

THE EFFECT OF EXTRA FUNDING FOR DISADVANTAGED PUPILS ON ACHIEVEMENT

Edwin Leuven, Mikael Lindahl, Hessel Oosterbeek, and Dinand Webbink*

Abstract—This paper evaluates the effects of two subsidies targeted at schools with large proportions of disadvantaged pupils. The first scheme gives primary schools with at least 70% disadvantaged minority pupils extra funding for personnel. The second scheme gives primary schools with at least 70% pupils from any disadvantaged group extra funding for computers and software. The cutoffs provide a regression discontinuity design that we exploit in a local difference-in-differences framework. For both subsidies we find negative point estimates, which are for some outcomes significantly different from 0. Extra funding for computers and software seems especially detrimental for girls' achievement. The negative effects of extra funding for computers and software are consistent with results from other recent studies casting doubt on the efficacy of computers in schools.

I. Introduction

EDUCATION is often regarded as a potentially powerful tool to combat poverty and reduce social inequality. At the same time it remains difficult to reduce the skill gap between students from different social backgrounds. A country with a long-standing tradition in attempting to promote equality in education is the Netherlands. For purposes of educational policy, two (main) disadvantaged groups have been defined. The first one is pupils who belong to an ethnic minority and whose parents attained a low level of education (no more than lower secondary school). A vast majority of these are children or grandchildren of immigrants from Turkey and Morocco, most of whom came to the Netherlands in the period 1970–1990. The second group consists of pupils with Dutch parents who attained a low level of education. In the remainder of this paper we will refer to these two groups as “disadvantaged minority” and “disadvantaged native,” respectively.

The government's main funding scheme for primary schools assigns extra funding to schools for each pupil belonging to one of the disadvantaged groups. Relative to the funding for nondisadvantaged pupils, the extra funding is 90% for disadvantaged minority pupils and 25% for disadvantaged native pupils. A school with all of its pupils from the disadvantaged minority group thus receives almost twice as much public funding as a school with all its pupils being nondisadvantaged. In the Netherlands, primary education is almost completely publicly funded. In the total population of primary school pupils, 18% belong to the

disadvantaged native group and 13% to the disadvantaged minority group. In 2000 the total amount spent on this compensatory program was \$234 million for 450,000 disadvantaged pupils.

The present paper evaluates two subsidies that were motivated by the belief that the compensation from the main scheme is insufficient, especially for schools with large shares of disadvantaged pupils.¹ This belief was grounded on the observation that despite the generous compensation scheme, disadvantaged pupils perform much worse on standardized tests in eighth grade. The first subsidy provided an extra payment per teacher of about 10% of gross salaries during two consecutive years. Only schools with at least 70% of the pupils in the disadvantaged minority group 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 for instance to hire extra teachers or to give teachers an extra payment. The second subsidy provided a one-time payment of \$90 per pupil, which is about 17% of schools' annual nonpersonnel budget. This money was earmarked for computers, software, and language materials. Only schools where at least 70% of the pupils belong to any of the disadvantaged groups were eligible for this subsidy.

For both interventions the 70% 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 (where the second difference relates to before and after treatment). To this end we combine administrative data with data on the achievement of eighth graders in nationwide exams.

Point estimates of the effects of both subsidies are negative and in some cases significantly so. For both subsidies, the results rule out even modest positive effect on pupil achievement. For both subsidies the treatments that generate these (non) effects are well defined. In the case of the personnel subsidy, treatment provides schools with a specific amount of extra funding per teacher to improve working conditions. In the case of the computer subsidy, treatment gives schools a specific amount of extra funding per

Received for publication August 23, 2004. Revision accepted for publication July 25, 2006.

* Leuven and Oosterbeek are affiliated with the University of Amsterdam, School of Economics, and the Tinbergen Institute. Lindahl is affiliated with the Swedish Institute for Social Research, and Webbink with CPB Netherlands Bureau for Economic Policy Analysis.

We would like to thank seminar participants in Amsterdam, Dublin, Sevilla, Stockholm, and Uppsala, two anonymous referees, and the editor Daron Acemoglu for constructive and stimulating comments. We also thank the Dutch Ministry of Education and CITO (Centraal Instituut voor Toetsontwikkeling, Central Institute for Test Development) for supplying the data used in this paper.

¹ A number of studies evaluate the effects of the compensatory element of the main funding scheme, many of these commissioned by the Dutch government. These studies do not 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.

pupil for computers and software. We will present evidence that schools spent the extra money as intended.

An explanation for the poor performance of the personnel subsidy is that the marginal value of an extra teacher in the targeted schools is low. Since the main funding scheme is generous—the pupil-teacher ratio in these schools is below 14—these schools might have difficulties to spend the extra money in an effective way. Although schools spent 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 nonpermanent increase in teachers' salaries that is not connected to an incentive scheme, it may fail to improve teacher effort or attract better teachers to these schools.

The negative 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 (cf. section II). 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 III provides details of the two programs and describes the data. Section IV outlines the estimation strategy. Section V presents and discusses the empirical findings. The final section summarizes and concludes.

II. Related Studies

A. Educational Resources and Pupils' Achievement

Until fairly recently the consensus among economists was that extra resources have no strong or systematic impact on pupils' achievement. The surveys by Hanushek (1986, 1994, 1996) contributed much to this view. Most of the studies that Hanushek reviewed, however, ignore endogeneity issues. Recent studies, which arguably use more convincing identification strategies, however, also report 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 fourth graders (but not for eighth graders) he finds improved test scores, especially for low-scoring students. Papke (2005) exploits a similar equalization scheme in Michigan and uses panel data to identify the effect on fourth-grade pass rates and seventh-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) use nationwide data to analyze the effects of school finance reforms on the distribution of school spending across richer and poorer districts.

They find that equalization of spending narrows the difference in test score out comes across family background groups. Chay, McEwan, and Urquiola (2005) use an RD design to estimate the impact of extra resources for poor-performing schools in Chile. Correcting for mean-reversion (important because cutoffs were based on schools' mean test scores), they find significant effects on test scores in fourth grade.

A recent study that fails to find an impact of extra resources on achievement of (disadvantaged) pupils is Van der Klaauw (forthcoming), 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 discusses possible explanations for his 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. Bénabou, Kramarz, and Prost (2004) investigate the effects of a compensatory funding scheme directed at schools with disadvantaged students in France, using 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 test scores.

B. Computers and Pupils' Achievement

The evidence on the effect of computers in schools on pupils' achievement is much more 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, they conclude that the value of this research is limited because it does not take endogeneity issues into account.

Three recent studies do address endogeneity, and each of these studies finds 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 eighth graders and for scores on Hebrew, the estimated effects are mostly negative although not significantly different from 0.

Goolsbee and Guryan (2006) 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 and Krueger (2004) study the effects of an instructional computer program called Fast ForWord (FFW). The authors find no evidence that the use of FFW results in 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 and Krueger 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.

III. Programs and Data

A. The Two Programs

In February 2000 the Dutch Ministry of Education announced a personnel subsidy for schools with at least 70% disadvantaged minority pupils. Eligibility was based on the percentage of these pupils that a school had on October 1, 1998, as counted in administrative data. The extra funding amounted to \$2,625 per teacher per year over a two-year period and was paid to schools between May 2000 and March 2001. This amount is roughly equal to 9% of the average annual gross salary of Dutch primary school teachers, and 11% of the annual gross salary of young teachers. This is a substantial intervention, given that personnel costs are about 80% 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; compensation 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, which stipulated that schools with at least 70% of their pupils belonging to any disadvantaged group would receive extra funding in the amount of \$90 per pupil. This is about \$1,250 per class in the eligible schools and equal to roughly 20% of the nonpersonnel budget of these schools. For this scheme the percentage of disadvantaged pupils of a school was based on administrative data counted on October 1, 1999.

