Tournament incentives for teachers: the case of Chile.[*]

Tomás Rau[†]
Dante Contreras[‡]

April 2011

Abstract

In this paper we evaluate the introduction in Chile of monetary incentives for teachers, based on a school performance tournament. This is particularly relevant since this is the only scaled-up incentive program for teachers in the world. We evaluate the incentive effect, i.e. the effect of introducing the incentive scheme on all participant schools, both winning and losing. We use double robust methods and panel data estimation. We explore the heterogeneous impact of the treatment through the distribution of winning. The results indicate a positive and significant tournament effect specially for schools that are *on the money*.

JEL Classification: I21,I28
Keywords: Education, teacher incentives, tournaments, government policy.

---

[*] We are indebted to the useful comments of David Card, Patrick Kline, Raj Chetty, Marcela Perticará, Peter Orazem, Francisco Gallego, two anonymous referees, and seminar participants at UC Berkeley, Universidad de Chile, U Alberto Hurtado, PUCP, the annual meeting of the Chilean Economic Association (SECHI), and the T. Paul Schultz Festschrift. We also thank the funding granted by Iniciativa Científica Milenio "Centro de Microdatos", Proyecto P07S-023-F. This research used as a source of information the SIMCE databases, Ministry of Education of Chile. The authors thank the Ministry of Education the access to that information. All the results of the study, however, are our own and in no way committed to this institution.

[†] Pontificia Universidad Católica de Chile. Vicuña Mackenna 4860, Santiago, Chile, email: trau@uc.cl

[‡] Departamento de Economía and Centro de Microdatos, Universidad de Chile.

1 Introduction

The provision of education is a topic that has received a great deal of attention in recent years. Many studies have shown the importance of education as a source of increased earnings. In Latin America, the evidence indicates that education can help to reduce income inequality, alleviate poverty and increase social mobility (World Bank (2006)). Thus, policies to improve education constitute a key area of improvement in achieving sustainable economic growth and social development. In recent decades, debates on how to improve access to education and education quality have been intense and controversial.

Since 1990, Chile has significantly increased its educational expenditure. Public education expenditure measured as a percentage of GDP increased from 2.6% in 1990 to 4.3% in the year 2000. As a consequence, profound and widespread reforms of the school system have been implemented, including decentralization, demand subsidies, standardized evaluations such as the *Sistema de Medición de la Calidad de la Educación* (SIMCE test), an increase in educational quality and equity improvement programs, educational programs targeted to the poorest schools and the extension of the school day. However, the empirical evidence for evaluating such programs is still limited.[1] Since 1996, the Ministry of Education has incorporated a monetary-based productivity bonus called The National System of School Performance Assessment (SNED).[2]This is a rank-order tournament directed towards all municipal and private subsidized schools in the country, which represent 90% of enrolled students. This is particularly relevant since this is the only one (to the best of our knowledge) scaled-up incentive program for teachers in the world.

This program seeks to improve teacher performance (productivity) via a monetary incentive (bonus). This incentive is allocated at the school level and awarded to teachers mainly on the basis of the pupils' results on the SIMCE. The program is a competitive system in which schools with similar external characteristics are grouped into homogenous school groups. The competition takes place within each distinct group. Thus, the SNED is a group incentive program in which schools compete on the basis of their average performance and monetary rewards are distributed equally among all teachers in the winning schools.

There is a usual theoretical model to explain the relationship between teacher incentives and educational performance. Tournaments may change the incentive structure of teachers and the competition may lead to more motivated teachers, improved quality of education, and hence an increase in participant schools' mean test scores.

The SNED program has some of the optimality properties described in the theoretical

---

[1] Bellei (2009) evaluates the extension of the school day

[2] See Mizala and Romaguera (2002)

work of Barlevy and Neal (2009). The authors propose an incentive pay scheme that links educator compensations to the ranks of their students within appropriately defined comparison sets. Consistently with Barlevy and Neal (2009), in the SNED program teachers in the same schools do not compete against each other. The competition scheme provides incentives for effective cooperation. In addition, in this program, teachers compete with teachers working in similar schools (guaranteed by the homogenous groups' definition).

Although performance-related pay for teachers is being introduced in many developed and developing countries, still there is significant debate on the outcomes on educational performance. On the one hand, advocates of teacher incentives programs argue in favor of strengthen weak incentives. Harbinson and Hanushek (1992), Hanushek, Kain, and Rivkin (1999) argue that teacher are mostly paid based on educational attainment, training and experience rather than performance. On the other hand, opponents argue teacher's tasks are multidimensional and only some aspects are measured by test scores. Linking compensation to test scores could cause teachers to sacrifice curiosity and creative thinking. In addition, it may lead to different ways of corruption such as teaching to the test.

Glewwe, Ilias, and Kremer (2010) discuss teacher incentives schemes in the context of an randomized experiment in Kenya, where some local schools committees strengthen teacher incentives by providing bonuses to teachers whose students performs well on exams. The authors found students in the treatment schools were more likely to take exams and scored higher, at least on some exams. However, they found little evidence of more teacher effort aimed at preventing dropouts or increasing long run learning.

Lavy (2002) examines two programs in Israel to evaluate the impact of teacher incentives on students performance. The first program is based on direct monetary rewards for teachers (incentives program) and the second is based on resources for the school (resources program). The results suggest that teacher's monetary incentives had some effect in the first year of implementation (mainly in religious schools), and it caused significant gains in many dimensions of student's outcomes in the second year (in religious and secular schools alike). However, endowing schools with more resources, also led to significant improvement in student performance. The comparison, based on cost and effectiveness, suggests that the teacher incentive program is much more cost-effective.

Muralidharan and Sundararaman (2009) investigate the effects of teacher incentives on student's achievement in India. They present results from the Andhra Pradesh Randomized Evaluation Study (AP) that considered two alternative approaches to improving primary education. The first was to provide schools with additional "smart inputs" that were believed to be more cost-effective than the status quo, and the second was to provide performance-based bonuses to teachers on the basis of the average improvement in test scores of their students. To address this, the authors designed and conducted an experiment with four different treatments. The experiment consisted in randomly allocating the programs across a representative sample of 500 government-run schools in rural AP with 100 schools in each of the four treatment groups and 100 control schools serving as the comparison group. Their paper presents results from the first year (2005) of all four interventions, but focuses on the two teacher incentive programs. The evidence indicates that students in incentive schools performed significantly better than those in control schools by 0.19 and 0.12 standard deviations in math and language tests respectively. However, incentive schools also performed better on subjects for which there were no incentives. Interestingly, they find no significant difference in the effectiveness of group versus individual teacher incentives. In addition, incentive schools performed significantly better than other randomly-chosen schools that received additional schooling inputs of a similar value.

