Expanding School Resources and Increasing Time on Task:

Effects of a Policy Experiment in Israel on Student Academic Achievement and Behavior*

August 2012

Victor Lavy

The Hebrew University of Jerusalem, University of Warwick and NBER

Abstract

In this paper, I examine how student academic achievements and behavior were affected by a school finance policy experiment undertaken in elementary schools in Israel. Begun in 2004, the funding formula changed from a budget set per class to a budget set per student, with more weight given to students from lower socioeconomic and lower educational backgrounds. The experiment altered teaching budgets, the length of the school week, and the allocation of time devoted to core subjects. The results suggest that spending more money and spending more time at school and on key tasks all lead to increasing academic achievements with no behavioral costs. I find that the overall budget per class has positive and significant effects on students' average test scores and that this effect is symmetric and identical for schools that gained or lost resources due to the funding reform. Separate estimations of the effect of increasing the length of the school week and the subject-specific instructional time per week also show positive and significant effects on math, science, and English test scores. However, no cross effects of additional instructional time across subjects emerge, suggesting that the effect of overall weekly school instruction time on test scores reflects only the effect of additional instructional time in these particular subjects. As a robustness check of the validity of the identification strategy, I also use an alternative method that exploits variation in the instruction time of different subjects. Remarkably, this alternative identification strategy yields almost identical results to the results obtained based on the school funding reform. Additional results suggest that the effect on test scores is similar for boys and girls but it is much larger for pupils from low socioeconomic backgrounds and it is also more pronounced in schools populated with students from homogenous socioeconomic backgrounds. The evidence also shows that a longer school week increases the time that students spend on homework without reducing social and school satisfaction and without increasing school violence.

msvictor@huji.ac.il
v.lavy@warwick.ac.uk

JEL No. I21,J18,J24

* Special thanks go to Adi Shani and Meir Amit for outstanding research assistance. I benefited from comments and suggestions by Arthur Blouin, Karen Brandon, Michael Freedman, Yona Rubinstein, Tali Regev, Analia Schlosser, and seminar participants at Hebrew University, LSE, National University of Singapore, PSE, Tel Aviv University, and University of Warwick.

I. Introduction

Questions about how students' academic achievements are affected by school resources, such as the amount of time they spend in school and the amount of time they devote to specific subjects, are of compelling interest for policymakers. Research on these questions are important since marginally increasing instructional time is relatively simple to do, and the potential to make such changes would seem to be possible in many countries as international statistics on the annual number of school days and on the distribution of weekly instructional time across subjects reveal large differences among and within countries.[1] For example, one of the main educational strategies of the "No Excuses" charter schools in New York, Boston and other places in the US is to emphasize the importance of increased instructional time (Dobbie and Fryer 2011).[2] In addition, evidence on these questions is valuable for improving the efficiency of school resource allocation, in particular the instruction budgets of schools across various subjects and activities.

While an extensive literature exists regarding the effects of school resources on student outcomes, much of the evidence is inconclusive. [3] For example, regarding the impact of school resources, Hanushek (2006) notes in an extensive survey that "the evidence – whether from aggregate school

[1] The OECD Program for International Student Assessment (PISA) 2003, 2006, and 2009 data reveal the great extent of these differences among the more than 60 countries that participated in this project. The reaction to these differences suggests that political leaders do pay attention to the results. For example, President Barack Obama cited the gap between the length of the school year in the United States and other countries as one reason he advocates the expansion of U.S. schools' instruction time among his key educational policy goals (March 10, 2009, at a speech to the U.S. Hispanic Chamber of Commerce).

[2] The so-called "No Excuses" schools are loosely defined as schools that emphasize strict discipline, extended time in school, and an intensive focus on building basic reading and math skills. In fact, Thernstrom and Thernstrom (2004) and Whitman (2008) argue that the "No Excuses" schools are more effective due to more instructional time, a zero tolerance disciplinary code, high academic expectations for all students, and an emphasis on teaching basic math and reading skills. However, these impressions are based on correlative analysis which does not permit causal interpretation.

[3] For instance, using TIMSS data and citing few recent studies, Wößmann (2003) reports that international differences in pupil test scores (in mathematics and science) are not caused by differences in school resources. In a French program that allocates extra financial resources to schools in disadvantaged zones, Benabou et al. (2004) find that this allocation has no significant impact on student outcomes. Häkkinen et al. (2003) uses the dramatic changes in the school spending caused by the 1990s recession in Finland and finds that changes in teaching expenditure did not have a significant effect on the test scores. Hanusheck (2003) report that inputs based schooling policies in the US failed to improve students' test scores. Lavy (2002) finds that providing resources to high schools led to significant gains in test scores and lower dropout rates. Focusing on labor market outcomes, Card and Krueger (1996) find positive effects of school resources on earnings, whereas Heckman et al. (1996) do not find significant effects, and Betts (2001) finds mixed results.

outcomes, econometric investigations, or a variety of experimental or quasi-experimental approaches – suggests that pure resource policies that do not change incentives are unlikely to be effective. Importantly, the results appear similar across both developed and developing countries." In addition, specific evidence on time use in schools is limited largely to the effect of the number of school days per year on student's achievements.[4] Thus, this evidence does not tell us much about the effect of the number of hours children spend in school every week; or the time they spend in specific activities, such as math or reading; or whether there is a complementarity or a spillover effect across subjects.

In this paper, I investigate the causal effect of school resources – the total teaching budget per class, the number of hours that children spend weekly in school and the amount of time devoted to core subjects – on pupils' achievements. I exploit a unique school finance policy experiment that changed the formula used to determine the teaching budget of primary schools in Israel. Until 2003, schools were funded based on the number of classes, irrespective of class size. In September 2004, the funding rules were changed and from then on schools received funding per student enrolled where a deprivation index was used to determine the amount of each "student voucher," with more money channeled toward students from the lowest economic and educational backgrounds.[5] This experimental reform generated a sharp and exogenous change in the teaching budget of many schools. Some schools gained resources while other schools experienced no change or even a decline in resources. Naturally, schools with a high proportion of students from a deprived background or with large classes were the main beneficiaries from this reform.

[4] For example, Grogger (1996), and Eide and Showalter (1998) estimated the effect of the length of the school year in the United States and found insignificant effects, perhaps due to selection and omitted variables. Card and Krueger (1992) and Betts and Johnson (1998) used state-level data in the United States to examine the same effect and found positive and significant effects on earnings which converge to zero once school quality is added as a control in the regression. Pischke (2007) used a natural experiment in West Germany and found that a shorter school year increased grade repetition and lowered enrollment in higher school tracks, but it had no effect on earnings and employment later in life. Based on school day cancelations due to snow, Hansen (2008) reports that more instructional time increases student performance, and Marcotte and Hemelt (2008) find that years with substantial snowfall are associated with lower pupil performance.

[5] In 2009, the funding rules were changed again to a system very similar to that used prior to September 2004.

For estimation, I use data from 2002-2005 for fifth-grade pupils for all the schools in the country. I observe each school twice, either in 2002 and 2004 or in 2003 and in 2005. The key feature of these data is that they present an opportunity to observe each school under the two different funding systems. This allows one to estimate the resulting changes in the instructional time budget per class, the length of the school week, and weekly instructional time by subject.

I employ two main identification strategies in the paper. The first identification strategy compares schools in the years before and after the reform. The advantage of this strategy is that it guarantees that no other fundamental changes occurred in schools during this period. As a robustness check, I also use a difference in differences estimation based on restricting my sample to schools that either gained or experienced no change in resources due to the reform, and secondly a restricted sample of schools that either lost or experienced no change in resources due the reform. As an additional identification strategy, I use an alternative method that exploits variation in the instruction time of different subjects. The consistency of results obtained from these different identification strategies strengthens the causal interpretation of the estimates reported in this paper.

In this paper, I consider the impact of several key explanatory variables – total teaching budget per class, the length of the school week, and the effect of instructional time – on the average scores in math, science and English. I compare the various estimates, show that they are consistent, and discuss the implications for the quality of identification of each of these estimates. I also evaluate the effect of instructional time on students' homework time allocation, on their overall satisfaction with school, their social satisfaction in class, on their involvement in violent behavior, and on their fear of school bullying. These topics offer an indication whether social factors and student behavior are affected by the length of time that students spend in school during the day and week.