A common feature of these two interventions is that they specify a minimum percentage of disadvantaged pupils that schools need to have to qualify for the extra compensation. The personnel subsidy requires at least 70% of disadvantaged minority pupils, and the computer subsidy requires at

least 70% pupils from any disadvantaged group. All treated schools received the same amount per teacher or per pupil.²

B. Data

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 an indicator of 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, the school's denomination, and pupils' gender.³ From another source we merged school-level data of the share of female teachers and teachers' average age.

More than 80% of primary schools participate in a nationwide testing round.⁴ Pupils who are in the highest (eighth) grade take 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, which parents use in their choice of primary school. This gives schools an

² The personnel subsidy is a fixed amount *per teacher*. Because the pupil-teacher ratio decreases with the share of disadvantaged minority pupils (because of the compensation in the main funding scheme), this implies a higher payment *per pupil* at schools with a higher share of disadvantaged minority pupils. Since this per pupil variation within the treatment group varies one-to-one with the share of disadvantaged minority pupils, there is no natural way to exploit it.

³ Unfortunately gender is the only background characteristic registered in the test data. The social background of individual pupils who take the test is unknown.

⁴ 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.

TABLE 1.—TIMING OF EVENTS

October 1, 1998	Reference date for personnel subsidy
February 1999	Nationwide test 1999
October 1, 1999	Reference date for information and communication technology subsidy
February 2000	Nationwide test 2000
February 2000	Decision and announcement of personnel subsidy
May 2000	Payment of \$2,225 per teacher as personnel subsidy
November 2000	Decision and announcement of ICT subsidy
November 2000	Decision and announcement of extra payment of 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

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 attention to it. A common perception is that is many schools preparing and taking this test is the main activity of pupils in their last two years in primary school (seventh and eighth grade).

For our analysis we use data of the test scores from preintervention years 1999 and 2000 and from postintervention years 2002 and 2003.⁵ In the empirical analysis, the scores of schools' pupils on the language, arithmetic, and information processing parts serve as the outcome variables. The scores for world orientation are excluded from the analysis because participation on this part of the test is optional. To standardize the estimated effects, the scores are divided by their standard deviations and normalized to mean 0 relative to the whole population.

Table 1 gives an overview of the timing of the relevant events. This shows 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 payment of the personnel subsidy. We use the 2002 and 2003 test scores as relevant outcome measures for both subsidies.

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 this requires them to anticipate the personnel subsidy one-and-a-half years and the computer subsidy one year prior to the announcements. 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% cutoff and a rise just above. Figure 1 shows the frequency distributions of

schools in the range of 10 percentage points around the cutoff levels of 70%. These distributions are fairly symmetric around the cutoffs, thereby giving no evidence that schools anticipated the implementation of the two programs.⁶

Figure 2 shows a scatterplot of schools' share of disadvantaged minority pupils in 1998 against their share of disadvantaged pupils from any group in 1999. Different symbols are used to indicate which subsidies schools actually received. Schools marked with a dot (“.”) received neither subsidy, schools marked with a circle (“○”) received the computer subsidy but not the personnel subsidy, schools marked with a triangle (“△”) received the personnel subsidy but not the computer subsidy, and schools marked with a plus (“+”) received both subsidies. The dashed lines in the graph indicate the 70% threshold levels for the two subsidies and divide the graph in four parts. Under perfect compliance, all symbols in the southwest part of the plot should be dots, all symbols in the northwest part should be triangles, all symbols in the southeast part should be circles, and all symbols in the northeast part should be pluses.

Actual assignment is very close to perfect compliance. In 1998 there were 7,045 primary schools in the Netherlands, of which 270 (4%) had at least 70% of their pupils belonging to the disadvantaged minority group, thereby qualifying for the personnel subsidy. Out of these 270, 267 schools (99%) actually received the personnel subsidy. Seven schools (0.1%) that fell below the 70% threshold mistakenly received this subsidy. In 1999 there were 7,028 primary schools, of which 564 (8%) had at least 70% of their pupils belonging to any disadvantaged group. 551 (98%) of these schools received the computer subsidy. Sixteen schools (0.2%) with less than 70% of their pupils belonging to any disadvantaged group mistakenly received this subsidy. We do not know the reasons for the misclassifications. In the empirical analysis we deal with the misclassified schools by using eligibility status as instrument for actual treatment.⁷

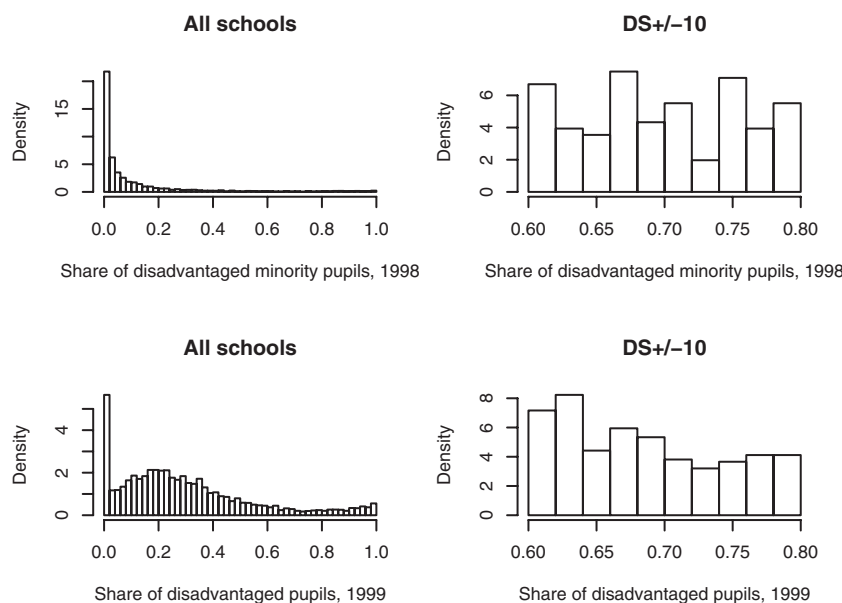
Schools that have at least 70% disadvantaged minority pupils in 1998 are also very likely to have at least 70% disadvantaged pupils in 1999. In other words, schools that qualify for the personnel subsidy are also very likely to qualify for the computer subsidy. In the empirical analysis we focus on schools with their shares of disadvantaged minority pupils or disadvantaged pupils from any group at most 10 percentage points away from the 70% thresholds. In figure 2 these schools are located in the areas enclosed by the dotted horizontal lines (for the personnel subsidy) and the dotted vertical lines (for the computer subsidy).

⁶ Distributions for adjacent years look almost the same.

⁷ Figure 2 demonstrates the first-stage effects of eligibility on actual assignment for both subsidies. Regressing actual assignment on eligibility status (controlling for a third-order polynomial in the relevant share of disadvantaged pupils) results in coefficients of 0.93 (s.e. 0.006) for the personnel subsidy and 0.95 (s.e. 0.007) for the computer subsidy.

⁵ We do not use data from 2001 because it is unclear whether this is a pre- or a postintervention year.

FIGURE 1.—DISTRIBUTION OF SCHOOLS



For the personnel subsidy, 125 schools fall in this 60%–80% range; only five of these schools (4%) were not eligible for the computer subsidy. On the other hand, 328 schools fall in the 60%–80% range for the computer subsidy and only eleven (3%) of these were also eligible for the personnel subsidy. This implies that we evaluate the impact of the personnel subsidy conditional on schools (treatment and control) also receiving the computer subsidy, and that the computer subsidy is evaluated conditional on schools (again treatment and controls) not receiving the personnel subsidy.