In Latin America teacher pay for performance incentives has been recently introduced. McEwan and Santibañez (2005) examines whether the program "Carrera Magisterial" in Mexico induced teachers to improve their student's test scores. The national office of Carrera Magisterial collects data that are used to calculate teachers' assessment scores. The sample is further limited to include primary teachers in grades third to sixth, given that students of teachers in grades first and second did not take any tests. The empirical strategy relies on the fact that some participating teachers of "Carrera Magisterial" face weaker incentives, because they face insuperable barriers to promotion in a given year, allowing them to serve as a comparison group in a regression-discontinuity framework. The essential premise of this paper is that many Mexican teachers –85% of their sample– faced weaker incentives, even when participating in the Carrera Magisterial assessment. Putting aside the empirical results, this by itself is a sobering conclusion. It suggests that a revised assessment process could focus upon groups of teachers that are most likely to face strong incentives. The paper found evidence of relatively small effects, at least for a subset of teachers near the relevant discontinuity. The empirical analysis suggests an effect on mean test scores of around 0.15-20 points in the vicinity of the discontinuity, equivalent to less than 10% of a standard deviation. The effect is robust to a variety of alternative specifications and subsamples, although some evidence suggests that the effect could be a lower bound.

In Chile Contreras, Flores, and Lobato (2005) attempt to evaluate the ex-post benefits of winning the prize. How does winning the award affect schools' mean test scores ex-post. They consider a regression analysis and a difference-in-difference estimator and find a small and positive impact of winning the SNED on future test scores. However, neither of the two techniques seems to exploit the potential quasi- experimental design of the program. While the regression analysis helps to shed light on statistical correlations between winning the SNED and future average test scores, it is not possible, in general, to identify a causal effect. Finally, the authors do not attempt to examine the tournament effect related to the program.

Finally, Gallego (2008) estimates the effect of winning the SNED on next period test score using a regression discontinuity design, this is called an ex-post effect which turns out to be negligible around 0.03 standard deviations. Then, in the same regression he aggregates a measure of the probability of winning and the density of the competition group and argues that an ex-ante effect can be computed with this approach. He found an ex-ante effect about 0.08 standard deviations.

In this article we provide an evaluation of the effect of the SNED on all participant schools, winners and losers. We contribute to the literature on teacher incentives on academic performance in Chile by estimating the incentive or tournament effect of the introduction of the program, i.e. the effect of the program over all schools affected by it, winners and losers. The effect of the incentive on standardized test scores at the school level is estimated by regressing the change in test score on a set of covariates and a treatment dummy variable (affected by the tournament or not). In order to correct for potential endogeneity we follow a double robust method (DR) which combines inverse probability reweighting with bias adjustment incorporating the covariates included in the propensity score. In addition, panel data estimation is pursued. Our results indicates a significant tournament effect on participant schools between 0.1 and 0.29 standard deviations for language and math test scores.

The rest of this paper is organized into six sections. Section 2 provides a brief description of the SNED teaching incentive program. The methodology and empirical strategy are discussed in section 3. Section 4 describes the data. The results are presented in section 5. In the final section we present the conclusions.

2 The Program

Prior to 1980, the administration of the Chilean school system was fully centralized in the Ministry of Education. The Ministry was not only responsible for the curriculum of the entire education system, but also for the administration of public schools, which accounted for 80 percent of all Chilean schools. The ministry also appointed public school teachers and principals, as well as approving and paying expenses and salaries. The decentralization process initiated in the early 1980s transferred the administration of public-sector schools to municipalities. Additionally, the reform opened the way for the private sector to participate as a provider of publicly financed education, by establishing a voucher-type, per-student subsidy. In Chile, schools are divided into three school administration types, based on funding source: (a) Public schools with public funding and administration; (b)

Private state-subsidized schools, in which the financing for each student is provided by the state but with private administration; and (c) Private fee-paying schools, in which both funding and administration are provided by the private sector. The voucher system gives families complete freedom to choose schools for their children. They can choose a subsidized school, either municipal or private. Alternatively, they can choose a fee-paying private school.[3]

The National System of School Performance Assessment (SNED) is directed to all primary and/or secondary subsidized schools in the country and is financed by the government. Thus, the private fee-paying schools are excluded. In the year 2000, 90% of all schools were municipal or public subsidized private schools. The SNED, which is a supply side incentive, was created with two objectives. First, to improve educational quality provided by state subsidized schools through monetary rewards to teachers. This strategy, defined as a pay-for-productivity wage compensation, seeks to change the fixed salary structure. The second objective was to provide the school community, parents and those responsible for children with information on the results and the progress of schools. It was expected that the school administrations and teachers would thus receive feedback on their teaching and administrative decisions.[4]

The SNED program is defined as follows. Schools are grouped by region. Then, they are classified according to location (urban/rural area), and as primary or secondary schools. Once these groups are defined, they are then subcategorized by vulnerability and socio-economic characteristics according the official classification provided by the Ministry of Education: high, medium-high, medium, low-medium and low levels. The ministry refers to the sets of schools associated together as homogeneous groups and thus investigates differences based on groups. This method is used because it is considered inappropriate to compare the performance of schools with adverse external conditions, such as low parental educational level, low family income and high social vulnerability, with the performance of schools with good external conditions. Therefore, following a tournament design, the competition among schools is supposed to take place within each homogenous group.

Once the group has been defined, the SNED index is computed for each school within its homogenous group and the schools are ranked according to this index. Top schools accounting for 25% of the enrollment in each homogeneous group are chosen for the Teaching Excellence Subsidy. These funds are distributed directly to the teachers as follows: 90% of the total bonus goes directly to all teachers at the rewarded school, based on the number of hours worked. The other 10% is allocated by the school as a differential bonus

---

[3] The school choice is limited by the school selection criterion and tuition fees.

[4] See Mizala and Urquiola (2007) for an evaluation of this second objective.

for those teachers whose contribution were more significant in achieving the performance goals or whose work was noteworthy. Payments are made quarterly. For the 1996-97 SNED competition, the yearly amount received by each teacher at awarded schools was about US$370. This is approximately 40% of a teacher's monthly income, equivalent to an increase of 3.33% in annual income.[5]

The factors determining the SNED index are the following:

1. Effectiveness, that is the educational results achieved by the school in relation to the population served. This considers the average SIMCE score in both language and mathematics during the past evaluation. For the 1996-1997 SNED competition this variable corresponded to the 1995 SIMCE score in eight grade and the 1994 SIMCE score in fourth grade. This factor was weighted of 40% in that years SNED index and has been decreased to 37% in the following rounds of the tournament.

2. Improvement consists of the differentials in educational achievement obtained over time by the school. It was weighted 30% in the 1996-1997 SNED and decreased to 28% in the following rounds. This measure of improvement varies according to the previous SIMCE score at the school level. For schools whose previous SIMCE test was in fourth grade this variable measures the average difference between 1994 and the 1992 SIMCE score. For those schools with previous information from eight grade testing, the comparison considered was 1995-1993.

3. Initiative, that is the capacity of the school to incorporate educational innovations and involve external agents in its teaching activities. It is measured through educational projects, teaching workshops, agreements with institutions and/or companies for work placement, and other related activities. The source used for this indicator is the SNED survey. It has a weight of 6% in all SNED rounds.

4. Improvement of working conditions and operation of the school. The indicators that make up this factor are complete permanent teaching staff and substitutes for absent teachers. This factor only has a 2% weight for all SNED rounds.