My results show that additional instructional budget per class, a longer school week, and additional instructional time spent on the different subjects have positive and significant effects on students' academic achievements. These estimates are in contrast to the "naïve" OLS estimates which are actually negative, reflecting a negative selection pattern of allocating higher instructional budgets to

potentially low achieving schools (low SES). This is the first paper to show that the biased estimated effect of school resources and instructional time are reversed from negative to positive once potential selection and endogeneity of school resources are fully accounted for.[6] The first differences and difference in differences estimates of the effect of the instruction budget per class, the length of the school week, and of instructional time in each subject are mutually consistent and yield very similar elasticities of test scores with respect to any of the three measures of instruction time at school. The estimates show that the boost in test scores is modest. For example, increasing instructional time in math, science, and English by one hour per week increases the average test scores in these subjects by a 0.053 standard deviation. The respective difference in differences estimate of this effect are 0.073 and 0.059, revealing symmetry in the effect of changing instruction time of the core subject. Estimating the effect separately for each subject yields an effect size of 0.041 in math, 0.043 in science and 0.056 in English. The average result of these three estimates is 0.047, only marginally lower than the estimated average result. Allowing for treatment heterogeneity, the growth in test scores is similar for boys and girls and is larger for pupils from low socioeconomic status. Overall, the main results presented in the paper are very robust to a variety of robustness checks with respect to their identification assumptions and to threats to their validity. Further, the alternative identification strategy based on pupil fixed effect and on variation in time of instruction across subjects yields surprisingly identical results: the estimated test score growth from increasing average weekly time of instruction in math, science and English by one hour per week is 0.058, almost identical to the respective estimate of 0.053 obtained from the natural experiment based on the funding policy reform. It is remarkable that the two identification methods which are based on different assumptions yield such strikingly similar results. These estimates are also identical to estimates that I obtained using PISA data of all OECD countries and another sample of all East European countries where the identification strategy was based on pupil fixed effect and variation in time of instruction across subjects (Lavy 2012). The estimate of the effect of hours of instruction on math, science and

[6] Angrist and Lavy (1999) show a similar reversal in sign with respect to the effect of class size. The "naïve" OLS estimates of class size were actually positive, reflecting a negative selection pattern of allocating smaller class size to potentially low achieving schools (low SES).

English in the OECD sample, is 0.058, and in the east European sample it is 0.059. Remarkably, these two estimates are identical to the respective estimate based on the data from Israel.

The funding reform was also used to study how additional time spent on instruction affects several behavioral outcomes of students. The results suggest that increasing instruction time in different subjects also increases modestly the time students spend doing homework. Furthermore, even with the additional time spent at school, students' overall satisfaction from school and from its social environment was unaffected. In particular, increasing the length of the school week did not have any effect on these behavioral outcomes; resulting in no change in students' satisfaction from school and its social environment, on their violent behavior, or on their fear of bullying. This is the first paper to show evidence of school resources and instruction time on behavioral outcomes of students.

The rest of the paper is organized as follows. Section II presents background on the reform. Section III presents our data and section IV presents our empirical strategy. Section V presents the evidence about the effects of school budgets on test scores and behavioral outcomes. Section VI presents the conclusions.

II. The 2004 Funding Reform

The budget for primary schools in Israel comes from two sources: the Ministry of Education and the local municipal authority. The Ministry of Education funding is provided according to the number of instruction hours, which is measured in units of one hour of instruction per week for the whole school year. This budget funds all teaching instruction costs, as well as the cost of internal (within school) and external (outside of school) teachers' training. The local authority funds all the administrative costs of the school such as the costs of secretaries, school supplies, and building maintenance. In the 2004 school year, the Ministry of Education introduced a school finance reform as a policy experiment that changed the formula used to determine the instruction time budget of primary schools in Israel. Until 2003, schools were funded based on the number of classes, irrespective of class size. Schools also received an additional instruction budget based on a per school deprivation index which was a weighted function of

the students' average parental education, family size, family income and number of immigrant students. Schools with a higher deprivation index received more resources. The overall instructional budget of this differential component amounted to about 10 percent of the overall funding of primary schools in the country. In September 2004, the funding rules were changed and from then on schools received funding per student enrolled. A deprivation index was used to determine the amount of each “student voucher,” with more resources channeled toward students from lower economic and educational backgrounds. The differential deprivation index is a per-student index calculated according to a needs-based formula (a larger budget is allocated to needy students according to the depth of their needs), with an added “national priority” element (a larger budget is allocated to students living in areas that were defined as 'national priority' areas, such as those near Israel’s borders). The elements that form this index and their weights are as follows: mother's years of schooling (15%), father's years of schooling (15%), number of siblings (10%), new immigrant status (20%) and immigrant from developing countries status (10%), national priority status (20%), and periphery location status (schools located far from the three largest cities in Israel) (10%). Most of the weights were derived from a variance decomposition regression that examines the correlation between students' background characteristics and students' achievement.

This experimental reform generated a sharp and exogenous change in the teaching budget of many schools. Schools with a large enrollment of students with a high deprivation index and schools with large classes gained resources, while others schools lost resources. The reform was intended to produce no changes in the overall resource distribution among schools in the Jewish public school system. However, the reform was designed to allow for an increase of about 15 percent to 20 percent in the overall budget of the Arab schooling sector, as this sector includes a much higher share of students with high deprivation index estimates and has larger classes. The schools that gained resources had on average lower pre-reform budgets per class while those who lost resources had higher pre-reform budgets per class. This experimental reform lasted until 2008 when it was changed again to a new system that was more similar to the pre-2004 rules.

While there are several inputs that may be affected by a change in a school's budget, I find that the budget per class is highly correlated with the length of the school week and with instructional time of math, science, and English and not with other inputs such as class size or extracurricular activities. For instance, the regression coefficient of budget per class on length of the school week is 0.280 (se=0.017) and it is very similar for schools that gained resources (0.242, se=0.023) and for schools that lost resources (0.324, se=0.019) due to the reform. The regression coefficient of budget per class on instructional time of math, science and English is 0.085 (se=0.009) and it is almost identical for schools that gained resources (0.087, se=0.012) and for schools that lost resources (0.097, se=0.012) due to the reform. [7] These relationships seem stable and yield the same estimates when estimated separately based on the pre and post reform samples. At the same time, the estimated coefficient of the budget per class on class size is -0.011 (se=0.034) suggesting that this is not a channel that schools used for spending their teaching budget. This conclusion is also evident when comparing the estimated effect of the budget per class on weekly hours of instruction or on core subjects instructional time obtained from two sub-samples stratified by actual class size or predicted class size (based on maximum class size of 40, see Angrist and Lavy, 1999). For example, the estimated coefficient of the school budget on length of the school week is 0.269 in the sample of above the mean predicted class size and it is 0.276 in the sample of schools with below the mean predicted class size.

III. Data

The data I use in this study are based on the Growth and Effectiveness Measures for Schools (GEMS - *Meizav* in Hebrew) datasets for the years 2002-2005. The GEMS includes a series of tests and questionnaires administered by the Division of Evaluation and Measurement of the Ministry of Education.[8] The GEMS is administered towards the end (from mid-May to mid-June) of each school year

[7] These results are not presented in the paper and are available from the author.

[8] The GEMS is not administered for school accountability purposes and only aggregated results at the district level are published. For more information on the GEMS see the Division of Evaluation and Measurement website (in Hebrew): http://cms.education.gov.il/educationcms/units/rama/odotrama/odot.htm.

to a representative 1-in-2 sample of all elementary and middle schools in Israel, so that each school participates in GEMS once every two years. The GEMS data include test scores of fifth- (primary school) and eighth- (middle school) grade students in math, science, Hebrew, and English. In principle, all students except those in special education classes are tested and the proportion of students tested is above 90 percent. The raw test scores used a 1-to-100 scale that I transform into z-scores to facilitate interpretation of the results. In this study I use only primary school data since the funding reform only affected primary level schools.

The test scores for the years 2002-2005 are linked to student administrative records collected by the Israel Ministry of Education. The administrative records include student demographics that I use to construct all measures of students' background characteristics. Using the linked datasets, I build a panel for elementary schools with test scores for the years 2002-2005. The sample is restricted to Jewish public schools (excluding Arab and religious Orthodox Jewish schools). There are 939 elementary schools with test score data. Since every school is sampled once in two years, I have two observations of the same school for more than 90 percent of the schools.

The GEMS also includes interviews with all teachers and the school principal. The questionnaire for home teachers of all classes included questions about classroom instructional time in each subject and the total instructional time per week. I use teachers' responses to these items to compute the school average for fifth-grade instructional time in each subject. Though there was very little difference between or among fifth-grade classes in a school in these time inputs, I still prefer to use the school-level mean per grade to avoid any biases that might be caused by sorting of students into certain classrooms and setting time allocations for given academic subjects according to those students' particular strengths and weaknesses. In any case, the grade- and class-level measures of these time inputs are very highly correlated.

I also use items from the GEMS student questionnaire that address various aspects of the school and their learning environment. I concentrate on two sections of the questionnaire: the first provides information on student satisfaction in school and on the violent behaviour of other students and the

second provides data on student allocation of time for homework by subject. In the first section students are asked to rate the extent to which they agree with a series of statements on a six-point scale ranging from *"strongly disagree"* to *"strongly agree"*. These items include: (1) "There are many fights among students in my classroom"; (2) "Sometimes I'm scared to go to school because there are violent students"; (3) "I am often involved in violent activities in school"; (4) "I feel well-adjusted socially in my class"; and (5) "I am satisfied in school". I transformed students' responses to these items into standardized z-scores. In the second section of the questionnaire, students are asked to report the number of hours per week that they spend at home doing homework in each of three subjects (math, science and English).