IV. 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.

A. Regression Discontinuity Design

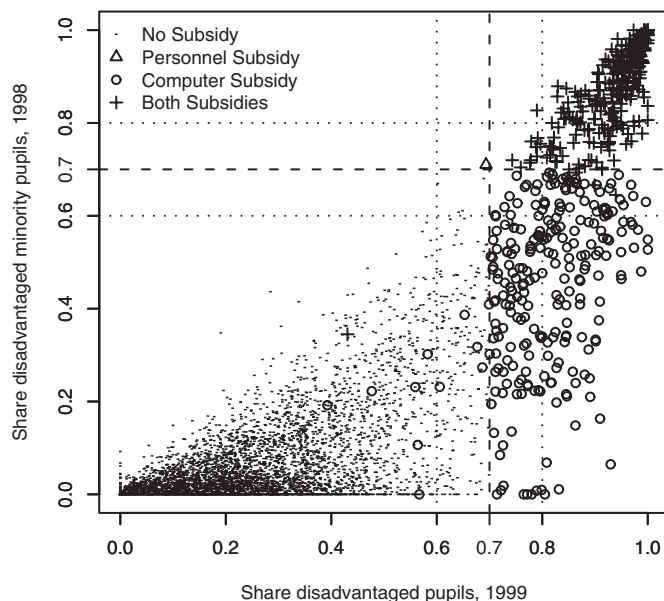
The eligibility rule of the personnel subsidy specifies that all schools with at least 70% disadvantaged minority pupils receive the subsidy and all schools with less than 70% of such pupils do not receive the subsidy. Without exceptions to this rule we would have a so-called sharp regression discontinuity design in which treatment depends in a deterministic way on the share of disadvantaged minority pupils.⁸

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% of disadvantaged pupils if there are no confounding discontinuities at the threshold.

Denote the share of disadvantaged 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:

⁸ Leuven and Oosterbeek (2004) provide an application of the sharp regression discontinuity design. Jacob and Lefgren (2004) is a recent example of a regression discontinuity design in the context of education.

FIGURE 2.—ASSIGNMENT OF SUBSIDIES



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

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, Todd, & Van der Klaauw, 2001)

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

where $y^+ \equiv \lim_{s \downarrow 0.7} E[y|s]$ and $y^- \equiv \lim_{s \uparrow 0.7} E[y|s]$.

B. Estimation

To fully exploit the available data (including outcomes from pre- and postintervention years), our preferred strategy is a difference-in-differences approach that increases the power and gives more precise estimates. In terms of implementation, we estimate fixed-effect regressions of the following form:

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

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 that equals 1 if school j received the subsidy, η_j is a school fixed effect, m_t are time effects (dummies), and ε_{ijt} is an error term.⁹ Note that the estimate of δ in equation (1) recuperates the standard difference-in-differences estimate when restricting the sample to one postintervention and one preintervention year. Below, we estimate equation (1) on a sample of two preintervention and two postintervention 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 $x\%$ discontinuity sample ($DS \pm x$) consists of the eligible group of schools with their percentage of disadvantaged minority pupils at most x percentage points above the cutoff of 70%, and the noneligible group of schools with their percentage of disadvantaged minority pupils at most x percentage points below the cutoff of 70%. Widening the bandwidths around the discontinuity increases

the number of observations but at the same time increases 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.

The major identifying assumption for our approach, to provide unbiased estimates of the effects of the subsidies, is that there are no other discontinuities around the cutoffs of 70%. This is an exclusion restriction with respect to the discontinuity. A nice feature of the regression discontinuity design is that it allows us to control for smooth functions of the variable determining eligibility (that is, the fractions of disadvantaged [minority] pupils). In the difference-in-differences framework, these fractions enter in the form of fixed effects. We will also present results where the fixed effects are replaced by polynomials in the shares of disadvantaged (minority) pupils (and other school characteristics).

As mentioned in the previous section, a few schools did (not) receive the personnel subsidy although they had less (more) than 70% disadvantaged minority pupils. Because the rule behind these exceptions is unknown, this breaks down the sharp regression discontinuity design. There is no longer a deterministic relation between treatment and the share of disadvantaged minority pupils. To address this issue we apply 2SLS in which treatment is instrumented by eligibility.

V. Results

A. Data Description

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 that we estimate are local, it is important to know how these samples compare with the population as a whole.

As seen in the first three rows of column 1, we standardized the test scores to have mean of 0 and standard deviation of 1 in the population. Compared with the average student, the pupils in the schools around the personnel cutoff score are on average more than half a standard deviation lower on both the language and the information-processing test. Performance in arithmetics is about one-third of a standard deviation lower in these schools compared with the population average. For the local samples around the computer eligibility cutoff, test scores for language and information processing are more than 0.4 of a standard deviation below the population average. For arithmetic, the difference is somewhat above 0.3 of a standard deviation.

The schools that are (almost) eligible for the personnel subsidy are in the two most disadvantaged groups of the socioeconomic classification index of the school population, whereas the schools in the computer subsidy sample have on average less-disadvantaged pupils. To compare, in the

⁹ For all the results we report heteroskedasticity robust standard errors while allowing for clustering at the school-year level.

TABLE 2.—SAMPLE MEANS FOR POPULATION AND ESTIMATION SAMPLES, 2002

	Population (1)	Personnel		Computer	
		DS $\pm$ 5 (2)	DS $\pm$ 10 (3)	DS $\pm$ 5 (4)	DS $\pm$ 10 (5)
Language (s.d.)	0.000 (1.000)	−0.538 (1.058)	−0.587 (1.048)	−0.428 (1.062)	−0.442 (1.065)
Arithmetic (s.d.)	0.000 (1.000)	−0.343 (1.074)	−0.412 (1.081)	−0.318 (1.057)	−0.320 (1.063)
Information (s.d.)	0.000 (1.000)	−0.549 (1.107)	−0.602 (1.083)	−0.441 (1.086)	−0.423 (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
Socioeconomic 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	150,821	1,817	3,392	3,954	8,263
Number of schools	5,938	63	124	150	328

whole population the vast majority of the schools have pupils 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 with 50% for the computer subsidy sample, and only 15% of the total population.

Finally, disadvantaged 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 with 32% in the population. In contrast, the denomination of the schools that find themselves around the cutoff for the computer subsidy is quite similar to those in the population.

We next investigate whether the samples just below and just above the cutoffs are different in terms of observed characteristics. Table 3 reports such information for the subsamples at most 10 percentage points below and 10 percentage points above the cutoffs. The top rows report the average of the school means for the pretreatment and posttreatment outcome variables.

The samples just below and just above the cutoffs for both subsidies are clearly very similar in terms of the pretreatment outcomes. In most cases the schools in the samples just above the cutoffs perform somewhat worse than the schools in the samples just below the cutoffs. This is not unexpected since the schools just above the cutoffs have higher fractions of disadvantaged pupils. Tested at the

school level, none of these differences is, however, significantly different from 0.

The averages of the school means of the posttreatment test scores are substantially lower in the just-above than in the just-below groups. These differences are not significantly different from 0 in the subsample around the cutoff for the personnel subsidy, but are significant in the subsample around the cutoff for the computer subsidy. The next two subsections examine whether these patterns are robust when the estimation procedure described in section IV is applied.

The remainder of the table reports the mean values for other school characteristics. Significant differences between schools just below and just above the cutoffs are observed for the shares of disadvantaged pupils and for the socioeconomic index. This is not surprising because the share of disadvantaged pupils determines schools assigned into the below and above groups, and schools' socioeconomic index is closely related to how disadvantaged a school's pupils are. For the schools around the cutoff for the personnel subsidy, there also appears to be a significant difference in terms of schools' denomination. The just-below sample has a larger share of public schools than the just-above sample. For schools around the computer subsidy cutoff, the degree of urbanization is somewhat different. Schools just above this cutoff are more often located in high-density areas than schools just below this cutoff. For all the other variables no significant differences between the just-above and just-below groups are observed.