5. Equality of opportunities, which consists of accessibility to facilities and permanence of the schooling population, as well as the incorporation of those with learning difficulties. It is measured through the retention rates, inclusion of of multi-deficit and severe deficit students, integration in development projects and the pass rate of students. The information is obtained from the enrollment and performance

---

[5] The monetary incentive has increased to about US$1,000 per year in the 2006-2007 round which is about 80% of a teacher's monthly salary.

statistics of the Ministry of Education, apart from the SNED survey. The weight for this index was 12% in the 1996-1997 round and increased to 22% afterwards.[6]

6. Integration and participation of teachers and parents in the development of the educational role of the school. This factor is calculated from two indicators. The first is the establishment of parental centers and the second is the acceptance of their work. This information comes from the SNED Survey and the questionnaire for parents of the SIMCE. This factor had a 10% weight in the 1996-1997 round and decreased to 5% in the following rounds.

Each of these factors is made up of a series of indicators. Those with the greatest relative weights are the SIMCE scores, representing 70% of the 1996-1997 SNED index. Table 1 shows the evolution of those proportions.

One concern one could have is about the incentives schools have to attempt to maximize SNED score, instead of focusing on progress of students. This has been pointed out by Carnoy, Brodziak, Molina, and Socías (2007) who argue that given that the SIMCE tests is an inter-cohort test, schools might have an incentive of increase performance on grades are being tested. They show that when intra-cohort measures are compared, most of awarded schools show little, if any, progress in terms of academic achievment.

A second concern is related to other programs that may be interacting (confounding) with SNED. The only program that might confound with SNED is the *Jornada Escolar Completa* (JEC) which was a program to increase the number of hours from half day to full day school. This program was launched in 1997, a year after the first round of SNED with a coverage of 19% of schools. We will control for this including a dummy variable taking whether the school is a full day school or not.

3 Evaluation and Identification Strategy

In order to evaluate the effect of SNED on test scores, we try to address the question: how competition for the prize increases, if at all, schools' mean test scores? According to the neoclassical models of incentives, the introduction of a tournament may change the incentive structure of teachers and competition for the prize may be reflected in more motivated teachers, improved quality of education, and hence, an increase in participant school's mean test scores.

This question is not trivial given the difficulties faced when trying to identifying a causal relationship. The construction of a valid control group given the design of the

---

[6] This component prevent the possibility of selecting only good students.

program is troublesome. Participating schools in the SNED tournament account for 90% of the total number of schools in Chile (the private fee-paying schools being non-eligible). It is natural to think that pre-treatment characteristics for a private fee-paying school control groups would be different from the pre-treatment characteristics of subsidized schools. One alternative is to construct a control group using a matching procedure and perform a difference-in-difference approach. This implies the choice of an algorithm to match treated and control observations such as nearest-neighbor or matching in the propensity score. For nearest neighbors methods (such as Abadie and Imbens (2006)) it is not clear how to choose the number of neighbors. In addition, in the case of using propensity score methods, a miss-specified propensity score may lead to bias in the treatment effects estimates. A second alternative is to pursue a double-robust method. These methods have the advantage of being robust to either a miss-specified propensity score or a miss-specified model. We prefer the second alternative but we can provide matching estimators upon request.

A second approach we follow in this sections is a panel data estimator for the average treatment effect on the treated (ATT).

Finally, we study if the tournament implies the presence of schools that are always *on the money* (top schools that systematically rank in the upper quartile or so) and schools that are *out of the money*. Then a reduced number of schools in the experimental sample are actually affected by the tournament. We propose a simple method to identify schools *on the money* by estimating the probability of winning the 1996-1997 tournament with pre-tournament data and then, computing the difference between actual and predicted test scores for groups with different probability of winning.

3.1 Double Robust methods

The first approach to shed light on the tournament effect on test scores is to implement a *double robust* estimator. This method was first introduced by Robins and Rotnitzky (1995) and consist on estimate a weighted regression of the outcome variable on the treatment dummy and the covariates. The weights are computed as a function of the propensity score. The advantage of this class of methods is that the estimator is consistent whenever one of this two things happens: the model is correctly specified or the propensity score is correctly specified.

The propensity score will be calculated by estimating a *probit* regression for the probability of being treated, i.e. the probability of being a private-voucher or public school against being a private school. With the results of the *probit*, we can obtain the propensity score and the weights. As described in Busso, DiNardo, and McCrary (2009), double robust methods allows to estimate average treatment effect (ATE) and average treatment on the treated effect (ATT) by adjusting the weighting scheme. We will focus on ATT effects given that the program affects 90% of schools in the country.

The weighting scheme we use to estimate the ATT is given by $w = (p_s/(1-p_s))/(\hat{p}/(1-\hat{p}))$ for untreated and $w = 1$ for the treated where p_s is the propensity score and $\hat{p}$ is the unconditional probability of being treated. This is the scheme IPW1 analyzed by Busso, DiNardo, and McCrary (2009). Then, we estimate the following weighted regression model

$$\Delta Y_{i,t} = X_{i,t}\beta + \alpha d_{i,t} + \epsilon_{i,t}$$

Where $\Delta Y_{i,t}$ is the difference between SIMCE test-score before and after the introduction of the program, $X_{i,t}$ are the covariates related to characteristics of the school before the treatment, $d_{i,t}$ corresponds to a dummy variable that is equal to one if the school participates into the program (public and private subsidized) and is equal to zero if school is private, and $\epsilon_{i,t}$ is the residual. Given the data constraints, the variables considered for the estimation of the propensity score are regional dummies, student-teacher ratio (STR), and average parent education dummies. For the regressions, we add a dummy variable equal to one if school is a full day school.

In order to avoid comparability issues between different test scores we standardize each measure subtracting the mean and dividing by the standard deviation of the control population. This implies that we will be able to identify only how the treated group does relative to the control population, but that is in the spirit of the design anyway.

This first approach will not exploit the panel data nature of the data and will measure only the effect of the program after the first round. A second alternative is to construct a panel data and is explored below.

3.2 Panel Data Estimation

In addition to the double robust estimation, we construct a panel of schools from 1990 to 1999. This contributes to the analysis since it allows us to measure the incentive effect after two rounds of the SNED have been implemented controlling for pre-treatment information, school fixed effects, and geographic region trends.

The construction of the panel is not straightforward since the data is in the form of school averages until 1997 and from 1998 on it is at the individual level. Hence we have to compute school averages and create some aggregate variables such as average parent's schooling, type of school, etc. Given the continuous changes in the questionnaires it is hard to create or keep track of some variables.

With all, we were able to create an unbalance panel with more than 6.500 schools with an average number of periods of 6.4 (minimum of 1 and maximum of 10).