In Table 1, I present summary statistics for the variables used in the analysis. Column 1 lists the results for our key variables in the pre-reform period of 2002-2003 and column 2 lists the results for the post-reform period of 2004-2005. Panel A presents the results for the budget per class and instructional time variable, all measured in terms of weekly hours of instruction. According to the table, the mean budget per class is the same in both periods, suggesting that the reform had no impact on the distribution of resources among the Jewish secular schools. The length of the school week is on average 35 hours, implying that 76 percent of the teaching budget of schools is used for classroom instruction. The rest of the teaching budget is used to fund teachers' training, to pay personnel for extracurricular activities in school, and after school remedial education programs. The average instructional time of the three core subjects of math, science and English is 14 hours a week, over two-fifths of it used for math instruction and the rest divided almost equally between the other two subjects. Overall, there seems to be little difference in instructional time in the years before and after the reform. Panel B presents the means for the average test scores and also for each subject. Panel C presents the means for school characteristics, which are almost identical over the two periods respectively.

IV. Empirical Strategy

The effects of unobserved correlated factors usually confound the effect of school budget or instructional time on student outcomes. Such correlations could result if self-selection and sorting of students across schools are affected by school resources, or if there is a correlation between school instructional time and other characteristics of the school that may affect student outcomes. The structure of the GEMS allows me to use an identification strategy that overcomes this potential problem because it is based on observing schools and their students at two points of time: before the funding reform (in 2002 or 2003) and after the funding reform (in 2004 or 2005). I take advantage of this feature and construct a longitudinal dataset at the school level to examine how changes in students' achievements are associated with changes in instructional time. Note that the change in instructional budget can only be due to the funding reform because there is no school choice at the primary schooling level in any school districts in Israel and assignment to schools is based on pre-determined rules (mainly the family location of residence). As a result, the potential for selection bias due to sorting of students across schools based on instruction budget or time is very small in this context.

To develop the relationships of interest using the panel data, I first specify the following standard education production function that links pupils' achievements and their relevant determinants:

$$Y_{ij0} = \alpha_j + \gamma\, W_{j0} + \beta\, X_{ij0} + \delta\, S_{j0} + u_{ij0} \qquad (1)$$

where Y_{ij0} is the average achievement of the ith student in math, science and English, in the jth school in period zero (pre-reform), W_{j0} is the total budget of instructional time per class in the jth school in pre-reform period. X_{ij0} is a vector of characteristics of the ith student, S_{j0} is a vector of time varying characteristics of the jth school, α_j is a school fixed effect that captures everything about the school that is not observed and does not vary between the two years that each school is observed (2002 and 2004 or 2003 and 2005) and u_{ij0} is the unobserved error term. Observing schools in more than one time period allows expanding equation (1) to the post reform period in the following equation:

$$Y_{ij1} = \alpha_j + \gamma\, W_{j1} + \beta\, X_{ij1} + \delta\, S_{j1} + u_{ij1} \qquad (2)$$

where *1* denotes the post reform period. Differencing equations (1) and (2) yields a first difference equation that is equivalent to the following school fixed effect model that we can estimate with a panel data on schools:

$$Y_{ijt} = \alpha_j + \gamma\, W_{jt} + \beta X_{ijt} + \delta S_{jt} + u_{ijt} \qquad (3)$$

where t denotes the time period. In this model, the identifying assumption is that W_{jt} could have changed between the two periods for each school only because of the change in the funding rules and that the average value of the X's (which are used to compute the deprivation index) remained unchanged. Therefore, conditional on a school fixed effect, the change in W_{jt} is not correlated with the change in u_{ijt}.

Equation (3) can also be used for difference in differences estimation once we take advantage of the unique feature of the funding reform that benefited some schools who gained resources while it harmed other schools that lost resources. Exploiting this feature of the reform, I run two sets of difference in differences estimations, which are variations of equation (3). In the first estimation, I restrict the sample to include only schools that either gained resources (treatment group) or had no change in resources (control group). In the second estimation, I restrict the sample to include only schools that either lost resources (treatment group) or had no change in resources (control group). Beyond providing a robustness check to the identification strategy that I use, the estimates from these two distinct sets of difference in differences estimation can also shed light on the very interesting and policy-relevant question of whether gaining or losing resources has a symmetric effect on test scores.

V. Empirical Results

A. Main Results

Table 2 presents our baseline results on the relationship between class budgets and student achievements in math, science, and English. The table estimates equation (3) with varying degrees of control variables. The estimates presented in column 1 are from OLS regressions which include only subject and year fixed effects as controls. The estimates presented in column 2 are from regressions that

include also school fixed effects. Column 3 also controls for student characteristics and column 4 controls for time varying school characteristics. In Panel A, I report estimates from a regression in which the instructional budget per class measured in weekly hours is the treatment variable. The mean of instructional budget per class is 46.6 (sd=5.99). The OLS estimate in column 1 is negative (-0.015) and significant (sd=0.002) which means that school resources and test scores are negatively correlated. This is most likely a biased estimate since schools with lower potential outcomes receive compensatory resources. The bias could also result from omitted variables that are correlated with student performance. However adding the school fixed effect reverses the sign of the estimate to be positive (0.007) and statistically significant (sd=0.003). This estimate is unchanged in the other two specifications (columns 3-4). This suggests that conditional on the school fixed effects, the instructional budget per class is not correlated with student and time varying school characteristics such as enrollment. This confirms the identification assumption that the school characteristics used in the budget formula have not changed during the two years between the pre- and post-funding reform. Therefore, we can be confident that the change in the school instructional budget reflects only the change in the weights of these characteristics in the funding rules.

In Panels B and C, I present the two sets of the difference in differences estimates. For these estimates I divide the sample into three groups: schools that gained resources following the reform, schools that had fewer resources following the reform, and schools that had no change in resources. [9] Note that dividing the sample according to the extent of change in school resources is appropriate since this variable is exogenous to potential outcomes of students (conditional on school fixed effects). It is also important to note that the schools that experienced no change in resources between the two periods serve as the comparison group and schools that experienced a change in resources in the second period

[9] Since there were few schools who experienced no change in resources, I expanded this category to include schools who experienced less than a +/- 2 percent change in resources. After this change, the mean percentage change in instructional hours per class for schools that experienced 'no change' was zero percent and the standard deviation was 1 percent. I also define schools that gained resources as schools that experienced a budget increase of more than 2 percent, and schools that lost resources as schools that experienced a budget decrease of more than 2 percent. Significantly, I found that the results are not sensitive to widening this range to -/+3 or narrowing it to -/+1 percent.

(either a gain or a loss) are the treatment group.[10] Accordingly, in schools that gained resources, the mean percentage change in instructional hours per class is 8.4 percent and the standard deviation is 10.2 percent. Furthermore, in schools that lost resources, the mean percentage change in instructional hours per class is -6 percent and the standard deviation is 4.2 percent.

According to Table 2, the respective estimates from the two sets of difference in differences estimation are remarkably similar. The simple OLS estimate obtained from the 'increase' sample in Panel B is -0.013 while the respective estimate in Panel C for the 'decrease' sample is -0.017. Similarly, the estimated effects in column 4 are 0.005 in the 'increase' sample and 0.006 in the 'decrease' sample. This indicates that the estimates in Panels B and C are very similar to the respective estimates presented in Panel A, though these are much more precisely estimated. This similarity demonstrates not only that school resources have a positive effect on test scores, but that this effect is also fully symmetric in terms of an increase or a decrease in resources.

It should be noted that the results of Table 2 estimate the effects of the reform only one or two years after its implementation. Therefore, a valid question is whether the changes we observe in schools and the estimates of the effect of school resources are representative of a longer run effect. Two pieces of evidence suggest that the estimates in Table 2 do reflect longer term adjustments. First, estimating the effect of school resources separately based on the contrast of 2002-2004 and 2003- 2005 yield almost exactly the same estimates, suggesting that the estimated effect based on experiencing one or two years of reform is the same. Second, the results of my alternative identification strategy which is based on cross section data analysis and reflects long term estimates (See Table 6) are identical to those presented in Table 2.

Another possible concern is whether the results from Table 2 are biased due to the convergence of underachieving schools towards the level of high performing schools. In other words, if schools with a lower than average budget per class (who benefited more from the funding reform and presumably had

[10] A similar approach for difference in differences estimation is applied in Duflo (2001) where regions in Indonesia that experienced very low rate of school construction were defined as control areas while regions that had many new constructed schools were used as the treatment group.

lower average test scores in the pre-reform period) had a higher improvement rate of test scores due to a convergence effect, such convergence would be positively correlated with the resource gain from the funding reform and therefore will bias upward the estimated effect of the budget per class on test scores. To check the extent of this possible bias, I divided the sample into two groups based on budget per class in the pre-reform period and re-estimated equation (3) in each of these samples separately. The estimated effect of budget per class on the average test score obtained from the sample of schools with above average budget per student is 0.008 (se=0.006). The respective estimate obtained from the sample of schools with below mean budget per student is 0.006 (se=0.003). Stratifying the sample into four groups based on budget per student yields a similar pattern. This evidence suggests that it is very unlikely that the resource effect that we estimated reflects test score convergence. To examine this potential threat further, I replicated this estimation by stratifying the sample based on the average test score in the first period. The estimated effects from the first and third quartiles are identical (0.007, sd=0.004), and those obtained from the second and fourth quartiles are lower. Even though it is not correct statistically to stratify the sample based on an endogenous variable (school average test score), these estimates also suggest no difference in the estimated effect of budget per class in high- and low-achieving schools.[11] In the next section, I estimate some particular channels through which school instructional resources affect student performance, in particular the length of the school week and classroom study time of core subjects.