TABLE 3.—SAMPLE MEANS JUST ABOVE AND JUST BELOW SAMPLES

	Personnel			Computer		
	DS - 10 (1)	DS + 10 (2)	<i>p</i> -value (3)	DS - 10 (4)	DS + 10 (5)	<i>p</i> -value (6)
Language pretreatment	-0.571	-0.570	0.996	-0.409	-0.424	0.728
Language posttreatment	-0.585	-0.684	0.093	-0.397	-0.535	0.001
Arithmetics pretreatment	-0.440	-0.477	0.559	-0.333	-0.327	0.897
Arithmetics posttreatment	-0.389	-0.470	0.170	-0.287	-0.375	0.032
Information pretreatment	-0.648	-0.693	0.428	-0.409	-0.421	0.453
Information posttreatment	-0.574	-0.663	0.160	-0.370	-0.489	0.005
Share minority 1998 (s^{98})	0.648	0.751	0.000	0.284	0.423	0.000
Share disadvantaged 1999 (s^{99})	0.843	0.871	0.034	0.646	0.751	0.000
Socioeconomic Index			0.000			0.000
1 (least disadvantaged)	0.000	0.000		0.000	0.000	
2	0.000	0.000		0.000	0.000	
3	0.000	0.000		0.034	0.008	
4	0.000	0.000		0.301	0.148	
5	0.031	0.017		0.447	0.320	
6	0.708	0.267		0.209	0.426	
7 (most disadvantaged)	0.262	0.717		0.010	0.098	
Urbanization School Area			0.151			0.048
Very high	0.554	0.667		0.286	0.385	
High	0.277	0.233		0.272	0.238	
Median	0.077	0.067		0.131	0.180	
Modest	0.046	0.033		0.194	0.107	
Low/none	0.046	0.000		0.117	0.090	
School Denomination			0.018			0.873
Public	0.615	0.333		0.505	0.484	
Catholic	0.185	0.383		0.262	0.303	
Protestant	0.169	0.250		0.199	0.189	
Montessori/Daltonian	0.015	0.033		0.029	0.025	
Other	0.015	0.000		0.005	0.000	
Share female teachers	0.752	0.738	0.354	0.747	0.742	0.633
Average age teachers	42.0	41.4	0.362	41.4	41.5	0.882
Number of test-taking students	29.3	24.8	0.093	24.8	25.8	0.516
School size 1998	268.8	245.9	0.338	225.3	241.1	0.216
School size 1999	271.1	251.0	0.279	224.6	240.7	0.233
Number of schools	65	60		206	122	

To summarize, the results in table 3 show that schools above and below the cutoffs are very similar in terms of pretreatment outcome variables and in terms of most background characteristics. They differ in background characteristics that determine treatment assignment (or are closely tied to that). Moreover, there appear to be some differences in terms of denomination (around the personnel cutoff) and urbanization (around the computer cutoff). These differences are controlled for in the empirical strategy in the form of covariates or fixed effects.

B. Effects of the Subsidies

Standard RD results. Postintervention outcomes conditional on the share of disadvantaged (minority) pupils can be estimated using local linear regression (Cleveland, 1979). Following Hahn, Todd, and Van der Klaauw (2001), this is done by stratifying the sample on whether schools have at least 70% disadvantaged (minority) pupils, before doing the local linear regression (using a bandwidth of 0.05). Figure 3 plots the fractions of disadvantaged (minority) pupils against school averages of the outcomes. The smoothed test scores in the graphs show a small but clear

downward jump at the 70% cutoffs. This is a first indication that the subsidies had no positive impact on achievement. To check whether such jumps are not an artifact of the stratification, we additionally stratified the sample at other discontinuities (10%, 30%, 50%, 90%). No jumps are observed at these arbitrary cutoffs.

It is difficult to unambiguously detect effects from casual inspection of the graphs because the size of effects (if any) will be small relative to the variance in the data. For this reason we proceed in panel A of table 4 with presenting standard regressions discontinuity estimates for the whole sample of schools. These results are obtained by regressing individual test scores on two dummies for treatment status. The first dummy equals 1 for pupils in schools that obtained the personnel subsidy, and 0 otherwise. The second dummy equals 1 for pupils in schools that obtained the computer subsidy, and 0 otherwise. Actual receipt of the subsidies is instrumented by eligibility status. The regressions include controls for polynomials of the fractions of disadvantaged minority pupils and pupils from any disadvantaged group. The regressions also include year dummies and interactions of the polynomials in fractions of disadvantaged pupils and

year dummies. These interactions should capture the effects of differential trends in the achievement of disadvantaged pupils. (It turns out, however, that inclusion of these interactions does not influence the results.) Furthermore, the regressions include dummies for pupils' gender, denomination, and degree of urbanization, and teachers' age and gender, number of pupils taking the test, and school size in 1998/9 as covariates.

Of the 5,938 schools that participated in the nationwide tests in 2002 and 2003, 254 were eligible for both subsidies, 257 were eligible for the computer subsidy but not for the personnel subsidy, and only one school was eligible for the personnel subsidy but not for the computer subsidy (this is a school with 71.0% disadvantaged minority pupils at October 1, 1998, and 69.3% pupils from any disadvantaged group at October 1, 1999). This distribution makes it impossible to disentangle the effects of receiving only the personnel subsidy and receiving the personnel subsidy and the computer subsidy jointly. We could therefore not include an interaction term for receipt of both subsidies.

Table 4 presents results from three different specifications. The first specification includes only linear controls for the share of disadvantaged minority pupils in 1998 and share of pupils from any disadvantaged group in 1999. In the second and third specifications this is increased to quadratic and cubic controls, respectively. Some of the point estimates change substantially when second-order terms are added. Adding third-order terms causes no further changes. These point estimates indicate that both subsidies had a negative impact on achievement. The effects of the personnel subsidy are significantly different from 0; for the computer subsidy the effect estimates lack precision.

Panel B of table 4 presents results from comparable regressions when the samples are restricted to pupils in schools that are at most 10 percentage points away from the 70% cutoffs. As we mentioned, there is (almost) no variation in eligibility status for the other subsidy within these discontinuity samples. We therefore consider both subsidies in isolation. All estimates have the same negative sign as the estimates based on the entire sample of schools. The estimates of the effects of the personnel subsidy are, however, smaller in absolute size. Furthermore, the estimates change only marginally when higher-order terms of the fractions of disadvantaged pupils are added.

DID Results—Personnel Subsidy Based on Discontinuity Samples. We continue by reporting the results from the difference-in-difference procedure outlined in section IV using the $\pm 5\%$ and $\pm 10\%$ samples around the discontinuity. Table 5 reports the findings for the personnel subsidy on the three outcome variables. We report the effects for the postintervention years 2002 and 2003 separately since, strictly speaking, these are different outcomes. We also report a pooled estimate for 2002 and 2003 that is more precise, and the statistic of the test for equality of the effects for the separate years.

First notice that the results reported in table 5 are not very different from those reported in table 4. All estimates have the same sign and their magnitudes and precision are in the same ballpark.

Next 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 -0.156 . The effects for the subsample with a wider bandwidth around the discontinuity, $DS \pm 10$, are also negative but somewhat smaller. The pooled estimate is -0.098 with a standard error of 0.050. We can therefore rule out positive effects on language scores with a 95% probability. The point estimate of -0.098 should be interpreted as the total effect of the personnel subsidy. Increasing schools' budget for personnel by 9% for two consecutive years reduces the average language test score measured two to three years later by 9.8% of a standard deviation. All other point estimates pertaining to the personnel subsidy can be interpreted similarly.