The equation we estimated is

$$\begin{aligned} Y_{i,t} &= \alpha_1 d\_treat_{i,t} + \alpha_2 d\_after_{i,t} + \alpha_3 d\_treat_{i,t} \times d\_after_{i,t} + X_{i,t}\beta + \\ &+ \gamma_1 trend_{i,t} + \sum_{j=1}^{12} \gamma_{2j} d\_region\_j_i \times d\_after_{i,t} \times trend_{i,t} + \epsilon_{i,t} \end{aligned}$$

where $d\_treat_{i,t}$ is a dummy taking the value 1 if the school is eligible for SNED (public or private subsidized) and 0 if it is private fee-paying. The dummy $d\_after_{i,t}$ is a binary variable equal to one if year is 1996 or after. The dummies d_region_j are twelve geographic region dummies. The variable *trend* is a time trend and $\epsilon_{i,t} = \eta_i + u_{i,t}$. The parameter of interest is the one accompanying the interacted dummies for eligibility and after the SNED started. As the error structure shows, we estimate this equation with school fixed effects and school clustered standard errors as well. Hence, we are able to identify the parameters for time variants covariates. In addition to the school fixed effects, we have dummies per level of parent's schooling, and we add geographic region trends interacted with the before-after dummy.

We perform this estimation in level and first difference, unweighted and then we reweight in the same way as we do in previous section.

3.3 Identifying schools *on the money*

As reported by McEwan and Santibañez (2005), teachers facing weaker incentives (about 85% of the sample) could explain why the effect of the Carrera Magisterial program are negligible. Then, we attempt to identify schools who are likely to lose and win with high probability to see of there is heterogeneity in the impact of SNED. One way of identifying such schools is to estimate the probability of winning with pre-tournament data and then compare predicted and actual test score (post treatment) for groups with different probability of winning.

We follow a procedure similar to Neal and Schanzenbach (2010). We compute the mean difference between predicted test score with information prior to the SNED program with actual test score following the introduction of the program. We perform this strategy separating by groups according to their probability of winning the tournament.

In order to determine the probability of winning the tournament, we estimate a linear model of the 1996 index on the lagged value of the math test scores and its second difference as follow

$$sned_{i,t} = \beta_1 simce_{i,t-1} + \beta_2 \Delta_{i,t-1} simce + \beta_3 \Delta_{i,t-2} simce + \beta_4 X_{i,t} + \epsilon_{i,t}$$

These variables capture the level and improvement factors defined in the formula of the SNED index. Given that we do not have the rest of the data tracked by the SNED index we add more controls such geographic region and urban/rural dummies. Then we predict the SNED index and compute for each homogeneous group the probability of winning. This is done by computing the cumulative distribution after sorting the schools (ascending) by the predicted SNED index in each homogeneous group.

Then, to find out about the presence of schools *on the money* and *out of the money* we compare the post-tournament test scores with their prediction using pre-tournament information. The distribution of this "prediction error" across the probability of winning (computed with pre-treatment data) may indicate the presence of schools *on-the-money* and a tournament effect for at least a sub-population of eligible schools.

In order to do this, we construct a panel data of eligible schools (public and private-voucher) from 1989 to 1995. Then, we estimate a linear dynamic panel data model of test scores on characteristics (such as school size, parental schooling, expenditure in tuition, and lags of the dependent and independent variables) following Arellano and Bond (1991).

With our estimated model we predict the 1996 test scores and compute their deviation from the true 1996 test scores. Hence, we can observe the distribution of this prediction error across the previously computed probability of winning. The presence of sure losers would be reflected in the presence of marked (fat) lower tail. Conversely, the presence of sure winners would be reflected in the presence of a upper tail.

Since this particular prediction error is between the post-tournament test score in 1996 and the results predicted with pre-tournament data (until 1995), if the tournament was ineffective the prediction error and the probability of winning should not be related. In the results section of this article we show that there would be a large group of sure losers and apparently no sure winners.

4 Data

This paper uses information from the national SIMCE (1989-1999) test. The data sets contain information for the period 1989-1999. Tests are conducted for students attending fourth, eighth or tenth grade. We have aggregate data, at the school level, from 1989 to 1997. Since 1998, student level data is available. However, we work with school level data since the tournament is at the school level. SIMCE data sets also include information about family and school characteristics. The continuous changes in the questionnaires during the period analyzed, makes the availability of covariates for the estimation limited. Hence, we are able to construct geographic region dummies, dummies for parents education level, and student-teacher ratio (STR).[7]

Table 2 presents the main school characteristics and performance levels by administrative school type: public, private subsidized and private fee-paying. The table summarizes information for the years 1996 and 2006. It indicates that private fee-paying schools have students of higher socioeconomic status than private subsidized and public schools do. Private fee-paying schools show the highest household income and parent education levels. Meanwhile, public schools have the lowest family income and parental education levels. Consistently, school performance in mathematics and language are lower in public schools compared to private subsidized and private fee-paying schools. Finally, there was a change in the SIMCE scoring scale in 1998. In 1996, the SIMCE test has an average around 70 points with standard deviation of about 10 points. Since 1998, the SIMCE test switched scale. The test exhibits an average of 250 points with a standard deviation of 50 points. Since then, SIMCE tests have been comparable over time, using same scale and grading.

Table 3 summarizes the same variables discussed above for winners and losers schools. This information is presented for 1996 and 2006. In both years we do not observe any significant differences in educational performance and socioeconomic characteristics between winner and losers schools. At first sight, it looks random, but these results should be interpreted carefully. First, given that competition occurs within a homogenous group, we expect to observe similar socioeconomic characteristics among schools in a particular group. Second, the simple average in performance is not capturing differences between homogenous groups. In other words, given that competition occurs within groups, differences in performance need to be observed between schools in the same homogeneous group. However, by comparing performance between winners and losers between groups, differences tend to be reduced in the 1996-1997 and 2006-2007 rounds of the tournament.

Table 4 shows the distribution of schools according to the number of awards received over time. This table indicates that 38% of schools have never been awarded the SNED bonus. Only a small fraction of schools have won the SNED several times. In other words, according to the evidence there are some schools that are *out-of-the-money* permanently. [8]

---

[7] For STR we are to construct the series for years 1993,1994,1995 and 1996

[8] This table slightly differs with Mizala and Urquiola (2007) who find a higher percentage of never winning schools but overall the findings tend to agree.

5 Results

In this section we present the results of the evaluation strategies discussed in section 3. We present results for the evaluation of the tournament effect using double robust methods and panel data.

As mentioned in section 4, in order to evaluate the tournament effect on test scores we estimate equation 1 by using double robust methods which consists in reweighting by a function of the propensity score and adding the covariates included in the propensity score estimation. These methods have been reported to perform well in finite samples and has the advantage of being robust to a misspecification of the propensity score or the model, separately. The set of variables considered are region dummies, student-teacher ratio, and average parents' education dummies.

The treatment group includes public and private subsidized schools and the control group includes the private schools.[9]. When implementing the double-robust method we use robust standard errors to account for heteroskedasticity and in the panel data estimates we used school fixed effects and school clustered standard errors as well.

The estimation of the propensity score can be found in the appendix, as well as, the weighted means for treated and untreated. As can be appreciated, the propensity score appears to be very well specified (pseudo R^2=0.73, sensitivity above 97% and specificity above 83%.). The covariates are relatively well balanced since the pairwise mean difference tests for each covariate rejects the null hypothesis of equality in five of sixteen cases. However, the joint hypothesis of equality is rejected.