B. Identification and Estimation of Time on Task

A school's instructional budget is largely spent on the length of the school week. Thus, I estimate equation (3) by replacing the instruction budget per class with the length of weekly school instruction (in terms of hours per week). The weekly instructional time is divided among different subjects. For this reason, I also estimate equation (3) where the sum of weekly hours of instruction of the three core subjects (math, science and English) is the treatment measure. One possible problem with this approach

[11] Results are available upon request.

is that these two measures of instructional time are choice variables and, therefore, could be endogenous in equation (3). If the choice made by schools of how much to allocate from the instructional budget to any of these two measures is a only a function of fixed characteristics of the school, then the school fixed effect model will identify the causal effect of any of these two treatment measures. However, if these choice decisions are correlated with the error term in equation (3), our results might be biased. I discuss this issue in greater detail in my Table 3 results.

In the first row of Table 3, I estimate the effects of the length of the school week (number of weekly hours in school) on the average test score. The mean number of weekly hours in school is 35.0 (sd=3.2). Similar to our Table 2 results, the OLS estimate is initially negative (-0.02) and statistically significant, and becomes positive (0.007) once schools fixed effects are added to the regression. In addition, the estimates in columns 2-4 are nearly identical, implying that adding the student and school characteristics as controls has no effect on the estimates and their standard errors. The estimated effect is 0.008, and it is statistically significant (se=0.004).

The estimated effect of the length of the school week and that of the budget per class can be compared based on the elasticity of the average test score. The elasticity of the instructional budget per class is 0.080, and the elasticity of the length of the school week is also 0.079.[12] This implies that the instructional budget per class has an effect on test scores mainly through the increase in length of the school week. A validation of this result is also shown in the second row of Table 3, where I present estimates of the effect of the budget of weekly hours of instruction beyond the length of the school week (simply the difference between the instructional budget per class and the length of the school week) on the average test score. The mean of this measure is 11 weekly hours and its estimated effect presented in column 4 is 0.003 (sd=0.002). Its elasticity with respect to the average test scores is 0.01, confirming that

[12] These estimates are presented in Table A1 and they show exactly the same pattern that is shown in Tables 2 and 3. Since the mean of the standardized test score is zero, I compute the two elasticities based on estimates of equation (3) where the dependent variable is the actual grade (scale 1 to 100) instead of the z score. The elasticity of the budget per class is computed as [0.127 x (46/70)] while the elasticity of the length of the school week is computed as [0.157 x (35/70)], both equal to 0.079.

the effect of class budget on average test scores of the three subjects, beyond what is allocated to the length of the school week, is indeed very small.

Another important implication of the similarity of these two estimated effects is with regard to the interpretation of the estimated effect of the length of the school week as causal. Given that the change in the instructional budget per class is exogenous, conditional on schools' fixed effects, its estimated effect is clearly unbiased. Therefore, the similarity in the two point estimates and in their implied elasticities is suggestive evidence that the estimated effect of the length of the school week is unlikely to be biased due to selection or endogeneity. A related point is that if the effect of length of the school week were biased, upward or downward, then the effect of the difference between the instructional budget per class and the length of the school week should have been biased in the opposite direction. Instead, we find that this estimate is practically zero.

In the third row of Table 3, I present estimates of the effect of the average weekly hours of instruction in math, science, and English. This average is equal to 4.6 hours per week (sd=1.70). The OLS estimate in column 1 is positive and significant. However adding the school fixed effects to the estimated equation almost double the estimated coefficients, from 0.029 to 0.055. Remarkably, however, the latter estimate remains unchanged as I add controls to the school fixed-effect regressions: the point estimate in the second column is 0.055 (sd=0.023) and is 0.053 (se=0.023) in columns 3 and 4. This robust estimate implies that adding one hour of instruction in each of the three subjects raises the average score by 0.053 standard deviations.

Remarkably, this result is very similar to the estimates that Dobbie and Fryer (2011) obtain from their sample of charter schools in New York City (NYC). They find that schools that add 25 percent or more instructional time have an annual gain that is 0.059 of a standard deviation higher in math. Note that a one hour increase in instruction time in our sample is approximately 25 percent (given that the respective mean is 4.6 hours) and our estimated effect is 0.053, almost identical to the NYC estimate. However, the authors emphasize that their estimates of the relationship between school inputs – including instructional time – and school effectiveness are unlikely to be causal given the lack of

experimental variation in school inputs. However, in a recent study of public schools in Houston, Texas Fryer (2012) reports similar size effects of instructional time: schools that add 25 percent or more instructional time compared to traditional public schools have annual gains that are 0.084 sd higher in math and 0.043 sd higher in English, and these results are based on controlled data. Moreover, the estimate of the effect of instructional time obtained from a sample of over 50 countries in Lavy (2010) is exactly equal to the estimate obtained in this paper and to the effect size presented in Dobbie and Fryer (2011). In the concluding section of the paper, I will discuss further this apparent 'empirical regularity' in the relationship between instructional time and test scores.

Furthermore, this estimate of 0.053 yields an elasticity of 0.21 which is almost identical to the elasticity of the length of the school week after we adjust for the difference in the means of the two instructional time measures.[13] This result has two important implications. The first is that other time that children spend in school during the week in pursuits outside of math, science, and English classes, does not affect at all their achievement in these subjects. In other words, the effect of the length of the school week on average test scores is only a reflection of its correlation with the instructional time of these particular subjects. The implication is that whatever skills students acquire during the time in school spent outside of math, science, and English classes (60 percent of their total school time) are immaterial to their academic progress in these three core subjects, at least as reflected in the short-term math, science and English test scores. Perhaps we should not be surprised that knowledge in other subjects, such as history, geography and literature, is irrelevant for better achievement in math or science. However, students may acquire and enhance non-cognitive skills, such as socialization, confidence and determination, during longer school weeks. Thus, one might have expected potential spillover effects to surface for a wide array of academic pursuits.

The similarity of the estimated effects of the length of the school week and of the instructional time in math, science and English has a second important implication: The effect of instructional time in

[13] The estimated effect of the average instruction time of these core subjects on the row test score is presented in the fourth row of appendix Table A1. The elasticity of this time measure is computed as [1.05 x (13.7/70)].

math, science and English is very unlikely to be biased. If the change in instructional time of these subjects between 2002/03 and 2004/05 were determined selectively with respect to potential outcomes in these subjects, we would have expected that the estimated effect of the length of the school week and of weekly hours of instruction of these subjects to be different. Instead, they are almost identical. A validation of this result is shown in the fourth row of Table 3, which presents estimates of the effect on the average test score of the number of weekly hours of instruction in all other subjects and activities in school. This measure of instructional time is simply the difference between the length of the school week and the instructional time of math, science and English, and its mean is 22 weekly hours. Remarkably, the point estimates in columns 2-4 are practically zero. This result confirms that there are no spillover effects in school in Israel from instruction of all other subjects on achievements in math, science and English. It also confirms that conditional on school fixed effects, the allocation of instructional time to math, science and English, given the length of the school week, is not correlated with potential outcomes in these subjects. Finally, if the estimated effect of instructional time of the core subject were biased, upward or downward, we should have expected that the effect of all other instructional time during the week to be biased. However, I do not observe such bias as the estimated effect of instruction on non-core subjects on average test scores for core subjects is nearly zero.

In Table 4, I rerun the results of Table 3 where the sample is stratified into schools who gained from the reform and schools who lost from the reform. Overall, the estimates from the two subsamples are very similar. For example, the estimated effect of the weekly instruction hours spent on the core subjects (presented in the third row of Table 4) is 0.073 in the 'increase' sample and 0.059 in the 'decrease' sample. It is quite remarkable that increasing the teaching time of these subjects by one hour lead to a test score gain that is almost the same as the decline in test scores due to reducing instructional time of these subjects by one hour. I also view this similarity as another indication that these estimates are not biased since it is very unlikely that the selection in the reaction to a decline in resources will lead to a bias which will be the same as the bias induced by an endogenous reaction to an increase in resources.