For the arithmetics scores, a very similar pattern emerges. All point estimates have a negative sign, but only the effect on the score in 2003 in $DS \pm 5$ is significantly different from 0. Using the pooled estimate from $DS \pm 10$, we can rule out effects larger than 5% of a standard deviation with 0.95 likelihood.

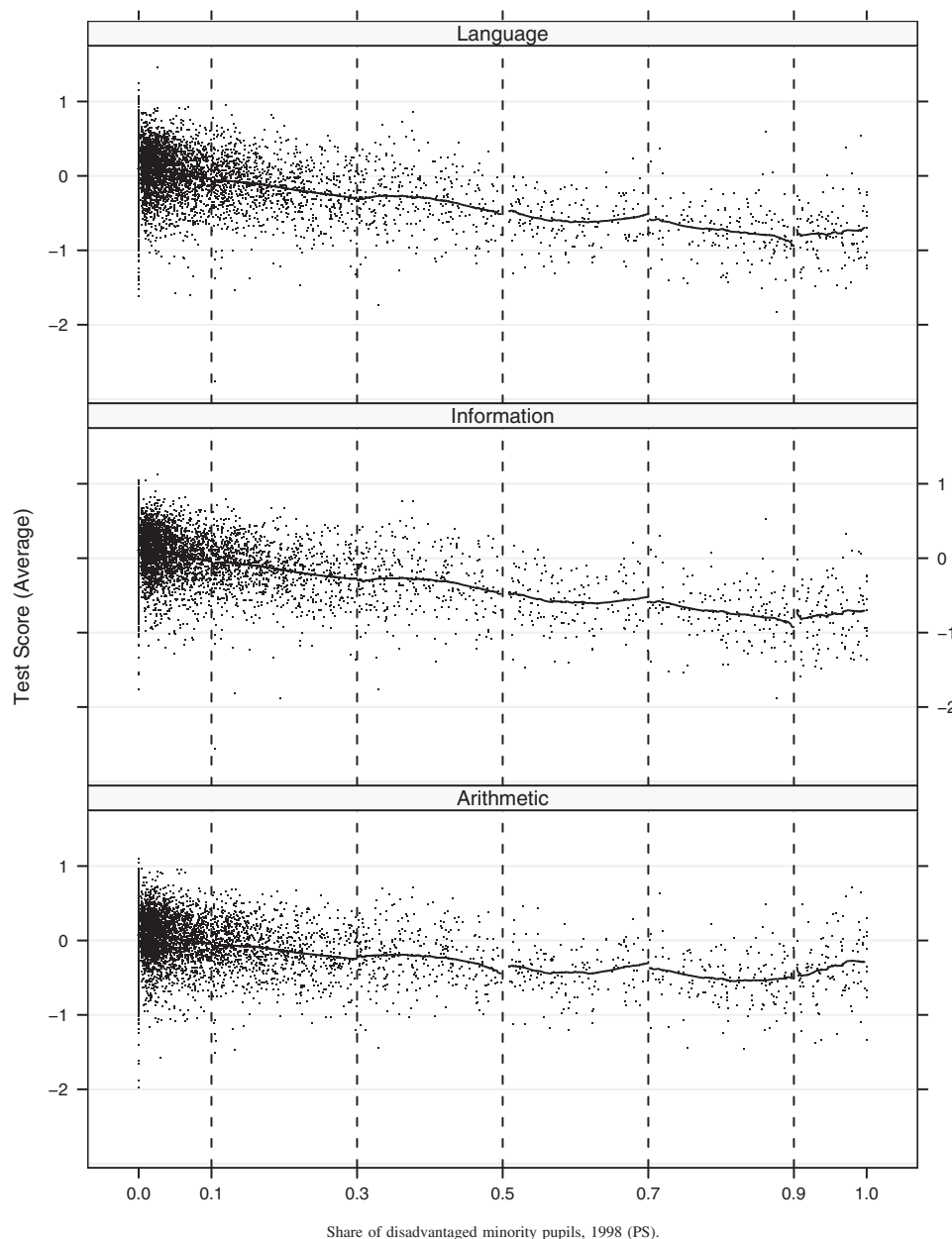
For the scores on the information-processing items, we find quite large negative effects for $DS \pm 5$ that are significantly different from 0 for the 2003 test score and the pooled estimate. Increasing the bandwidth around the cutoff to 10% reduces the size of estimated effects considerably. For 2002 the effect disappears while for 2003, although the point estimate is reduced by a factor two, it still is -0.120 .

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 $DS \pm 10$ fall within the 95% confidence interval of the $DS \pm 5$ estimates.

In subsequent analyses we estimated the regressions separately for girls and boys. The point estimates are a bit more negative for girls than for boys, but the differences are never significant. We also regressed the difference between the test scores at the 90th and 10th percentiles within a school on receipt of the personnel subsidy. This could indicate

FIGURE 3.—AVERAGE SCHOOL TEST SCORE IN 2002/2003 BY SHARE OF DISADVANTAGED PUPILS, DISCONTINUITY SAMPLE



whether the subsidy had differential impacts on pupils at different parts of the ability distribution. The effects are, however, never significant nor is there any consistency in the signs.¹⁰

As a robustness check, we conducted the same analyses using the arbitrary (“fake”) cutoff at 50% disadvantaged minority pupils. The results of this exercise are reported in the appendix. For these analyses we had to assume that treatment status coincides with eligibility. As they ought to, these results show no systematic pattern at all. None of the estimated effects is significantly different from 0, and the

¹⁰ The same holds if differences between the 75th and 25th percentiles are considered.

numbers of positive and negative point estimates are almost equal (8 versus 10). This is consistent with the pattern in figure 3. This is further evidence that the extra resources from the personnel subsidy did more harm than good.

Summarizing, all 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.