Table 5 presents the average treatment effect on the treated (ATT) for math and language scores. The outcome variable is the 1996-1995 pairwise difference in standardize test score. We also use the 1997-1995 pairwise difference in standardize test score since schools might have take some time to react to the introduction of the program. As it can be appreciated, the effect is positive but insignificant when using the 1996-1995 pairwise difference as outcome variable. When the outcome variable is the 1996-1995 pairwise difference in standardize test scores the effects are 0.14 and 0.27 standard deviation for language and math respectively. When using the 1997-1995 difference the effects are stronger and statistically significant between 0.25 and 0.29 standard deviation for math and language respectively. [10]

---

[9] We also considered excluding the public schools in order to increase comparability between the treatment and control groups obtaining similar results and a slightly better balance of covariates after reweighting. However, by excluding public schools we would be losing schools we care about from a policy perspective.

[10] In the appendix we provide the results for the propensity score estimation, the reweighted regressions and descriptive statistics to exam the balance of the covariates after reweighting.

The results are in line with those found by Muralidharan and Sundararaman (2009) in India and about two times those found by McEwan and Santibañez (2005) in Mexico.

The panel data evidence is consistent with what was found with double robust methods as can be appreciated in Table 6. The variable d_treat is the treatment dummy variable and d_after is a dummy variable that takes the value 1 from 1996 on. Since the estimation is with school fixed effects only time variants covariates are identified (then regional dummies are excluded from the estimation). In column one we present the results for math controlling by school fixed effects and region dummies interacted with the before after dummy and a trend. Hence, we allow for different trends before and after the introduction of the SNED. In column two we perform the same fixed effects estimation for math test scores but we also reweight in the same way we did in the Double Robust approach. In columns three and four we do the same as columns one and two but for language tests. As we can observe, for math and language we find a significant effect about 0.16 and 0.14 standard deviation respectively when no weights are introduced. The effect rises up to 0.22-0.25 standard deviations when we reweight. To explore the presence of heterogeneity between public and private subsidized schools we added a dummy variable of being a public school interacted with the before-after dummy(d_public x d_after). The results show no significance of the interacted dummy. Also separated estimations were performed, i.e. public against private and private subsidized agains private finding no significant difference.

In Table 6 we see also a dummy variable controlling for full-day schools which is positive and significant ranging from 0.06 to 0.07 standard deviations depending on the specification. The variables measuring average parent's schooling results positive but nonsignificant.

Now, we perform the panel data estimation with the outcome variable in first difference. The results can be observed in Table 7. As can be appreciated, the effect of SNED on math scores does not change in a significant manner. However, the effect of the SNED on language test scores decreases to 0.05 standard deviations being nonsignificant when reweighting.

Related to the probability of winning, in Figure 1 we can see the box plots of the prediction error of test scores across the predicted probability of winning. The probability of winning is categorized into 20 categories. The first category includes schools with probability of winning between 0 and 0.05, and so on.[11] Then, it can be seen that the

---

[11] In case the reader is not familiarized with this type of plots, each box contains 50% of the data for each category, from the 25th to the 75th percentile. The line in the middle of the box represents the median or 50th percentile, and the other lines (whiskers) are 1.5 times the inter-quartile ratio (distance from the 25th to the 75th percentile). Observations lying outside the whiskers are considered outliers.

tournament seems to affect schools with probability of winning greater than 0.6 or 0.65. This suggests the existence of sure losers, schools that were not affected by the tournament according to the prediction error.

Now, in order to see if non-eligible schools show the same pattern as in Figure 1, we repeat the exercise for private schools (false experiment). Then we predict their SIMCE test score for 2006 using pre-treatment information and compute the probability of winning on "artificial" *homogeneous groups.* These groups were constructed using geographic region and urban/rural status and the empirical probability of winning is computed for each group. Figure 2 shows the box plots of the prediction error of test scores across the predicted probability of winning. It is interesting to note that the pattern observed in Figure 1 is not observed here. Thus, non-eligible schools are not subject to the tournament. This fact validates our identification strategy.

Finally, we estimate the double robust model for schools *on-the-monety* according to what we observe in Figure 1. Hence, we restrict the estimates to schools with probability of winning between 0.6 and 0.95. The results are presented in Table 8. As we see the ATT raises for all estimates increasing the statistical significance. Of course this is a reduced number of schools and these results are not generalizable to the whole population but show that schools *on-the-money* may have higher incentive effect than those *out-of-the-money.*

In sum, the evidence indicates that the SNED program had a positive and significant effect on the educational achievement for a sub-population of eligible schools.

6 Conclusion

Although performance-related pay for teachers is being introduced in many developed countries, little evidence has been provided based for measured effects in LDCs. This article contributes with empirical evidence on the effects of performance-related incentive pay for teachers based on school academic performance. We examine the effect of a rank-order tournament, the The National System of School Performance Assessment (SNED), on standardized test scores. A major feature of this program is that it is a scaled-up one affecting about 90% of the schools in the country.

Double robust methods and panel data estimation are pursued in order to estimate the incentive effect of the introduction of the program on test scores. We find a significant effect of the program on standardized math and language test scores. The results are comparable to those found by Muralidharan and Sundararaman (2009) in India with the distinction that the SNED is a scaled-up program and the experiments in India affected a small fraction of the schools in the country. The results are robust to different specifications and vary between 0.14 and 0.25 standard deviations for math and from 0.9 to 0.23 for language.

The empirical evidence presented in this paper provides support for educational policies oriented towards greater differentiation in the salary structure for teachers. In many countries where teachers unions are very important (in particular in Latin America and less developed countries), a wage structure which recognizes pay-for-productivity would be theoretically efficient. This paper provides evidence supporting a wage structure for teachers that is more related to productivity as a mechanism to increase student achievement. However, this paper also shows that this type of tournaments is only productive for a certain subset of schools, given the existence of sure winners and losers. In the case of Chile, nearly half of eligible schools have never won the award after eleven years of implementation. Thus, the evidence shows that such rewards system may only create improvements in a fraction of schools. Further research is needed to evaluate different designs and incentive mechanisms to affect a broader range of schools.

References

Abadie, A., and G. Imbens (2006): "Large Sample Properties of Matching Estimators for Average Treatment Effects," *Econometrica*, 74(1), 235–267.

Arellano, M., and S. Bond (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application," *Review of Economic Studies 58*, 58, 277–297.

Bank, W. (2006): *World Development Report 2006: Equity and Development.* The World Bank and Oxford University Press, Washington.

Barlevy, G., and D. Neal (2009): "Pay for Percentile," *FRB of Chicago Working Paper No. 2009-09.*

Bellei, C. (2009): "Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile," *Economics of Education Review*, 28(5), 629–640.

Busso, M., J. DiNardo, and J. McCrary (2009): "Finite Sample Properties of Semiparametric Estimators of Average Treatment Effects," *Mimeo, Department of Economics, University of California-Berkeley.*

Carnoy, M., I. Brodziak, A. Molina, and M. Socías (2007): "The Limitations of Teacher Pay Incentive Programs Based on Inter-Cohort Comparisons: The Case of Chile's SNED," *Education Finance and Policy*, 2(3), 189–227.

