources increase 4th graders test scores by about 20% of a standard deviation. Greenwald et al. (1996), who perform a meta analysis of (among other things) expenditures per-pupil on test scores using 27 estimates from 14 studies, find (recalculating all expenditures in 1994 dollars for all the studies) that a 10 percent increase in resources generate about 15% standard deviation higher test scores. Given these previous results, with effect sizes between 15% and 20% of a standard deviation, our estimates are very informative since we can rule out much smaller effects.

A very important difference between the circumstances analyzed, is that in the US situation the extra resources were given to schools with relatively few resources whereas in the Dutch situation the extra funds come on top of an already generous compensatory funding scheme. It seems as if the additional resources for disadvantaged primary schools in the Netherlands have reached some threshold point representing resource adequacy (cf. Burtless 1996, p.19).

5.3 Effects of computer subsidy

Table 5 repeats the analysis of the previous subsection, but now for the computer subsidy.

Considering first the effects of the computer subsidy on language scores in DS$\pm$5, we find point estimates which are negative, but not significantly different from zero. Increasing the bandwidth to 10 percent increases the precision of the estimates, which become somewhat more negative. As a result, the negative effects are now statistically significant at the 5% level. Equality between the 2002 and 2003 effects cannot be rejected and the pooled estimate is -0.079 with a standard error of 0.030 and is therefore significant at the 1% level.

Based on DS$\pm$5, the results for the effects on the arithmetics test score are very similar. The effect in 2002 is very close to the one for 2003 and equality cannot be rejected. The pooled estimate is -0.072 of a standard deviation with a standard error 0.047, this rules out positive effects in excess of 2% of a standard deviation with 95 percent confidence. Increasing the bandwidth does not substantially change the picture. The effect is more negative (and significant) in 2003 than in 2002, suggesting that the negative effect is not a short term phenomenon. The pooled estimate of -0.061 is statistically significant at the 10 percent level.

The estimates for information processing are all negative but not statistically significant. The size of the effects is smaller than those on the language and arithmetic domains. The pooled estimate for DS$\pm$10 rules out positive effects larger than 3% of a standard deviation with 95 percent likelihood.

Our findings for the effects of the computer subsidy indicate that extra funds for computers and software do not have a positive impact on pupils' achievement and even seem to have a negative effect on language and arithmetics scores. This finding accords with results from the other recent studies cited in subsection 2.2.

5.4 First stage relations

The analysis so far deals with the effects of the subsidies on pupils' achievement. Like in most policy evaluations, our estimates are not informative about the underlying process that translates subsidies into outcomes. However, from a policy perspective these estimates are very relevant since they inform policy makers about the effect of providing extra resources on the ultimate outcomes of interest.

Nevertheless, one might be interested in how schools actually used the provided subsidies, where it should be noted that it is difficult to draw strong conclusions from this information. This is because how schools allocate money over the different spending categories is obviously a choice variable. Different schools will make different choices depending on their needs. Comparing pupils' achievement between schools that spent the subsidy in different ways to do causal inference is therefore problematic.

To learn more about the anatomy of spending, descriptive information on school spending is interesting, if only to establish that schools actually spent the extra money. In this subsection we present therefore information about how schools used the two subsidies. For the personnel subsidy this information was collected

Table 5: Difference-in-differences estimates of the effect of the Computer Subsidy on test scores

	2002		2003		Pooled		Test $\hat{\delta}_{2002} = \hat{\delta}_{2003}$	
	$\hat{\delta}$	s.e.	$\hat{\delta}$	s.e.	$\hat{\delta}$	s.e.	χ^2	p-val
Language								
DS±5	-0.065	(0.052)	-0.013	(0.051)	-0.039	(0.043)	0.795	0.373
DS±10	-0.088	(0.036)	-0.071	(0.036)	-0.079	(0.030)	0.177	0.674
Arithmetics								
DS±5	-0.065	(0.055)	-0.079	(0.058)	-0.072	(0.047)	0.047	0.829
DS±10	-0.027	(0.037)	-0.095	(0.041)	-0.061	(0.033)	2.483	0.115
Information								
DS±5	-0.021	(0.053)	-0.001	(0.050)	-0.011	(0.043)	0.116	0.733
DS±10	-0.013	(0.036)	-0.053	(0.037)	-0.032	(0.030)	0.942	0.332

Note: Standard errors (in parentheses) take into account clustering at the school*year (class) level and are heteroscedasticity robust.

Table 6: Allocation of personnel subsidy

	%
Hiring and recruitment of extra personnel	36
Teacher training	5
Extra payment of personnel	22
Extra facilities	20
Other	5
Reservation	12

Source: Beerends and van der Ploeg (2001).

by other researchers, for the computer subsidy we sent out a brief questionnaire to all schools in DS$\pm$5.

Personnel spending

Beginning of 2001 the Dutch ministry of education commissioned a research project to gather information about the personnel subsidy. To this end Beerends and Van der Ploeg (2001) contacted the (vice-) principals of all 285 schools that were eligible for the subsidy, in order to have a telephonic interview. Ultimately, they received responses from 65 school principals, who answered questions about how they actually allocated the personnel subsidy.