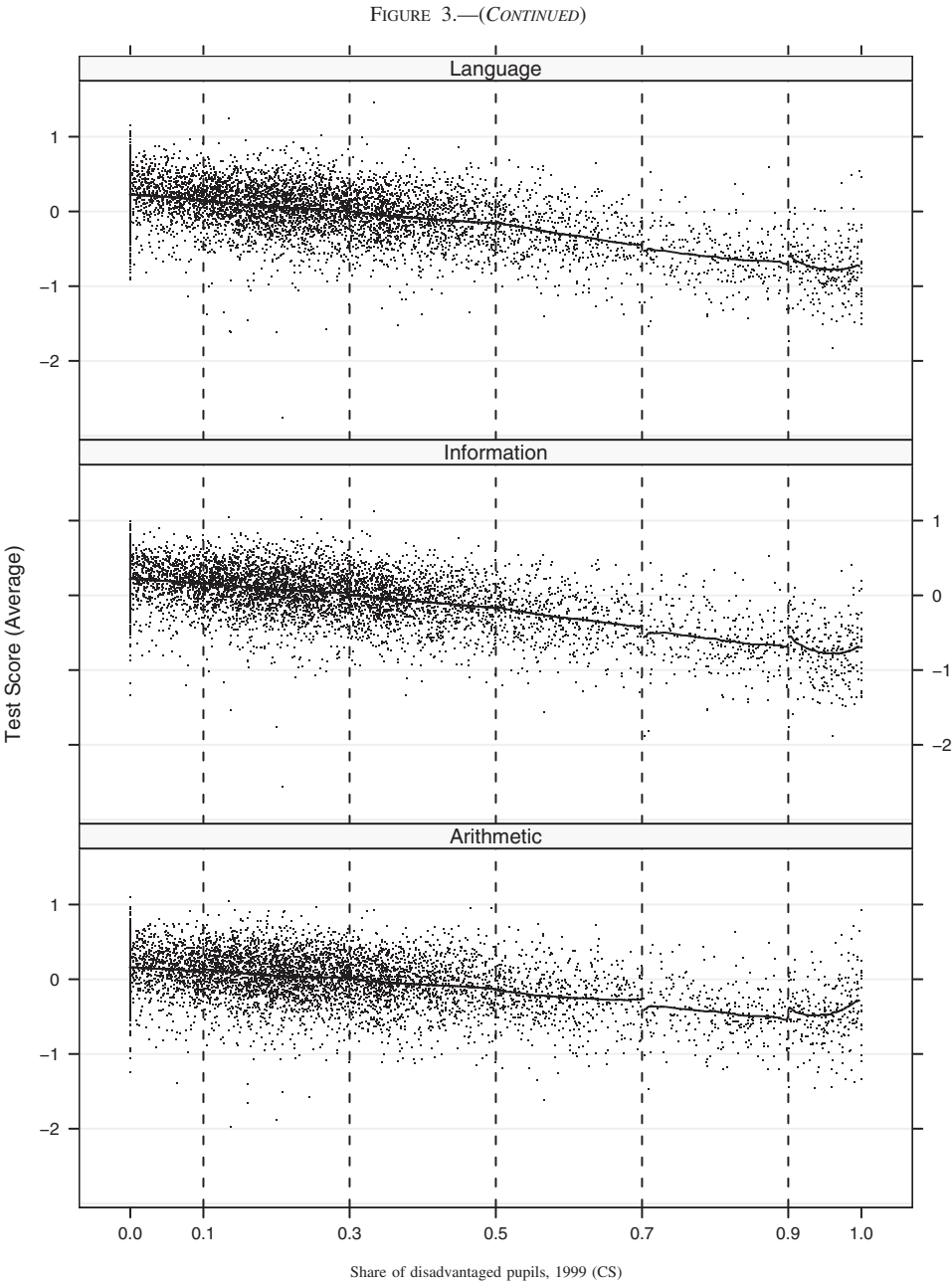


TABLE 4.—RD ESTIMATES OF THE EFFECT OF THE PERSONNEL SUBSIDY (PS) AND THE COMPUTER SUBSIDY (CS) ON TEST SCORES

	(1)		(2)		(3)	
	PS	CS	PS	CS	PS	CS
A. All Schools						
Language	−0.012 (0.033)	−0.086 (0.025)	−0.138 (0.058)	−0.057 (0.034)	−0.118 (0.059)	−0.055 (0.038)
Arithmetics	0.084 (0.034)	−0.057 (0.024)	−0.202 (0.058)	−0.059 (0.034)	−0.206 (0.060)	−0.046 (0.038)
Information	−0.020 (0.034)	−0.072 (0.025)	−0.132 (0.059)	−0.023 (0.034)	−0.116 (0.060)	−0.026 (0.038)
B. DS ± 10						
Language	−0.059 (0.079)	−0.070 (0.042)	−0.059 (0.078)	−0.066 (0.042)	−0.079 (0.080)	−0.066 (0.043)
Arithmetics	−0.015 (0.078)	−0.055 (0.043)	−0.014 (0.078)	−0.048 (0.044)	−0.027 (0.081)	−0.037 (0.046)
Information	−0.033 (0.079)	−0.038 (0.042)	−0.032 (0.078)	−0.035 (0.042)	−0.041 (0.080)	−0.030 (0.044)
Degree polynomial in fractions disadv.		1st		2nd		3rd

Note: Standard errors (in parentheses) take into account clustering at the school × year (class) level and are heteroskedasticity robust. Other covariates are dummy for pupils' gender, dummies for denomination and degree of urbanization, teachers' age and gender, number of pupils taking the test, school size in 1998/9, and interactions of the polynomials of the fractions of disadvantaged pupils and year dummies. For the entire sample, the estimates of the effects of the personnel and computer subsidy for a given specification and outcome come from a single regression. For the discontinuity samples, these come from separate regressions.

TABLE 5.—DIFFERENCE-IN-DIFFERENCES IV ESTIMATES OF THE EFFECT OF THE PERSONNEL SUBSIDY ON TEST SCORES

	$\hat{\delta}_{2002}$ (s.e.)	$\hat{\delta}_{2003}$ (s.e.)	$\hat{\delta}_{Pooled}$ (s.e.)	Test $\hat{\delta}_{2002} = \hat{\delta}_{2003}$	
				χ^2	p -value
Language					
DS $\pm$ 5	-0.124 (0.081)	-0.187 (0.089)	-0.156 (0.071)	0.466	0.495
DS $\pm$ 10	-0.075 (0.061)	-0.121 (0.061)	-0.098 (0.050)	0.459	0.498
Arithmetics					
DS $\pm$ 5	-0.050 (0.091)	-0.160 (0.061)	-0.106 (0.083)	1.220	0.270
DS $\pm$ 10	-0.026 (0.062)	-0.087 (0.069)	-0.056 (0.055)	0.724	0.395
Information					
DS $\pm$ 5	-0.111 (0.075)	-0.296 (0.091)	-0.204 (0.070)	4.078	0.045
DS $\pm$ 10	-0.004 (0.060)	-0.120 (0.064)	-0.061 (0.051)	2.759	0.097

Note: Standard errors (in parentheses) take into account clustering at the school $\times$ year (class) level and are heteroskedasticity robust.

Our results contrast with those reported by others. Guryan (2000), for example, estimates that a 10% increase in resources increases fourth graders' test scores by about 20% of a standard deviation. Greenwald, Hedges, and Laine (1996), who perform a metaanalysis of (among other things) expenditures per pupil on test scores using 27 estimates from fourteen studies, find (recalculating all expenditures in 1994 dollars for all the studies) that a 10% increase in resources generates 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 for an intervention of comparable size.¹¹

An important difference between the circumstances analyzed is that in the U.S. 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 the threshold point of resource adequacy (cf. Burtless, 1996, p. 19).

DID Results—Computer Subsidy Based on Discontinuity Samples. Table 6 repeats the previous analysis, but now for the computer subsidy. Again the estimates are comparable to those reported in table 4. Considering first the effects of the computer subsidy on language scores in DS $\pm$ 5, we find point estimates that are negative, but not significantly different from 0. Increasing the bandwidth to 10% 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 -0.078 with a standard error of 0.031 and is therefore significant at the 1% level. The point estimate of -0.078 should be interpreted as the total effect of the computer subsidy. That is: the one-time subsidy of

\$90 per pupil for computers and software leads to a reduction of language test score measured one to two years later by 7.8% of a standard deviation. All other point estimates pertaining to the computer subsidy can be interpreted similarly.

Based on DS $\pm$ 5, the results for the effects on the arithmetics test score are very similar. The effect in 2002 is smaller than the one for 2003, but equality cannot be rejected. 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 is -0.050 of a standard deviation with a standard error 0.034, this rules out positive effects in excess of 2% of a standard deviation with 95% confidence.

With one exception, the point 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% likelihood.

Also for the computer subsidy, we conducted separate analyses for girls and boys. These separate results are reported in the appendix. The evidence clearly indicates that the effects of the computer subsidy are more negative for girls than for boys. For all three outcomes, the effects for girls obtained from DS $\pm$ 10 are significantly negative, whereas for boys none of the effects differs significantly from 0. Apparently girls suffer from the availability of additional resources for computers and software, whereas boys don't.

We also estimated regressions of the school differences between the 90th and 10th (and 75th and 25th) percentiles, to see whether the computer subsidy had differential impacts on pupils at different parts in the ability distribution. Like for the personnel subsidy, we find no indication for such differential impacts. Again we repeated the analyses using the arbitrary cutoff at 50% disadvantaged pupils. The results are reported in the appendix. Again these results show no systematic pattern. This is further evidence that also the extra resources from the computer subsidy did more harm than good.