Contreras, D., L. Flores, and F. Lobato (2005): "Monetary Incentives for Teachers and School Performance: Evidence for Chile," *Mimeo, Universidad de Chile.*

Gallego, F. (2008): "Efectos del SNED en Resultados del Proceso Educativo," *Mimeo, Catholic University of Chile.*

Glewwe, P., N. Ilias, and M. Kremer (2010): "Teacher Incentives," *American Economic Journal: Applied Economics*, 2(3), 205–227.

Hanushek, E., J. Kain, and S. Rivkin (1999): "Do Higher Salaries Buy Better Teachers?," *NBER Working Paper 7082.*

Harbinson, R., and E. Hanushek (1992): *Educational Perfomance of the Poor: Lessons from Rural Northeast Brazil.* NY: Oxford University Press.

Lavy, V. (2002): "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement," *Journal of Political Economy*, 110(6), 1286–1317.

McEwan, P., and L. Santibañez (2005): *Teacher and Principal Incentives in Mexico, en E. Vegas (ed.): Incentives to Improve Teaching: Lessons from Latin America.* Washington DC,The World Bank.

Mizala, A., and P. Romaguera (2002): "Evaluación del desempeño e incentivos en la educación Chilena," *Cuadernos de Economía*, 118, 353–394.

Mizala, A., and M. Urquiola (2007): "School Markets: The Impact of Information Approximating Schools' Effectiveness," *NBER Working Paper*, (13676).

Muralidharan, K., and V. Sundararaman (2009): "Teacher Performance Pay: Experimental Evidence from India," *Working Paper 15323, NBER.*

Neal, D., and D. W. Schanzenbach (2010): "Left Behind by Design: Proficiency Counts and Test-Based Accountability," *The Review of Economics and Statistics*, 92(2), 263–283.

Robins, and Rotnitzky (1995): "Semiparametric Efficiency in Multivariate Regression Models With Missing Data," *Journal of the American Statistical Association*, 90(429), 122–129.

SIMCE (1989-1999): "Base de Datos SIMCE (1989-1999)," *Santiago, Chile: SIMCE, Ministerio de Educacin.*

Table 1: Description

Factor	SNED weighting 96-97	*SNED weighting 98-99*
Effectivity	40%	37%
Improvement	30%	28%
Initiative	6%	6%
Improvement of working conditions	2%	2%
Equality of opportunities	12%	22%
Incorporation of parents	10%	5%

Source: Ministry of Education

Table 2: School characteristics, administrative dependency and performance

	1996			**2006**		
Variables by School	Private	Private Subsidized	Public	Private	Private Subsidized	Public
SIMCE Score						
SIMCE Mathematics	83.61	69.59	65.61	288.09	243.75	231.47
	(7.31)	(11.42)	(10.21)	(28.34)	(34.10)	(31.94)
SIMCE Language	84.48	70.38	65.43	289.33	252.84	243.77
	(6.14)	(11.74)	(9.78)	(24.84)	(28.45)	(28.44)
Household Variables						
Average Schooling of Parents	4.44	2.70	2.18	4.01	3.20	2.69
	(0.58)	(0.76)	(0.47)	(0.20)	(0.69)	(0.52)
Average Schooling of Mothers	-	-	-	4.01	3.20	2.72
	-	-	-	0.18	0.70	0.52
Average Schooling of Fathers	-	-	-	4.14	3.23	2.71
	-	-	-	0.37	0.68	0.54
Average Household Income	-	-	-	1045554.00	290863.10	148104.90
	-	-	-	(205955.40)	(196710.30)	(82136.08)
Schools Variables	-	-	-			
Rural	0.01	0.16	0.51	0.04	0.21	0.60
	(0.12)	(0.36)	(0.50)	(0.19)	(0.41)	(0.49)
Number of students taking the test	43.15	56.94	46.18	35.28	38.61	25.60
	(33.17)	(49.59)	(44.26)	(30.83)	(36.86)	(30.06)

Source: Authors calculation based on SIMCE data set

Table 3: Schools performance: winners and losers

	1996		2006	
Variables by School	Winers	Losers	Winers	Losers
SIMCE Score				
SIMCE Mathematics	68.27	66.24	249.37	248.27
	(11.19)	(10.52)	(28.44)	(25.73)
SIMCE Spanish	68.49	66.33	257.11	255.92
	(11.28)	(10.37)	(24.12)	(22.84)
Household Variables				
Average Schooling of Parents	2.38	2.33	3.04	3.15
	(0.67)	(0.59)	(0.58)	(0.61)
Average Schooling of Mothers			3.03	3.16
			(0.59)	(0.61)
Average Schooling of Fathers			3.05	3.15
			(0.59)	(0.60)
Average Household Income			232262.00	250254.10
			(159815.90)	(159661.40)
Schools Variables				
Rural	0.43	0.40	0.36	0.30
	(0.49)	(0.49)	(0.48)	(0.46)
Number of students taking the test	48.50	49.85	41.83	45.43
	(46.14)	(46.06)	(36.11)	(37.47)

Source: Authors calculation based on SIMCE data set

Table 4: Schools by number of awards (6 rounds participant)

Number of awards	Frequency	Percent
0	3,108	38.64
1	2,085	25.92
2	1,339	16.65
3	802	9.97
4	427	5.31
5	215	2.67
6	68	0.85
Total	8,044	100

Table 5: Tournament effects in Math and Language, Double Robust

Language :

Difference	ATT	SD	T-test	Obs
96-95	0.14	0.14	1.04	1784
97-95	0.29	0.09	3.34	1762

Math :

Difference	ATT	SD	T-test	Obs
96-95	0.27	0.26	1.06	1792
97-95	0.25	0.10	2.46	1762

Table 6: Tournament effects, Panel 1990-1999

Coefficient	Math		Language	
	(1)	(2)	(3)	(4)
d_after	0.177***	0.067	0.210***	0.147***
	(0.036)	(0.058)	(0.033)	(0.055)
d_treat x d_after	**0.157*** **	**0.252*** **	**0.138*** **	**0.220*** **
	(0.022)	(0.033)	(0.020)	(0.039)
d_public x d_after	0.001	-0.002	-0.007	-0.008
	(0.014)	(0.015)	(0.013)	(0.014)
d_primary	0.027	0.032	-0.001	0.010
	(0.049)	(0.050)	(0.040)	(0.041)
d_secondary	0.066	0.058	0.074*	0.095**
	(0.050)	(0.052)	(0.042)	(0.047)
d_college	0.062	0.055	0.085**	0.070
	(0.049)	(0.051)	(0.041)	(0.044)
d_full_day	0.072***	0.072**	0.056***	0.065**
	(0.016)	(0.032)	(0.014)	(0.030)
constant	-0.094*	0.127***	-0.090**	0.108***
	(0.048)	(0.049)	(0.040)	(0.040)
School Fixed effects	Yes	Yes	Yes	Yes
Region trends	Yes	Yes	Yes	Yes
Reweighted	No	Yes	No	Yes
N	43270	43270	43231	43231