C. Estimated effect of instructional time in each subject

In this section, I specify and estimate a school fixed-effect model where both the dependent variable and the time of instruction per week are subject-specific, as follows:

$$Y_{kijt} = \mu_j + \rho\, T_{kjt} + \eta\, X_{ijt} + \varepsilon\, S_{jt} + u_{kijt} \qquad (4)$$

where Y_{kijt} is the achievement in the kth subject, in the jth school, of the ith student, and T_{kjt} is instructional time in the kth subject in the jth school. The unobserved error term u_{kijt} is now subject's specific. In addition, I estimate an alternative specification that also includes as treatments the hours of instruction in each of the other two subjects and the total instructional time of all other subjects:

$$Y_{kijt} = \mu_j + \rho\, T_{kjt} + \eta\, X_{ijt} + \varepsilon\, S_j + \lambda\, T_{2jt} + \theta\, T_{3jt} + \sigma\, T_{ojt} + u_{kijt} \qquad (5)$$

where T_{2jt} and T_{3jt} represent instructional time in the other two subjects and T_{ojt} is instructional time in all other subjects. λ, θ, and σ are the cross subjects parameters. Note that the sum of T_k, T_2, T_3, and $T_{o'}$ is equal to the length of the school week in terms of hours of instruction. By comparing the estimates obtained using equation (5) to those obtained based on using equation (3) and overall instructional time of the three subjects, I hope to strengthen the causal interpretation of the evidence.[14]

In Table 5, I present the results of estimating equations (4) and (5) for each of the three subjects separately. Each parameter presented in the table is estimated in a separate regression. Each regression includes as controls school fixed effects, year dummies, student characteristics, and school characteristics. The three estimates presented in the first row are positive and precisely measured. The effect of an hour of instruction in math is 0.041 (sd=0.018), in science it is 0.043 (sd=0.016), and in English it is higher, 0.056 (sd=0.020). The average of these three estimates is 0.048, only slightly lower than the estimated average effect (0.053) reported in the third column of Table 2. Based on the close similarity of these two estimates, this implies that the distribution of the overall teaching time of all three subjects to each subject is not correlated with potential outcome. Overall, I conclude that the estimates

[14] I also compare the estimates of equation (4) to respective parameters based on a completely different identification methodology (See Table 6), and I will argue that the similarity in the estimates across methods strengthen their causal interpretation.

based on separate regression for each subject are fully consistent with the estimates obtained where the dependent variable is the average test scores of all three subjects and the treatment measure is the average instructional time in these subjects.

In the second row, I present the results of estimating equation (5) where each regression includes all three subjects' specific weekly hours of instruction as well as the sum of instructional time in all other subjects. The set of four estimates presented in each column is obtained from one regression. The table indicates that the estimates of the effect of instructional time of each subject on the same subject test score are very similar to the respective estimates reported in the first specification, though they are marginally higher for all three subjects. The biggest gap is in English, for which the estimate increases by 7 percent, from 0.056 to 0.063. The table also indicates that the cross effects results are all positive, though very small and not significantly different from zero. For example, the effects of math and science instructional time on English test scores are, not surprisingly, practically zero (0.007 and 0.001). The largest cross effect is that of math instructional time on science test score, 0.035. However, this effect is measured very imprecisely (sd=0.027) and therefore not statistically significant from zero. In comparison, the reciprocal cross effect of science instructional time on math is not important as this estimate is 0.011 (sd=0.013). Interestingly, the table also confirms our earlier finding that instructional time of non-core subjects (representing 60 percent of the length of the school week) has no effect on achievements in any of the three core subjects: the estimates in the fourth row of the second panel of Table 5 are all positive but very small, and they are not significantly different from zero. This result is also consistent with the evidence presented in Table 3.

Overall, these findings strengthen our previous results, as I did not find any cross effects within subjects, and between each of these subjects and instructional time of non-core subjects. This can be viewed as additional suggestive evidence that the estimated effect of instructional time of each of the subjects is not biased due to selection or endogenous determination of these educational inputs (that is not accounted for by our natural experiment and our school fixed effect difference in differences framework).

D. An alternative identification strategy: between subjects variation in instructional time

In this section, I present additional estimates to those presented above that are based on an alternative identification method that can account for potential confounding factors in the estimation of instructional time. Here I rely on within-student variations in instructional time across various subjects of study to examine whether differences in student performance in three subjects are systematically associated with differences between subjects in instructional time. The basic identification strategy is that student characteristics and the school environment are the same for all three subjects except for the fact that some subjects receive more instructional time. It important to emphasize that the pupil fixed-effect identification method proposed here does not exploit any variation in instructional time due to the funding reform. I use the cross-section variation since and observe students when they are exposed only to one regime of funding. Based on this approach I present within student estimates of the effect of instructional time on individual test scores using the following panel data specification,

$$Y_{kij} = \mu_i + \gamma\, T_{kj} + \beta\, X_{ij} + \delta\, S_j + (\varepsilon_j + \eta_k) + u_{ijk} \qquad (6)$$

Where Y_{kij} is the achievement in the kth subject of the ith student in the jth school, T_{kj} is instructional time in the kth subject in the jth school, X is a vector of characteristics of the ith student in the jth school and S_j is a vector of characteristics of the jth school. ε_j and η_k represents the unobserved characteristics of the school and the subject, respectively, and u_{kij} is the remaining unobserved error term. The student fixed effect μ_i captures the individual's family background, underlying ability, motivation, and other constant non-cognitive skills. Of course, a specification that includes μ_i will not include the term βX_{ij}. Note, that by controlling for this individual fixed effect and using within-student across subjects' variation in test scores, I also control for the school fixed effect ε_j. Therefore, exploiting within-student variation allows one to control for a number of sources of potential biases related to unobserved characteristics of the school, the student, or their interaction. One potential source of bias is that students might be placed or be sorted according to their ability across schools that provide more (or less) instructional time in some subjects. For example, if more talented students attend better schools that provide more instructional

hours overall in each subject, it would cause γ to be biased downward unless the effect of student and school fixed effects are accounted for. Similarly, the bias will have an opposite sign if the less talented students are exposed to more instructional time. Identification of the effect of instructional time based on a comparison of the performance of the same student in different subjects is therefore immune to biases due to omitted school level characteristics, such as resources, peer composition and so on, or to omitted individual background characteristics, such as parental schooling and income. Equation (6) can be estimated with a single year's cross-sectional data for each school, or it can be estimated using two years of data for each school, the latter allows including in the model a school fixed effect in addition to the pupil fixed effect.

This identification strategy is also subject to several key assumptions. First, a necessary assumption for this identification strategy is that the effect of instructional time is the same for all subjects, implying that γ cannot vary by subject. This restriction seems plausible as the first identification method that I used estimated the treatment effect for each subject separately and all three estimates were relatively similar. Second, the effect of instructional time is "net" of instructional time spillovers across subjects. This assumption is also supported by the evidence presented in Table 3 which showed that there are no significant cross-subject effects of hours of instruction. Third, this identification strategy does not preclude the possibility that pupils select or are sorted across schools partly based on subject-specific instructional time. For example, the results would be biased if students who have a high ability for math may select or be placed in a school that specializes in math and has more instructional time in math. However, I believe that this concern is not relevant for three reasons. First, the pupils in the sample are in fifth grade of primary school in Israel where admission is based on neighborhood school zones without any school choice. Second, primary schools that specialize in a given subject are very rare in Israel. Third, tracking within schools is not allowed in primary and middle schools in Israel, and Ministry circulars reiterate this issue frequently. Even if some schools overlook this regulation and practice tracking within school, this is not a major concern as I measure instructional time in each subject by the school-level means and not by the class means or even the within school program-level means.

Therefore, omitted subject-specific student ability will not be correlated with subject instructional time in a given school.

Table 6 presents the estimates based on the within pupil estimation strategy.[15] Two different specifications are used. The regression results reported in the first column include year fixed effects, pupil demographic controls, and school characteristics. The regression reported in the second column includes also pupils fixed effects. In the first row I present estimates based on pooling all three subjects. The two estimates are positive and significantly different from zero. The estimate in the second column is lower than the estimate in the first column, suggesting that the OLS estimates are slightly upwardly biased. Significantly, the estimated effect when pooling the three subjects is 0.058 (sd=0.007), which is very similar to the estimated average effect of 0.053 reported in the third row of Table 3. This is a remarkable outcome since the two estimates are obtained from two very different identification strategies. As noted, in Table 3 the identifying assumption is to compare results within schools for two adjacent fifth-grade cohorts where nothing has changed except the funding rules, and, therefore the first difference estimation at the *school level* is appropriate. In Table 6, the identifying assumption is that conditional on pupil fixed effects, hours of instruction of each subject are not correlated with the potential outcome (the error term in equation (5)), and, therefore, estimation based on differences of all variables from pupil level means is accounting for all potential omitted variables. The similarity in the estimated average treatment effect of instructional time clearly contributes to the credibility of the interpretation of the estimates in Tables 2-4 as causal. This similarity also suggests that the short-term evidence presented in Table 3 is close to the estimated longer run effects presented in Table 6.

In the other rows of Table 6, I present estimates based on pooling two of three subjects. This is possible since all that is needed for this identification strategy is at least two or more observations per student. Remarkably, all three estimates in column 2 are very similar and range from 0.055-0.060,

[15] Since the treatment variable instructional time is measured at the school level the error term $u_{ijk,}$ is clustered by school to capture common unobservable shocks to students at the same school.

providing additional proof that the effects of instructional time on each of these subjects are not very different. In addition, the average of the three separate effects is 0.058 which identical to the estimate in the first row that pools all three subjects.