Table 6 reports the budget shares of different categories. This reveals that large shares of the subsidy were allocated to the hiring and recruitment of new (temporary) personnel and extra payments. Our interpretation of Table 6 is that schools spent almost the entire subsidy as it was intended. The first four categories are clearly consistent with the program's requirement of ”improving working conditions”. Also the category ”other” is not inconsistent with this requirement. Only the 12 percent that goes to the category reservations may not contribute to any estimated impact of the subsidy. It seems likely however that by 2003 (our latest outcome measure) this money was spent as well.

We were unable to obtain these data at the school level. The data collected by

Beerends and Van der Ploeg are however not very suitable for further empirical analysis. First of all, no information was collected among schools which did not receive the personnel subsidy. Second, the share of schools that responded to the interview is very small.

Computer spending

A comparable study as the one conducted by Beerends and Van der Ploeg for the personnel subsidy, is not available for the computer subsidy. We therefore collected information about computer use by sending out a brief questionnaire to 171 schools belonging to DS$\pm$5.[14] This was done in the Spring of 2003. After having approached non-respondents of the written questionnaire by telephone, we obtained information from 153 schools.[15] Sixty-three of these schools were eligible for the computer subsidy; 90 were not. The questionnaire contained no more than 6 questions to keep the effort required from respondents as small as possible (we believe that this contributed to the high response rate). The questions asked about: the number of pupils in highest three grade levels; the number of computers in the school available for these pupils; the age of the computers; and the average numbers of hours per week pupils in the highest three grade levels make use of these computers in total and separately for language and math.[16]

Table 7 reports, for the computer use variables, the mean values and standard deviations separately for the treated and non-treated groups. It also reports the differences with and without controlling for the share of disadvantaged pupils. These

[14]Notice that this number of schools exceeds the schools in DS$\pm$5 in the analysis of achievement. The reason is that we also sent the questionnaire to schools that did not participate in the nationwide test.

[15]This implies a response rate of almost 0.90. This is even a lower bound on the actual response rate because some schools may have closed down between October 1999 (the pupil count date for the computer subsidy) and May/June 2003 (when we interviewed the schools).

[16]Before we designed the questionnaire, we visited some schools and talked to the headmasters in order to find out what could reasonably be asked. Based on this experience we concluded that it was not sensible to ask questions about how up-to-date the schools' software is. Consequently, we have no information on this although schools could spend the computer subsidy on software.

differences and their standard errors are from a WLS-regression of the row variable on a dummy for treatment (and the share of disadvantaged pupils), where the reported numbers of pupils in the three highest grades are the weights.

The first rows in the table show no significant differences between treated and non-treated schools in terms of the computer-pupil ratio and the average age of the computers. The computer-pupil ratio is slightly higher and the computers slightly newer among treated schools than among non-treated schools, but this is reversed when we control for the share of minority pupils. Hence, there are no significant first-stage effects of the subsidy on the computer-pupil ratio and the age of computers.

It is important to note that, independently of the computer subsidy, schools already have nearly one computer for every five pupils. This is high compared to the ”official” target of the government to have one computer for every ten pupils in primary schools. It seems that the hardware needs of the schools in both groups are already satisfied. The computer subsidy is not used to buy more computers or to replace old computers by newer ones.

Although the subsidy does not seem to improve the computer hardware resources in the treatment schools, the next three rows of Table 7 reveal that pupils in the treatment group do spend more time using a computer than pupils in the control group. Controlling for the share of minority pupils, the difference amounts to slightly over fifty minutes per pupil per week. This difference is significant at the 5%-level. Although in absolute terms this is a small effect, relative it is quite substantial. Twenty minutes of this difference are allocated to language, and ten minutes to math. These latter disaggregated estimates lack precision.

The observation that treated schools did not spend their subsidy on hardware combined with the finding that pupils of treated schools use a school-computer more frequently, suggests that the subsidy has been spend to buy software or in-

Table 7: Effect of eligibility of computer subsidy in various intermediate variables

	Control (1)	Treatment (2)	(2)-(1) (3)	(2)-(1) (4)
Computer-pupil ratio	0.173	0.190	0.017	−0.018
	(0.093)	(0.102)	(0.016)	(0.033)
Age of computers (in years)	2.574	2.425	−0.149	0.028
	(1.461)	(1.398)	(0.239)	(0.493)
Computer use (hours p/w):				
-Total	1.543	1.643	0.100	0.851
	(1.208)	(1.234)	(0.205)	(0.418)
-Language	0.637	0.783	0.147	0.326
	(0.657)	(0.739)	(0.116)	(0.239)
-Arithmetics	0.461	0.496	0.035	0.149
	(0.396)	(0.392)	(0.067)	(0.139)
Controlling for s^{99}			No	Yes

vested in Internet connections.