Note: standard errors in parenthesis. Standard errors clustered by school.
Schools fixed effects included. Outcome variable in level.
* Statistical significance at the 10%, ** Idem, 5%, *** Idem, 1%

Table 7: Tournament effects, Panel 1990-1999

Coefficient	Math		Language	
	(1)	(2)	(3)	(4)
d_after	-0.282***	-0.729***	-0.056	-0.295
	(0.078)	(0.144)	(0.073)	(0.194)
d_treat x d_after	**0.153*** **	**0.286*** **	**0.047* **	**0.046**
	(0.027)	(0.064)	(0.026)	(0.055)
d_public x d_after	0.016	-0.002	-0.01	-0.029
	(0.014)	(0.027)	(0.014)	(0.021)
d_primary	0.029	0.035	-0.015	-0.006
	(0.106)	(0.108)	(0.083)	(0.085)
d_secondary	0.019	-0.065	0.01	0.014
	(0.109)	(0.141)	(0.086)	(0.092)
d_college	0.020	-0.118	-0.043	-0.275
	(0.115)	(0.215)	(0.092)	(0.197)
d_full_day	0.070**	0.149**	0.007	0.059
	(0.029)	(0.072)	(0.026)	(0.090)
constant	-0.048	0.025	-0.02	0.028
	(0.106)	(0.125)	(0.082)	(0.096)
School Fixed effects	Yes	Yes	Yes	Yes
Region trends	Yes	Yes	Yes	Yes
Reweighted	No	Yes	No	Yes
N	29315	29315	29271	29271

Note: standard errors in parenthesis. Standard errors clustered by school. Schools fixed effects included. Outcome variable in differences.

* Statistical significance at the 10%, ** Idem, 5%, *** Idem, 1%

Table 8: Tournament effects in Math and Language, Double Robust excluding sure losers and sure winners

Language :

Difference	ATT	SD	T-test	Obs
96-95	0.32	0.18	1.77	621
97-95	0.37	0.15	2.45	628

Math :

Difference	ATT	SD	T-test	Obs
96-95	0.42	0.21	1.94	621
97-95	0.32	0.11	2.98	628

Excludes schools with probability of winning less than 0.6 and more than 0.95

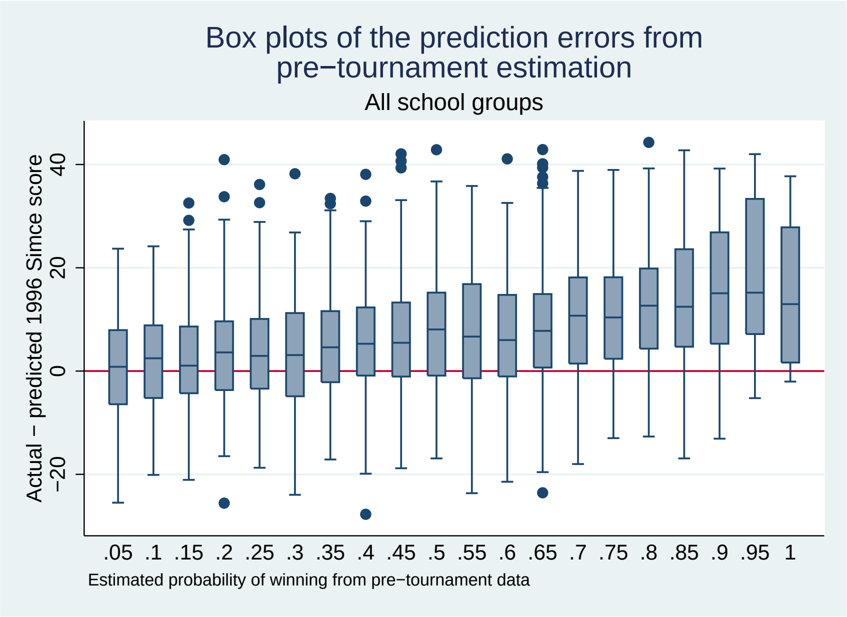

Figure 1: Box plots of the test score prediction errors across probability of winning groups: All eligible schools

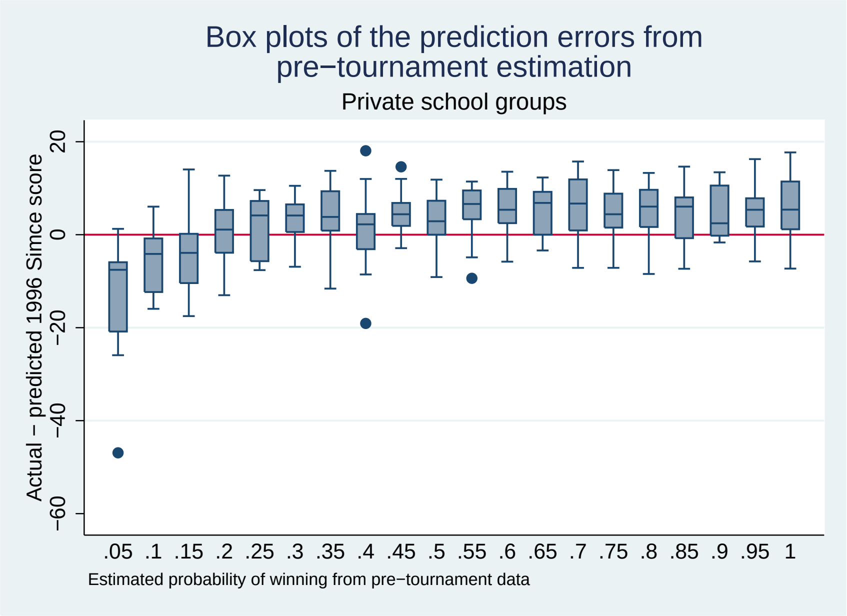

Figure 2: Box plots of the test score prediction errors across probability of winning groups: Non-eligible Schools

Appendix

Table 9: Probit, Propensity Score. Dependent variable: treated

Variable	Coefficient
d_primary	-4.003***
	(0.707)
d_secondary	-6.232***
	(0.710)
d_college	-8.294***
	(0.741)
d_region_2	0.796*
	(0.468)
d_region_3	1.660***
	(0.472)
d_region_4	0.763**
	(0.389)
d_region_5	-0.004
	(0.291)
d_region_6	0.135
	(0.363)
d_region_7	0.525
	(0.378)
d_region_8	0.403
	(0.319)
d_region_9	0.938**
	(0.415)
d_region_10	0.222
	(0.398)
d_region_11	-0.757
	(1.057)
d_region_12	0.403
	(0.455)
d_region_13	-0.015
	(0.280)
STR	0.010**
	(0.004)
d_full-day	-0.794***
	(0.273)
cons	5.882***
	(0.760)
N	4195
Pseudo R2	0.73