E. Heterogeneity in effect of instructional time

In this section, I consider the marginal productivity of instructional time in each of the three subjects. This is an important policy question as this parameter can be useful for allocating instructional resources across subjects and within schools. In Table 7, I report estimates from regressions where the continuous instructional time measure was converted to dummy indicators. Columns 1, 4, and 7 present the range of hours of instruction for each of the subjects. These ranges vary by subject. Math instructional time ranges from three to 10 hours a week, so it is measured by four such indicators as follows: five hours or less (mean=4.92), six hours, seven hours, eight hours or more (mean=8.20). Science instructional time ranges from one to seven hours a week, and, therefore, it is measured by four indicators as follows: two hours or less (mean=1.95), three hours, four hours, and five hours or more (mean=5.33). English instructional time ranges from two to eight hours a week, and it is also measured by four such indicators: three hours or less (mean=2.94), four hours, five hours, six hours or more (mean=6.12). Columns 2, 5, and 8 present the proportion of students in each range, which again varies by subject. For example, the lowest range in the three subjects includes similar proportions of students (from 15 percent in science and English to 19 percent in math). The highest range includes less than 5 percent of the students in math and English but almost 18 percent of the students in science.

The results reported in the Table 7 are based on a specification with student and school characteristics and year fixed effects. Columns 3, 6, and 9 are separate regressions where the treatments included in each regression are the respective subject indicators of hours of instruction. In column 3, a comparison of the marginal changes in productivity of an hour of math instruction suggests a moderate positive non-linearity: the gain from a sixth hour is 0.047, the seventh hour adds 0.080 and the eighth hour adds 0.090. The results are similar in columns 6 and 9 for science and English respectively. This suggests evidence of a slight increase in marginal productivity as hours of instruction increase.

In Table 8, I examine heterogeneity of the treatment results reported in Table 3. I present treatment effect estimates by gender, family education, and by the degree of heterogeneity in student ability in the classroom. Columns 1-2 of Table 8 show that the effect of additional instructional time for math is the same for boys and girls, but the effect of increasing science and English instructional time is much larger for boys.

In columns 3-4, I present heterogeneous treatment effect by parental education. The sample is divided by the median value of the sum of one's father's and mother's years of schooling (a proxy for socioeconomic background). Interestingly, the effect of math instruction is much larger among children from families with low levels of parent education. The gap is more than 200 percent (0.055 versus 0.023), suggesting that targeting of additional math instructional time to students from lower socioeconomic backgrounds will yield much higher returns. The average gap between the groups is about half a standard deviation, so adding two to three hours of math instruction per week to the lower socioeconomic group should help narrow this gap. The effect of science hours of instruction is also higher for students from lower socioeconomic backgrounds, but this gap is not very large. However, for English instruction, the results are reversed: the respective effect on English achievements is larger for the sample of students from families with high levels of education, though this difference is small.[16]

The productivity of school instruction may also vary by the heterogeneity in students' ability in the classroom. Since parental schooling is highly correlated with student ability, I measured class level heterogeneity by the standard deviation of the classroom distribution of fathers' schooling. In columns 5-6 of Table 8, I present the effect of instructional time in each of the subjects for two sub-samples. The first, denoted as "heterogeneous," includes classes where the standard deviation of the father's level of schooling is above the median for the sample of classes. The second, denoted as "homogenous," includes classes below the median. The productivity of instructional time is higher in homogenous classes in all

[16] Another interesting result is that the effect of math instruction on science achievement is much larger for students from backgrounds with low levels of parent education. Results are available upon request.

three subjects, with the largest difference being in math, but these differences are relatively small and not significantly different from zero.

F. Effect of school instructional time on pupils' homework effort and satisfaction in school

In order to fully assess the overall benefit of an extended school week program or of interventions that add instructional time to some subjects, it is important to consider other important questions. For instance, do students who "enjoy" a longer school week study harder at home and spend more time doing homework? Are they more satisfied in school? Are they better off socially in class? Do they feel more secured in school and get less involved in violence and bulling?

In Table 9, I estimate the effects of classroom instructional time in each of the subjects on homework time allocation. For each subject I report results from three different regression specifications. In all three I include school fixed effects and student and school characteristics. In the first specification (columns 1, 4, and 7) the only treatment in the regression is the respective subject instructional time. In the second specification (columns 2, 5, and 8) I add the effects of the instructional time in each of the other two subjects. In the third specification (columns 3, 6, and 9) I also add the instructional time in all other subjects. Both before and after the reforms, students spent approximately 3.2 hours per week doing math homework, 2.5 hours doing science homework, and 3 hours doing English homework.

The evidence presented in Table 9 suggests that homework time in each of the core subjects increases slightly with the subject's increased instructional time in school. This effect is significant for math and English, but it is only marginally significant for science. In addition, the effect sizes are relatively small. For example, an increase of an hour of school instruction in math or English leads to an average increase of four to five minutes of homework. Considering that students are engaged in homework for 2 to 3 hours per subject, the changes in students' time allocation in each of the three subjects are marginal. The table also indicates that these estimated effects are not sensitive to adding the measures of instructional time in other subjects. For example, when math instruction is the only included treatment, its effect on math homework time is 0.044 (sd=0.021), and it is unchanged when instructional

time of science and English are added to the regression. However, this estimate decreases to 0.041 when the regression includes instructional time for all other subjects.[17]

These results imply that added school instructional time for any subject does not crowd out the time that students invest at home for the study of other subjects; to the contrary, it even marginally increases the overall time spent on homework. The implication is that an increase in instructional time at school leads to a net expansion of overall study time at home and at school together. Significantly, the additional time spent on homework is most likely a mechanism for the effect of increased school instructional time on test scores.

An additional question linked to the school reforms concerns whether additional time spent in school would come at the expense of overall social and school satisfaction. Thus, it is important to assess and measure these effects for a more general equilibrium evaluation of the effect of extending the length of the school week. In Table 10, I present estimates of the effect of increased school instructional time per week on five behavioral outcomes: personal violence in school, mean level of classroom violence, personal fear of violence in school, satisfaction from school, and social satisfaction in school. The regression specification includes school fixed effects and student and school characteristics. Column 1 reports estimates based on the full sample. In columns 2-3, I stratify the sample by gender. In columns 4-5, I present the results for the high and low parental education samples, respectively, and in columns 6-7 for the samples of homogenous and heterogeneous schools samples. The table indicates that there is no systematic pattern of an effect of the length of the school week on any of the five behavioral measures, and none of the five reported estimates in the first column is significantly different from zero.[18] The

[17] I also estimate the effect of instructional time in each of the core subjects on the probability that some students receive additional instruction from privately funded tutors at home. Private tutors work with 14 percent of the students in math, 5 percent of students in science, and 27 percent of students in English. Additional school instructional time in math and science has positive effect on the propensity of getting private tutoring but these effects are small and insignificant for all three subjects. The estimate for math instructional time is 0.003 (se=0.003), for science 0.002 (se=0.001) and for English precisely zero. Results are available upon request.

[18] It should be noted that there are some results that are significant such as the negative estimated effect for boys on school satisfaction. In addition, three other statistically significant parameter estimates in Table 10 are: the negative effect of the length of the school week on classroom violence in heterogeneous school, the positive effect on

conclusion, therefore, is that extending the school week carries no negative repercussions in terms of satisfaction of students from school, the class social environment, or levels of school violence.

VI. Conclusions

In this paper, I estimate empirically the effect of school resources and various measures of increased instructional time in school on students' academic performance and behavior. In particular, I analyze the effects of the instructional budget per class, the length of the school week, and the instructional time of math, science and English on test scores in these subjects and on homework time of students. I also assess students' school satisfaction, social acclimatization, violence, and fear of bullying. I take advantage of a policy reform in Israel that began in 2004 which changed the rules of funding public primary schools. The system changed from funding based on number of classes to a system based on number of students, weighted to take into account their average socioeconomic status. The results based on fifth-grade students' test scores clearly indicate that school resources and the length of the school week have a positive and significant effect on pupils' performance in core subjects. Importantly, the effects of increasing or decreasing school resources or the length of the school week are fully symmetric, as they are identical in absolute terms though opposite in sign. The evidence also consistently shows that increasing the amount of instructional time of math, science, or English positively impacts test scores in that subject. Here as well, the effect of increasing or decreasing instructional time in the core subjects is symmetric. An alternative identification strategy based on a pupil fixed-effect estimation that exploits variation of time of instruction in different subjects yields exactly the same results. The evidence also strongly suggests that increased achievement in a given subject is the result of increased time on that particular arena of study alone. Results show that cross effects – the benefits of additional instructional time in one subject upon the achievements in another – are negligible. There is no spillover

---

personal violence in homogenous schools, and the negative effect on school satisfaction in homogenous schools. These however are very small effects and therefore not economically meaningful.

effect on math and English from increased instructional time in other subjects, and only modest increases in science, with the slight benefit stemming from increased math instructional time. In addition, there are no apparent cross effects from instructional time spent engaging in all other subjects such as history, geography, literature, and social studies on math, science and English test scores. The evidence also suggests some heterogeneity in the effect of instructional time. For example, increased instructional time in science and English leads to more dramatic academic achievement growth for boys, while additional math instructional time results in more pronounced academic achievement growth for girls. Pupils from families with low levels of parent education show higher gains in achievement from additional classroom instruction in math and science. Increasing instructional time in math, science or English actually leads to an increase in the time at home that students spend on homework. Finally, the evidence suggests that expanding the school week does not diminish the school and social satisfaction of students.