¹¹ The interventions are comparable in relative size: a 10% increase in resources. If the bases are very different, these increases may be very different in an absolute sense. But since average education expenditures in developed countries are not too different, comparable relative increases imply comparable absolute increases. Comparing in relative terms avoids complications due to differences in purchasing power.

TABLE 6.—DIFFERENCE-IN-DIFFERENCES IV ESTIMATES OF THE EFFECT OF THE COMPUTER SUBSIDY ON TEST SCORES

	$\hat{\delta}_{2002}$ (s.e.)	$\hat{\delta}_{2003}$ (s.e.)	$\hat{\delta}_{Pooled}$ (s.e.)	Test $\hat{\delta}_{2002} = \hat{\delta}_{2003}$	
				χ^2	p-value
Language					
DS $\pm$ 5	-0.047 (0.056)	-0.025 (0.050)	-0.036 (0.046)	0.130	0.718
DS $\pm$ 10	-0.080 (0.037)	-0.076 (0.037)	-0.078 (0.031)	0.008	0.929
Arithmetics					
DS $\pm$ 5	-0.027 (0.059)	-0.061 (0.063)	-0.044 (0.051)	0.255	0.614
DS $\pm$ 10	-0.011 (0.038)	-0.090 (0.043)	-0.050 (0.034)	3.073	0.080
Information					
DS $\pm$ 5	0.004 (0.057)	-0.010 (0.053)	-0.003 (0.046)	0.056	0.813
DS $\pm$ 10	-0.000 (0.037)	-0.060 (0.039)	-0.029 (0.032)	1.959	0.162

Note: Standard errors (in parentheses) take into account clustering at the school $\times$ year (class) level and are heteroskedasticity robust.

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. The results even indicate that girls' achievement has suffered from these extra funds.

C. 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 policymakers 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 therefore present information about how schools used the two subsidies. For the personnel subsidy this information was collected by other researchers, for the computer subsidy we sent out a brief questionnaire to all schools in DS $\pm$ 5.

Personnel Spending. At the 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) interviewed the (vice-) principals of the schools eligible for the subsidy by telephone. Ultimately, they received responses from 65 school principals, who answered questions about how they actually allocated the personnel subsidy.¹²

¹² The principals of these schools (which all received the personnel subsidy) were also asked whether they thought that the subsidy was

Table 7 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 7 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."¹³ Also, the category "other" (where respondents mentioned such things as audiovisual devices and teaching materials) is not inconsistent with this requirement. Only the 12% that goes to the category "not yet spent" 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 (the interviews were held before the final payment).

We were unable to obtain the data collected by Beerends and van der Ploeg. These data are however not very suitable for further empirical analysis. First, no information was collected among schools that did not receive the personnel subsidy. Second, the share of schools that responded to the interview is quite small. Unfortunately there are no other data sources having systematic information about the hiring of teachers on temporary contracts. This makes it impossible to examine what happened to pupil-teacher ratios. A possible scenario, which was brought to our attention when we discussed our findings with people in the field, is that temporary teachers replaced regular teachers and that these replaced regular teachers were assigned management tasks. This implies that the actual teaching was done by less experienced teachers, which could explain a negative effect of the personnel subsidy.

Another possible explanation for our finding of no effects of the personnel subsidy is that schools directed the extra funds to their younger pupils, thereby making it impossible to find a short-term effect. Although we have no data with which we can reject this explanation, we consider it rather unlikely that a substantial part of the treated schools followed such a

effective. This question did not specify how effectiveness should be measured. Over 80% of the interviewed principals responded that the subsidy was indeed effective. This "evaluation" later on played a role in the ministry's decision to continue this subsidy.

¹³ The category "extra facilities" may include such things as new coffee machines, compensations for housing costs, traveling costs, or child care facilities.

TABLE 7.—ALLOCATION OF PERSONNEL SUBSIDY

	%
Hiring and recruitment of extra personnel	36
Teacher training	5
Extra payment of personnel	22
Extra facilities	20
Other	5
Not yet spent	12

Source: Beerends and van der Ploeg (2001).

policy. In 1996 the Dutch government decided to supply primary schools with extra funding that was earmarked to reduce class size in grades 1 to 4. This extra funding was phased in during the period 1996–2003. Many primary schools objected against this policy because they wanted to use part of the extra funding for pupils in the higher grades. Until 2003 this was explicitly forbidden. Hence, if schools in these circumstances obtain the personnel subsidy, it seems more likely that they spend the larger share for the higher grades.

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 nonrespondents of the written questionnaire by telephone, we obtained information from 153 schools. Of these schools, 63 were eligible for the computer subsidy; 90 were not. The questionnaire contained no more than six 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 the 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.¹⁵

Table 8 reports, for the computer-use variables, the mean values and standard deviations separately for the treated and nontreated groups. It also reports the differences with and without controlling for the share of disadvantaged pupils. These 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 nontreated 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 nontreated schools, but this is reversed when we control for the share of disadvantaged 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 with 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 8 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 disadvantaged minority pupils, the difference amounts to slightly over 50 minutes per pupil per week. This difference is significant at the 5% level. Twenty minutes of this difference are allocated to language, and ten minutes to math. These latter disaggregated estimates lack precision.

It may seem that even if computer instruction in language and arithmetics is ineffective, twenty and ten minutes of extra computer instruction in these subjects per week is insufficient to bring test scores down. It is important to notice, however, that the full 50 minutes of computer use cannot be spent on regular instruction in language and arithmetics. Moreover, when all pupils in the class spend an extra 50 minutes per week using a computer, this may be disruptive for their classmates.

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 spent to buy software or to invest in Internet connections.

VI. 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% below which schools receive no extra funding. All schools with at least 70% disadvantaged pupils receive the same amounts per teacher or per pupil independent of the exact share of disadvantaged pupils. The cutoff at 70% was maintained quite strictly, and manipulation of shares by schools was not possible because the shares of disadvantaged pupils were determined on the basis of information from years prior to the announcement of the subsidies. Because of these features, the cutoffs provide

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 Before we designed the questionnaire, we visited some schools and talked to the headmasters 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.

TABLE 8.—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 (0.093)	0.190 (0.102)	0.017 (0.016)	–0.018 (0.033)
Age of computers (in years)	2.574 (1.461)	2.425 (1.398)	–0.149 (0.239)	0.028 (0.493)
Computer Use (hours p/w):				
Total	1.543 (1.208)	1.643 (1.234)	0.100 (0.205)	0.851 (0.418)
Language	0.637 (0.657)	0.783 (0.739)	0.147 (0.116)	0.326 (0.239)
Arithmetics	0.461 (0.396)	0.496 (0.392)	0.035 (0.067)	0.149 (0.139)
Controlling for s^{99}			No	Yes

very convincing opportunities to evaluate the effects of these two subsidies.

Point estimates of the effects of both subsidies are negative and in some cases significantly so. The computer subsidy seems to be particularly harmful for girls' achievement. For both subsidies, the results rule out even modest positive effect on pupils' achievement.

The personnel subsidy 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, they did not. 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, making the marginal value of additional teachers low. Recall that in the main funding scheme for Dutch primary schools, disadvantaged minority pupils have a weight of 1.9 times the weight of a nondisadvantaged pupil. A school with, say, 200 pupils of whom 150 are from the disadvantaged minority group receives from the main funding scheme the same personnel budget as a school with 317 nondisadvantaged pupils.¹⁶ Where the pupil-teacher ratio in entirely nondisadvantaged schools equals 22, this ratio will be below 14 at the school with 75% 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 spent 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 nonpositive effect of this intervention 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.