* Statistical significance at the10%

** Idem, 5%, *** Idem, 1%

Specificity 84%, Sensitivity 97%

Table 10: Balance of treated an untreated using reweighting

	Untreated		Treated		t-stat*
	mean	sd	mean	sd	
d_primary	0.83	0.38	0.81	0.39	0.33
d_secondary	0.14	0.35	0.16	0.37	0.35
d_college	0.03	0.17	0.01	0.08	2.01
d_region_2	0.24	0.43	0.04	0.19	6.19
d_region_3	0.00	0.03	0.02	0.14	0.89
d_region_4	0.01	0.08	0.04	0.19	1.00
d_region_5	0.05	0.22	0.15	0.36	1.79
d_region_6	0.05	0.23	0.04	0.20	0.32
d_region_7	0.01	0.11	0.04	0.20	1.02
d_region_8	0.16	0.37	0.11	0.32	0.94
d_region_9	0.01	0.09	0.05	0.21	1.14
d_region_10	0.11	0.32	0.05	0.22	1.85
d_region_11	0.01	0.07	0.02	0.14	0.70
d_region_12	0.00	0.06	0.02	0.13	0.65
d_region_13	0.31	0.46	0.39	0.49	0.96
STR	28.80	57.54	27.25	11.23	0.06
d_full-day	0.001	0.02	0.01	0.10	0.98
χ^2 Stat	=	57.12			
P-Value	=	0.000			

*Pairwise t-test for difference of means. ** Joint test for equality of means (rejected).

Table 11: Tournament effects, Math and Language

	Math		Language	
	$\Delta 96-95$	$\Delta 97-95$	$\Delta 96-95$	$\Delta 97-95$
treat	0.266	0.253**	0.141	0.293***
	(0.250)	-0.103	(0.136)	(0.088)
d_pub	0.015	-0.04	-0.041	-0.088***
	(0.044)	-0.034	(0.035)	(0.023)
d_primary	0.121	-0.082	-0.251*	-0.255***
	(0.237)	-0.376	(0.134)	(0.035)
d_secondary	-0.033	-0.142	-0.369**	-0.274***
	(0.231)	-0.378	(0.121)	(0.048)
d_college	0.132	-0.31	-0.234	-0.285
	(0.237)	-0.401	(0.137)	(0.208)
d_full_day	0.268***	0.304	0.156***	0.167***
	(0.018)	-0.264	(0.014)	(0.006)
STR	-0.003***	0.001	-0.003***	-0.001
	(0.001)	-0.001	(0.001)	(0.001)
d_region_2	0.034**	-0.005	-0.055***	0.167***
	(0.012)	-0.087	(0.006)	(0.005)
d_region_3	-0.138***	-0.163	-0.144***	-0.225***
	(0.040)	-0.108	(0.020)	(0.015)
d_region_4	-0.042	-0.142	-0.177***	-0.213***
	(0.031)	-0.105	(0.017)	(0.012)
d_region_5	-0.170***	-0.099	-0.248***	-0.132***
	(0.026)	-0.082	(0.017)	(0.008)
d_region_6	-0.213***	-0.04	-0.429***	-0.106***
	(0.009)	-0.134	(0.009)	(0.007)
d_region_7	0.100***	-0.031	-0.051**	-0.084***
	(0.028)	-0.099	(0.017)	(0.012)
d_region_8	-0.078***	-0.056	-0.132***	-0.070***
	(0.015)	-0.082	(0.007)	(0.007)
d_region_9	-0.119***	-0.046	-0.168***	-0.021
	(0.032)	-0.102	(0.018)	(0.012)
d_region_10	-0.195***	-0.188*	-0.247***	-0.018***
	(0.011)	-0.099	(0.005)	(0.003)
d_region_11	0.491***	-0.003	0.408***	0.585***
	(0.143)	-0.157	(0.070)	(0.051)
d_region_12	0.047	-0.068	-0.331***	-0.114***
	(0.035)	-0.146	(0.018)	(0.012)
d_region_13	-0.130***	-0.197**	-0.273***	-0.102***
	(0.010)	-0.077	(0.008)	(0.010)
constant	-0.064	-0.035	0.556***	0.150
	(0.182)	-0.394	(0.094)	(0.090)
N	1792	1762	1784	1762
R2	0.03	0.03	0.03	0.03

Robust standard errors in parenthesis.
Reweighting using weights from section 3.1

Table 12: Tournament effects, Math and Language excluding sure losers and sure winners

	Math		Language	
	$\Delta 96-95$	$\Delta 97-95$	$\Delta 96-95$	$\Delta 96-95$
treat	0.423*	0.324***	0.321*	0.366**
	(0.218)	(0.109)	(0.181)	(0.149)
d_pub	0.044	0.169	0.024	0.085
	(0.126)	(0.109)	(0.118)	(0.113)
d_primary	0.478	0.393	0.360	0.134
	(0.372)	(0.250)	(0.278)	(0.533)
d_secondary	0.505	0.409	0.255	0.08
	(0.406)	(0.269)	(0.310)	(0.545)
d_college	1.046	0.242	1.273	1.066
	(1.026)	(0.374)	(0.902)	(0.892)
d_full-day	-0.272	0.117	0.023	0.543
	(0.280)	(0.287)	(0.239)	(0.341)
STR	-0.005	0.001	-0.006**	-0.003
	(0.003)	(0.001)	(0.003)	(0.002)
d_region_2	-0.590***	-0.238	-0.587***	0.028
	(0.172)	(0.169)	(0.219)	(0.196)
d_region_3	-0.348	-0.349	-0.765**	-0.407*
	(0.389)	(0.288)	(0.344)	(0.222)
d_region_4	-0.575	-0.258	-0.573*	-0.451*
	(0.413)	(0.267)	(0.325)	(0.268)
d_region_5	-0.228	0.007	-0.330	-0.128
	(0.221)	(0.192)	(0.250)	(0.201)
d_region_6	-0.005	0.096	-0.830***	-0.347
	(0.546)	(0.351)	(0.227)	(0.273)
d_region_7	0.059	0.017	-0.007	0.082
	(0.256)	(0.246)	(0.276)	(0.252)
d_region_8	0.260	-0.030	-0.055	-0.076
	(0.279)	(0.186)	(0.249)	(0.219)
d_region_9	-0.761***	0.060	-0.732***	0.006
	(0.258)	(0.283)	(0.276)	(0.213)
d_region_10	-0.972***	-0.284	-0.852***	0.219
	(0.257)	(0.241)	(0.303)	(0.217)
d_region_11	0.228	-0.078	-0.179	-0.006
	(0.378)	(0.297)	(0.349)	(0.562)
d_region_12	-0.679*	-0.167	-0.968***	-0.062
	(0.359)	(0.348)	(0.347)	(0.269)
d_region_13	-0.467**	-0.373**	-0.538**	-0.145
	(0.221)	(0.187)	(0.241)	(0.216)
constant	-0.104	-0.419	0.279	-0.131
	(0.444)	(0.306)	(0.379)	(0.576)
N	637	621	628	621
R2	0.20	0.17	0.18	0.11

Robust standard errors in parenthesis.
Reweighting using weights from section 3.1
Excludes schools with probability of winning less than 0.6 and more than .95