This is the first paper that provides such detailed evidence on the causal effects of the school instructional resources, the length of the school week, and of instructional time in different subjects on students' academic performance and on important behavioral outcomes. The results are based on a sample that includes all the primary schools in Israel (with the exception of Arab and religious Orthodox Jewish schools). In this sample the means of the length of the school week and the time of instruction in math, science and English in Israel are very similar to the respective means of the OECD countries as observed in the PISA data from its various rounds in the previous decade. These two aspects provide an external validity appeal to the evidence presented here because they are relevant to many countries that seek ways to improve their education system. The evidence presented in the paper can also serve as a good benchmark for evaluating the effect and the cost-benefit of many "traditional" school interventions such as reducing class size, increasing teachers' training, and tracking student by ability. As well, it can also serve as a benchmark for evaluating more recent popular "progressive" interventions in schools such as pay for performance for teachers or for students, or using computer added instructions in the classroom.

References

Angrist J. and V. Lavy. (1999). "Using Maimonides' Rule to Estimate the Effect of Class Size on Children's Academic Achievement." *Quarterly Journal of Economics*, May 1999, 533-579.

Benabou, R., Kramarz, F. and Prost, C. (2009). "The French Zones d'Education Prioritaire: Much Ado about Nothing." *Economics of Education Review*, 28(3), pp. 345-356.

Betts, J. R. and Johnson, E. (1998). "A Test of Diminishing Returns to School Spending." mimeograph, University of California San Diego.

Betts, J. R. (2001). "The impact of school resources on women's earnings and educational attainment: findings from the National Longitudinal Survey of young women." *Journal of Labor Economics*, 19(3), pp. 635–57.

Card, D. and Krueger, A. (1992). "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy*, 100, pp. 1-40.

Card, D. and Krueger, A. B. (1996). 'School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina', *Journal of Economic Perspectives*, 10(4), pp. 31–50.

Dobbie, Will and Roland G. Fryer, Jr. (2011) "Getting Beneath the Veil of Effective Schools: Evidence from New York City", Harvard University.

Duflo, Esther (2001). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review*, 91(4), pp. 795-813.

Eide, E. and Showalter, M.H. (1998). "The Effect of School Quality on Student Performance: A Quantile Regression Approach," *Economics Letters*, 58, pp. 345-50.

Fryer, Roland G., Jr. (2012). "Injecting Successful Charter School Strategies into Traditional Public Schools: Early Results from an Experiment in Houston, January 2012 Code to Replicate Results." Harvard University.

Grogger, J. (1996). "Does School Quality Explain the Recent Black/White Wage Trend?" *Journal of Labor Economics*, 14, pp. 231-53.

Häkkinen, I.,, T. Kirjavainen and R. Uusitalo (2003). "School resources and student achievement revisited: new evidence from panel data", *Economics of Education Review*, 22(3), pp. 329–335.

Hanushek. E. A (2003) "The Failure of Input Based Schooling Policies", *Economic Journal* 113: F64-98.

Hanushek, E. A (2006). "School Resources", In Eric A. Hanushek and Finis Welch (Eds.), *Handbook of the Economics of Education, Volume 2*, pp. 865-908.

Hansen, Ben. (2008). "School Year Length and Student Performance: Quasi-Experimental Evidence", University of California Santa Barbara.

Heckman, J., Layne-Farrar, A. and Todd, P. (1996). "Does Measured School Quality Really Matter? An Examination of the Earnings-Quality Relationship", in Burtless, G. (ed.), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, Washington, DC: Brookings Institution Press.

Lazear, E., (2001) "Educational Production", *Quarterly Journal of Economics* 116: 777-803.

Lavy, Victor. (2002). "Evaluating the Effect of Teachers' Group Performance Incentives on Pupils' Achievements", *Journal of Political Economy*, 1286-1317.

Lavy, Victor, "Do Differences in School's Instruction Time Explain International Achievement Gaps in Math, Science, and Reading? Evidence from Developed and Developing Countries." NBER Working Paper w16227, revised version. 2012.

Marcotte, Dave E. and Steven Hemelt (2008). "Unscheduled Closings and Student Performance." *Education Finance and Policy*, 3(3), pp. 316-38.

Pischke, J. S. (2007) "The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years," *Economic Journal*, 117(523), pp. 1216-1242.

Thernstrom, Abigail, and Stephan Thernstrom. (2004). "No Excuses: Closing the Racial Gap in Learning." Simon & Schuster.

Whitman, David. (2008). "Sweating the Small Stuff: Inner-City Schools and the New Paternalism." Thomas B. Fordham Institute.

Wößmann, L. (2003). "Schooling resources, educational institutions and pupil performance: the international evidence". *Oxf. Bull. Econ. Statist.*, 65, pp. 117–170.

Table 1: Descriptive Statistics: Instruction Time, Test Scores, and School Characteristics

	Years	
	2002-2003	2004-2005
A. Instruction time		
Instruction budget per class (in weekly hours)	46.63 (5.99)	45.69 (5.86)
Length of the school week (in weekly hours)	35.06 (3.21)	34.76 (3.14)
Weekly instruction hours of math, science, and English	13.68 (1.67)	13.88 (1.72)
Weekly instruction hours in:		
Math	6.00 (0.79)	6.12 (0.81)
Science	3.60 (1.09)	3.60 (1.05)
English	4.07 (0.69)	4.15 (0.70)
B. Test scores		
Average test scores (math, science, and English)	70.93 (15.42)	75.14 (14.47)
Math	72.34 (19.16)	71.97 (16.96)
Science	65.74 (17.79)	77.14 (15.86)
English	73.23 (20.73)	75.00 (21.58)

Notes: Standard deviations are presented in parentheses. The sample includes all the Jewish secular and religious state schools. This sample includes over 60 percent of the schools and students in the country.

Table 1: Descriptive statistics: Instruction Time, Test Scores, Schools Characteristics

	Years	
	2002-2003	2004-2005
C. School Characteristics		
Enrollment	441.65	440.78
	(154.07)	(154.53)
Class size	28.11	27.96
	(4.10)	(4.04)
Religious school	0.22	0.22
	(0.41)	(0.41)
Number of schools	920	927
Number of students	53,981	55,633

Table 2: Estimated Effect of School Instruction Budget per Class on the Average Score in Math, Science, and English

	The Controls Included in the Regression			
	Year Control Only	School Fixed Effect	School Fixed Effect and Student Characteristics	School Fixed Effect, Student, and School Characteristics
	(1)	(2)	(3)	(4)
Panel A:				
Full Sample	-0.015	0.007	0.007	0.007
	(0.002)	(0.003)	(0.003)	(0.003)
Number of Schools	932	932	932	932
Number of Students	88,495	88,495	87,903	87,903
Panel B:				
"Increase" Sample	-0.013	0.003	0.005	0.005
	(0.002)	(0.004)	(0.004)	(0.004)
Number of Schools	447	447	447	447
Number of Students	42,331	42,331	42,033	42,033
Panel C:				
"Decrease" Sample	-0.017	0.006	0.006	0.006
	(0.002)	(0.005)	(0.005)	(0.005)
Number of Schools	673	673	673	673
Number of Students	64,652	64,652	64,269	64,269

Notes: Standard errors are presented in parentheses and are clustered at the school level. Each parameter presented in columns (1)-(4) is from a different regression. All specifications include year fixed effects. Student characteristics include: gender dummy, both parents' years of schooling, number of siblings, immigration status indicators, and ethnic origin indicators. School characteristics include: student enrollment and student enrollment squared.

Table 3: Estimated Effect of School Instruction Time on the Average Score in Math, Science, and English

	The Controls Included in the Regression			
Measures of Instruction Time	Year Control Only	School Fixed Effect	School Fixed Effect and Student Characteristics	School Fixed Effect, Student and School Characteristics
Subject	All Subjects	All Subjects	All Subjects	All Subjects
	(1)	(2)	(3)	(4)
Length of the school week (in weekly hours)	-0.020 (0.003)	0.007 (0.004)	0.008 (0.004)	0.008 (0.004)
Difference between instruction budget per class and length of school week	-0.012 (0.002)	0.004 (0.002)	0.003 (0.002)	0.003 (0.002)
Average weekly instruction hours of math, science and English	0.029 (0.017)	0.055 (0.023)	0.053 (0.023)	0.053 (0.023)
Weekly instruction hours of all other subjects	-0.023 (0.003)	-0.003 (0.005)	-0.001 (0.005)	-0.001 (0.005)
Number of Schools	932	932	932	932
Number of Students	88,495	88,495	87,903	87,903

Notes: See Table 2.

Table 4: Estimated Effect of School Instruction Time on the Average Score in Math, Science, and English

	The Controls Included in the Regression					
	School Fixed Effect	School Fixed Effect and Student Characteristics	School Fixed Effect, Student and School Characteristics	School Fixed Effect	School Fixed Effect and Student Characteristics	School Fixed Effect, Student and School Characteristics
	"Increase" Sample			"Decrease Sample"		
	(1)	(2)	(3)	(4)	(5)	(6)
Length of the school week (in weekly hours)	0.004	0.005	0.005	0.006	0.007	0.007
	(0.008)	(0.008)	(0.008)	(0.006)	(0.006)	(0.006)
Difference between instruction budget per class and length of school week	0.001	0.001	0.002	0.001	0.000	0.000
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Average weekly instruction hours of math, science and English	0.076	0.073	0.073	0.060	0.059	0.059
	(0.034)	(0.034)	(0.034)	(0.027)	(0.027)	(0.027)
Number of Schools	447	447	447	673	673	673
Number of Students	42,331	42,331	42,033	64,652	64,652	64,269

Notes: See Table 2.