¹⁶ The funding scheme gives no compensation for the first 9% of weighted students.

The recent studies by Angrist and Lavy (1999), Goolsbee and Guryan (2006), Rouse and Krueger (2004), and our results all point in the same direction. Instruction methods using computers are likely to be less effective than more traditional instruction methods, and than many policymakers and policy advisers believe them to be. We think that policymakers and policy advisers should take the results of the recent studies seriously. If not, they may jeopardize the skill development of future generations.

REFERENCES

- Angrist, J. D., and V. Lavy, "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics* 114:2 (1999), 533–575.
- , "New Evidence on Classroom Computers and Pupil Learning," *Economic Journal* 112 (2002), 735–765.
- Beerends, H., and S. van der Ploeg, "Onderzoek vergoeding schoolspecifieke knelpunten," Regioplan, Report OA-230 (2001).
- Bénabou, R., F. Kramarz, and C. Prost, "Zones d'Education Prioritaire: Much Ado About Nothing?" CREST-INSEE working paper (2004).
- Burtless, G., "Introduction and Summary" (pp. 1–42), in G. Burtless (Ed.), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success* (Washington, DC: The Brookings Institution, 1996).
- Card, D., and A. A. Payne, "School Finance Reform, the Distribution of School Spending, and the Distribution of Student Test Scores," *Journal of Public Economics* 83 (2002), 49–82.
- Chay, K. Y., P. J. McEwan, and M. Urquiola, "The Central Role of Noise in Evaluating Interventions that Use Test Scores to Rank Schools," *American Economic Review* 95 (2005), 1237–1258.
- Cleveland, W. S., "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association* 74: 368 (1979), 829–836.
- Goolsbee, A., and J. Guryan, "The Impact of Internet Subsidies in Public Schools," this REVIEW 88:2 (2006), 336–347.
- Greenwald, R., L. V. Hedges, and R. D. Laine, "The Effect of School Resources on Student Achievement," *Review of Educational Research* 66:3 (1996), 361–396.
- Guryan, J., "Does Money Matter? Regression-Discontinuity Estimates from Education Finance Reform in Massachusetts," NBER working paper no. 8269 (2000).
- Hahn, J., P. Todd, and W. Van der Klaauw, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica* 69:1 (2001), 201–209.
- Hanushek, E. A., "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature* 24 (1986), 1141–1177.
- , *Making Schools Work: Improving Performance and Controlling Costs* (Washington, DC: The Brookings Institution, 1994).
- , "School Resources and Student Performance," in G. Burtless (Ed.), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success* (Washington, DC: The Brookings Institution, 1996).
- Jacob, B. A., and L. Lefgren, "Remedial Education and Student Achievement: A Regression Discontinuity Analysis," this REVIEW 86:1 (2004), 226–244.

- Kirkpatrick, H., and L. Cuban, "Computers Make Kids Smarter—Right?" *TECHNOS Quarterly for Education and Technology* 7:2 (1998), 1–11.
- Leuven, E., and H. Oosterbeek, "Evaluating the Effects of a Tax Deduction on Training," *Journal of Labor Economics* 22:2 (2004), 461–488.
- Papke, L. E., "The Effects of Spending on Test Pass Rates: Evidence from Michigan," *Journal of Public Economics* 89 (2005), 821–839.
- Rouse, C. E., and A. B. Krueger, "Putting Computerized Instruction to the Test: A Randomized Evaluation of a Scientifically-Based Reading Program," *Economics of Education Review* 23 (2004), 323–338.
- Van der Klaauw, W., "Breaking the Link between Poverty and Low Student Achievement: An evaluation of Title I, *Journal of Econometrics* (forthcoming).

Appendix

Additional Results

TABLE A1.—DIFFERENCE-IN-DIFFERENCES ESTIMATES OF THE EFFECT OF "ELIGIBILITY" OF THE PERSONNEL SUBSIDY ON TEST SCORES
BASED ON ARBITRARY CUTOFF AT 50% DISADVANTAGED MINORITY PUPILS

	$\hat{\delta}_{2002}$ (s.e.)	$\hat{\delta}_{2003}$ (s.e.)	$\hat{\delta}_{Pooled}$ (s.e.)	Test $\hat{\delta}_{2002} = \hat{\delta}_{2003}$	
				χ^2	p-value
Language					
DS $\pm$ 5	0.001 (0.066)	−0.015 (0.067)	−0.006 (0.055)	0.044	0.834
DS $\pm$ 10	0.022 (0.046)	−0.081 (0.045)	−0.028 (0.039)	4.176	0.041
Arithmetics					
DS $\pm$ 5	0.039 (0.065)	0.093 (0.069)	0.065 (0.055)	0.502	0.479
DS $\pm$ 10	−0.047 (0.045)	−0.006 (0.048)	−0.027 (0.038)	0.634	0.426
Information					
DS $\pm$ 5	−0.004 (0.062)	0.036 (0.070)	0.016 (0.054)	0.282	0.596
DS $\pm$ 10	0.012 (0.043)	−0.019 (0.045)	−0.003 (0.037)	0.393	0.531

Note: Standard errors (in parentheses) take into account clustering at the school $\times$ year (class) level and are heteroskedasticity robust.

TABLE A2.—DIFFERENCE-IN-DIFFERENCES ESTIMATES OF THE EFFECT OF "ELIGIBILITY" OF THE COMPUTER SUBSIDY ON TEST SCORES
BASED ON ARBITRARY CUTOFF AT 50% DISADVANTAGED PUPILS

	$\hat{\delta}_{2002}$ (s.e.)	$\hat{\delta}_{2003}$ (s.e.)	$\hat{\delta}_{Pooled}$ (s.e.)	Test $\hat{\delta}_{2002} = \hat{\delta}_{2003}$	
				χ^2	p-value
Language					
DS $\pm$ 5	0.021 (0.032)	−0.006 (0.035)	0.007 (0.028)	0.547	0.459
DS $\pm$ 10	0.056 (0.023)	0.006 (0.024)	0.031 (0.019)	3.892	0.049
Arithmetics					
DS $\pm$ 5	−0.004 (0.033)	−0.018 (0.036)	−0.011 (0.029)	0.121	0.728
DS $\pm$ 10	0.045 (0.024)	−0.025 (0.024)	0.010 (0.020)	6.929	0.009
Information					
DS $\pm$ 5	0.015 (0.031)	0.027 (0.033)	0.021 (0.026)	0.106	0.744
DS $\pm$ 10	0.047 (0.022)	0.004 (0.022)	0.026 (0.018)	2.919	0.088

Note: Standard errors (in parentheses) take into account clustering at the school $\times$ year (class) level and are heteroskedasticity robust.

TABLE A3.—DIFFERENCE-IN-DIFFERENCES IV ESTIMATES OF THE EFFECT OF THE COMPUTER SUBSIDY ON TEST SCORES FOR GIRLS

	Pooled—Girls		Pooled—Boys	
	$\hat{\delta}$	s.e.	$\hat{\delta}$	s.e.
Language				
DS $\pm$ 5	−0.063	(0.053)	0.002	(0.054)
DS $\pm$ 10	−0.120	(0.037)	−0.038	(0.036)
Arithmetics				
DS $\pm$ 5	−0.128	(0.058)	−0.040	(0.053)
DS $\pm$ 10	−0.121	(0.040)	−0.022	(0.036)
Information				
DS $\pm$ 5	−0.025	(0.054)	0.001	(0.055)
DS $\pm$ 10	−0.077	(0.039)	0.000	(0.037)

Note: Standard errors (in parentheses) take into account clustering at the school $\times$ year (class) level and are heteroskedasticity robust.