6 Conclusion

This study evaluates two subsidies in primary education. One subsidy provides extra resources to improve teachers' working conditions. The other gives additional funding mainly for computers and software. Both subsidy schemes specify a cutoff level of disadvantaged pupils (differently defined) of 70 percent below which schools receive no extra funding. All schools with at least 70 percent disadvantaged pupils receive the same amounts per teacher or per pupil independent of the exact share of disadvantaged pupils. The cutoff at 70 percent was maintained quite strictly, and manipulation of shares by schools was not possible as the shares of disadvantaged pupils were determined on the basis of information from years prior to the announcement of the subsidies. Due to these features the cutoffs provide very convincing opportunities to evaluate the effects of these two subsidies.

The point estimates of the effects of both the personnel subsidy and the computer subsidy on achievement of 8th graders on language, arithmetics and information processing are negative. For the personnel subsidy these are in most cases not significantly different from zero. For the computer subsidy we find more evidence of negative effects.

The personnel was mainly spent on extra payments for current teachers and on recruiting and hiring extra teachers. While schools could have conditioned the extra payment on performance, this is not what they did. Consequently, the extra payment for current teachers does not provide an incentive to teachers to perform better so that pupils' achievement increases. Recruitment and hiring extra teachers potentially has a beneficial impact on pupils' achievement. That this is not corroborated by our findings is probably due to the fact that the schools targeted by the personnel subsidy already have sufficient (personnel) resources. Recall that in the main funding scheme for Dutch primary schools, minority pupils have a weight of 1.9 times the weight of a non-disadvantaged pupil. A school with say 200 pupils of whom 150 are minority receives from the main funding scheme the same personnel budget as a school with 317 non-disadvantaged pupils.[17] Where the pupil-teacher ratio in entirely non-disadvantaged schools equals 22, this ratio will be below 14 at the school with 75 percent disadvantaged pupils. In this situation it is unlikely that hiring a new teacher will result in a further reduction of average class-size. The most direct channel to increase pupils' achievement is then not used and a boost in achievement is less likely.

From the evidence provided above it seems that the computer subsidy was not used to invest in extra computers or to replace old ones. Given this, and the fact that pupils in treatment schools spend more time in school using a computer, we infer that the computer subsidy was used to buy new software or invest in Internet connections. One might be tempted to attribute the non-positive effect of this in-

[17] The funding scheme gives no compensation for the first 9 percent of weighted students.

tervention to the limited amount of time elapsed between the intervention and the measurement of the outcomes. This is however contradicted by our finding that the effect is more negative two years than one year after the intervention. To explain the negative impact of computers on test scores, Angrist and Lavy (2002) suggest that instruction methods using computers are less effective than other instruction methods. Our results provide additional support for this view.

References

Angrist, J. D. and Lavy, V. (2002). New evidence on classroom computers and pupil learning. *Economic Journal*, 112:735–765.

Beerends, H. and van der Ploeg, S. (2001). Onderzoek vergoeding schoolspecifieke knelpunten. Report OA-230, Regioplan.

Bénabou, R., Kramarz, F., and Prost, C. (2004). ”Zones d’Education Prioritaire”: Much ado about nothing? Working Paper. CREST-INSEE.

Burtless, G. (1996). Introduction and Summary. In Burtless, G., editor, *Does money matter? The effect of school resources on student achievement and adult success*, pages 1–42. The Brookings Institution, Washington, DC.

Card, D. and Payne, A. A. (2002). School finance reform, the distribution of school spending, and the distribution of student test scores. *Journal of Public Economics*, 83:49–82.

Goolsbee, A. and Guryan, J. (2002). The impact of internet subsidies in public schools. Working Paper Series 9090, NBER.

Greenwald, R., Hedges, L. V., and Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research*, 66(3):361–396.

Guryan, J. (2000). Does money matter? regression-discontinuity estimates from education finance reform in Massachusetts. Working Paper 8269, NBER.

Hahn, J., Todd, P., and Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, 24:1141–1177.

Hanushek, E. A. (1994). *Making Schools Work: Improving Performance and Controlling Costs*. The Brookings Institution, Washington, D.C.

Hanushek, E. A. (1996). School resources and student performace. In Burtless, G., editor, *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. The Brookings Institution.

Hanushek, E. A. (2002). Politics, and the Class Size Debate. In Mishel, L. and Rothstein, R., editors, *The Class Size Debate*, pages 37–65. Economic Policy Institute, Washington, DC.

Kirkpatrick, H. and Cuban, L. (1998). Computers make kids smarter - right? *TECHNOS Quarterly For Education and Technology*, 7(2):1–11.

Krueger, A. B. (2003). Economic Considerations and Class Size. *Economic Journal*, 113:F34–F63.

Leuven, E. and Oosterbeek, H. (2004). Evaluating the effects of a tax deducation on training. *Journal of Labor Economics*. Forthcoming.

Papke, L. E. (2003). The effects of spending on test pass rates: Evidence from Michigan. Manuscript.

Rouse, C. E., Krueger, A. B., and Markman, L. (2004). Putting computerized instruction to the test: A randomized evaluation of a ”scientifically-based” reading program. Working Paper Series 10315, NBER.

Shadish, W. R., Cook, T. D., and Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin Company, Boston.

Van der Klaauw, W. (2003). Breaking the link between poverty and low student achievement: Does Title I work? Manuscript, University of North Carolina at Chapel Hill.