Table 5: Estimated Effect of School Instruction Time by Subject on Test Score

	Test Score		
	Math	Science	English
	(1)	(2)	(3)
Specification I:			
Only own subject's hours of instruction is included as treatment	0.041	0.043	0.056
	(0.018)	(0.016)	(0.020)
Specification II:			
All three subjects' hours of instruction and total of other subjects' hours of instruction are included as treatments			
Math instruction hours	0.043	0.035	0.007
	(0.019)	(0.027)	(0.023)
Science instruction hours	0.011	0.046	0.001
	(0.013)	(0.016)	(0.014)
English instruction hours	0.017	0.009	0.063
	(0.021)	(0.021)	(0.021)
Other subjects total weekly instruction hours	0.003	0.005	0.010
	(0.006)	(0.007)	(0.007)
Number of Schools	932	932	932
Number of Students	100,834	99,842	99,395

Note: See Table 2. In Specification II, the set of four estimates presented in each column is obtained from one regression.

Table 6: Estimated Effect of School Instruction Time on Test Score Based on Within Pupil Regressions

Subject Combination	OLS Regression with Pupil and School Characteristics (1)	Regression with Pupil Fixed Effects (2)
Math + Science + English	0.069	0.058
	(0.007)	(0.007)
Math + Science	0.074	0.055
	(0.007)	(0.010)
Math + English	0.080	0.060
	(0.011)	(0.016)
Science + English	0.056	0.059
	(0.008)	(0.012)
Number of schools	933	933
Number of students	166,630	167,726

Notes: See Table 2. These regressions assume that the effect of instruction time is the same for all subjects.

Table 7: Estimated Non-Linear Effect of School Instruction Time by Subject on Test Score

Test Scores								
Math			Science			English		
Grouping Hours	Proportion in Sample	Estimate	Grouping Hours	Proportion in Sample	Estimate	Grouping Hours	Proportion in Sample	Estimate
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
5 hours or less (mean=4.92)	19.07%	-	2 hours or less (mean=1.95)	15.01%	-	3 hours or less (mean=2.94)	14.52%	-
6 hours	59.70%	0.047 (0.027)	3 hours	31.62%	-0.001 (0.037)	4 hours	62.25%	0.068 (0.035)
7 hours	15.78%	0.080 (0.041)	4 hours	35.80%	0.066 (0.042)	5 hours	19.73%	0.079 (0.045)
8 hours or more (mean=8.20)	5.45%	0.090 (0.064)	5 hours or more (mean=5.33)	17.57%	0.105 (0.055)	6 hours or more (mean=6.12)	3.50%	0.130 (0.071)

Note: See Table 2. All specifications include year fixed effects, student and school characteristics.

Table 8: Heterogeneity in Estimated Effect of School Instruction Time by Subject on Test Score: by Gender, Family Education and School Homogeneity

	Gender		Family Education		School Homogenity	
	Boys	Girls	High Education	Low Education	Heterogeneous Schools	Homogenous Schools
	(1)	(2)	(3)	(4)	(5)	(6)
Math instruction hours	0.042	0.046	0.023	0.055	0.034	0.049
	(0.020)	(0.021)	(0.019)	(0.022)	(0.028)	(0.028)
Science instruction hours	0.054	0.030	0.037	0.046	0.035	0.046
	(0.017)	(0.017)	(0.017)	(0.018)	(0.033)	(0.022)
English instruction hours	0.069	0.042	0.063	0.049	0.035	0.044
	(0.022)	(0.025)	(0.019)	(0.026)	(0.029)	(0.039)
Number of schools	910	913	927	932	608	596
Number of students	50,273	50,561	46,932	53,911	50,325	50,509

Note: See Table 2. All specifications include year fixed effects, student and school characteristics.

Table 9: Estimated Effect of School Instruction Time by Subject on Homework Hours

	Homework Hours								
	Math (mean=3.19, sd=1.55)			Science (mean=2.48, sd=1.55)			English (mean=3.00, sd=1.60)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Math instruction hours	0.044	0.044	0.041	-	0.003	0.005	-	0.016	0.012
	(0.021)	(0.021)	(0.022)		(0.021)	(0.022)		(0.022)	(0.023)
Science instruction hours	-	-0.009	-0.012	0.021	0.022	0.024	-	-0.005	-0.011
		(0.013)	(0.014)	(0.014)	(0.014)	(0.015)		(0.015)	(0.015)
English instruction hours	-	-0.005	-0.008	-	-0.002	0.000	0.048	0.048	0.044
		(0.018)	(0.019)		(0.018)	(0.019)	(0.018)	(0.018)	(0.019)
Other subjects total weekly instruction hours	-	-	-0.005	-	-	0.004	-	-	-0.008
			(0.007)			(0.007)			(0.008)
Number of schools		932			932			932	
Number of students		97,174			97,057			97,042	

Note: See Table 2. All specifications include year fixed effects, student and school characteristics.

Table 10: Estimated Effect of Length of School Week on Violence and Student Satisfaction.

		Sample					
		Gender		Family's Education		School Homogenity	
	Full Sample	Boys	Girls	High Education	Low Education	Hetrogenous Schools	Homogenous Schools
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Personal Violence	0.000	-0.007	0.006	-0.004	0.004	-0.004	0.011
(mean=1.97, sd=1.43)	(0.004)	(0.006)	(0.004)	(0.004)	(0.006)	(0.008)	(0.006)
Class Violence	-0.006	-0.009	-0.005	0.002	-0.009	-0.029	0.001
(mean=3.62, sd=1.54)	(0.007)	(0.008)	(0.008)	(0.007)	(0.008)	(0.014)	(0.010)
Fear from Violence at School	-0.005	-0.004	-0.006	-0.005	-0.005	-0.004	-0.001
(mean=2.02, sd=1.55)	(0.004)	(0.005)	(0.006)	(0.005)	(0.005)	(0.008)	(0.007)
School Satisfaction	-0.005	-0.015	0.004	-0.005	-0.006	0.003	-0.017
(mean=5.18, sd=1.25)	(0.004)	(0.006)	(0.005)	(0.006)	(0.006)	(0.009)	(0.007)
Social Satisfaction	0.000	0.000	-0.001	-0.003	0.001	0.003	-0.007
(mean=5.16, sd=1.24)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.008)	(0.007)
Number of schools	932	910	913	928	932	609	598
Number of students	97,664	48,323	48,812	45,321	51,814	48,290	48,845

Notes: See table 2. All specifications include year fixed effects, student and school characteristics. Each parameter presented in the table is from a different regression. Violence and the student satisfaction raw variables range between 1 (lowest) to 6 (highest). The estimates in columns (1)-(3) are based on these variables' z scores.

Table A1: Estimated Effect of School Instruction Time on the Average Score in Math, Sciences and English (non-standartized scores)

	The Controls Included in the Regression						
Measures of Instruction Time	Year Control Only	School Fixed Effect	School Fixed Effect and Student's Characteristic	School Fixed Effect, Student's and school's Characteristics			
Subject	All Subjects	All Subjects	All Subjects	All Subjects	Math	Science	English
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Instruction budget per class (in weekly hours)	-0.269	0.119	0.119	0.127	0.070	0.129	0.182
	(0.030)	(0.050)	(0.050)	(0.052)	(0.064)	(0.061)	(0.080)
Length of the school week (in weekly hours)	-0.378	0.135	0.156	0.157	0.142	0.142	0.187
	(0.055)	(0.094)	(0.092)	(0.092)	(0.117)	(0.109)	(0.142)
Difference between instruction budget per class and length of school week	-0.218	0.060	0.055	0.060	0.019	0.065	0.096
	(0.034)	(0.044)	(0.044)	(0.046)	(0.058)	(0.050)	(0.072)
Average weekly instruction hours of math, science and English	0.538	1.083	1.055	1.050	0.990	1.552	0.608
	(0.324)	(0.431)	(0.429)	(0.430)	(0.475)	(0.711)	(0.611)
Weekly hours instruction of all other subjects	-0.441	-0.045	-0.021	-0.019	-0.024	-0.110	0.075
	(0.055)	(0.095)	(0.094)	(0.094)	(0.112)	(0.118)	(0.142)
Number of schools	932	932	932	932	932	932	932
Number of students	88,495	88,495	87,903	87,903	87,903	87,903	87,903

Notes: Standard errors are presented in parentheses and are clustered at the school level. Each parameter presented in columns (1)-(7) is from a different regression. All specifications include year fixed effects. Student characteristics include: gender dummy, both parents' years of schooling, number of siblings, immigration status indicators and ethnic origin indicators. School characteristics include: enrollment and enrollment square.