Tournaments, gift exchanges, and the effect of monetary incentive for teachers: the case of Chile.*

Tomás Rau Binder†
Dante Contreras‡

First draft: August 2007
This draft November 2008

Abstract

In this paper we develop a simple model of school incentives and evaluate the introduction of monetary incentives for teachers, based on a school performance tournament in Chile. We evaluate the tournament effect, i.e. the effect of introducing the incentive scheme on all participant schools: winners and losers. We also evaluate the effect of winning the tournament on next period school performance that we call the gift-exchange effect. Matching and Regression Discontinuity techniques are used to identify both treatment effects. The results indicate a positive and significant tournament effect and a positive but nonsignificant gift-exchange effect.

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

*We are indebted to the useful comments of David Card, Patrick Kline, Raj Chetty, Marcela Perticará and seminar participants at UC Berkeley, Universidad de Chile, U Alberto Hurtado, and the annual meeting of the Chilean economic association (SECHI). We also thank the funding granted by Iniciativa Cientfica Milenio "Centro de Microdatos", Proyecto P07S-023-F.

†Departamento de Economía and Centro de Microdatos, Universidad de Chile. Diagonal Paraguay #257, Santiago, Chile. Tel: (562) 978-3634, Fax: (562) 978-3421, email: trau@econ.uchile.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 increasing earnings. In Latin America, the evidence indicates that education helps to reduce income inequality and alleviate poverty. Thus, policies to improve education constitute a key area of concern in achieving sustainable economic growth and 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 percent 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 SIMCE test, special educational quality and equity improvement programs, educational programs targeted to the poorest schools and extending the school day. However, little empirical evidence has been provided for the evaluation of such programs. Since 1996, the Ministry of Education has incorporated a monetary- productivity bonus called The National Subsidized School Performance Evaluation System (SNED). This is a rank-order tournament directed towards all the municipal and private subsidized schools of the country, which represent 90% of national enrollment. 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 competing mainly on the basis of their pupils results on the SIMCE. The program is based on 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 the monetary rewards are distributed equally among all teachers in the winning schools. Although performance-related pay for teachers is being introduced in many developed countries, little evidence has been provided based on measured effects in LDCs. There are at least two theoretical models to explain a relation between teacher incentives and educational performance. First, tournaments 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, increase in participant school's mean test scores. The second argument arise from the "gift exchange" or "reciprocal gifts" theory (Akerlof (1982) and Akerlof (1984)). In this model, the awarded teachers would exert more effort after obtaining the prize as a "gift exchange" with the principal or the community. Both models are empirically tested.

This paper provides evidence on the impact of this incentive on academic achievement in Chile. The effect of the incentive on standardized test scores at the school level is estimated by using both Matching in characteristics estimators and Regression Discontinuity analysis. These techniques are used to check two alternative theoretical models which may explain outcomes related to this type of incentives: a gift-exchange model and the total productivity model.

The rest of this paper is organized into six sections. Section 2 provides a brief description of the SNED teaching incentive program. Section 3 details the theoretical model. The methodology and empirical strategy are discussed in section 4. Section 5 describes the data. The results are presented in section 6. 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 whole education system, but also for the administration of the public schools, which accounted for 80 percent of all schools in the country. 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 the municipalities. In addition, 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 where their financing comes from: (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.[1]

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

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 socio-economic characteristics according the official classification provided by the Ministry of Education: high, medium-high, medium, low-medium and low income levels. The Ministry of Education refers to the sets of schools associated together in this manner as homogeneous groups and thus investigates differences in schools through the analysis

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

of conglomerates. 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 the teachers at the rewarded schools, based on the number of hours worked by each teacher. The other 10% is allocated by the school as a differential bonus for those teachers whose contribution was more significant in achieving the performance goals or whose work was noteworthy. For the 1996-97 SNED competition, the yearly amount received by each teacher on the awarded schools was about US$370. This is approximately 40% of a teacher's monthly income, which is equivalent to a increase of 3.32% of teachers' earnings. [2]The payment is made quarterly.

The factors determining the SNED index are the following:

1. Effectiveness: the educational results achieved by the school in relation to the population served. This considers the average SIMCE score in 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 had a weighting of 40% in that year SNED index and it decreased to a 37% in the following rounds of the tournament.

2. Improvement: Consists of the differentials in educational achievement obtained over time by the school. It had a 30% weighting in the 1996-1997 SNED and decreased to a 27% in the following rouds. The

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

measure of improvement varies according to the previous past SIMCE score at school level. For school which previous SIMCE test was at fourth grade primary education this variable measures the average difference between 1994-1992 SIMCE score. For those schools with previous information on eight grade primary education, the difference considered was 1995-1993.

3. Initiative: 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, among others. The information source used for this indicator was the SNED survey. It has have a weighting of 6% in all the 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% weighting in all SNED rounds.

5. Equality of opportunities: accessibility to facilities and permanence of schooling population, as well as the incorporation of groups with learning difficulties. It is measured through the retention rates, incorporation of multi-deficit and severe deficit students, differential groups in operation, integration in development projects and the pass rate of students. The information is obtained from the enrollment and performance statistics of the Ministry of Education, apart from the SNED survey. The weighting for this index was 12% in the 1996-1997 round and increased to a 20% afterwards.[3]

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 educational work. The information comes from the SNED Survey and the questionnaire for

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

parents of the SIMCE. This factor had a 10% weighting in the 1996-1997 round and decreased to a 5% in the following rounds.

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

3 Theoretical Model

In this section we present a model of school effort to show how monetary incentives may affect the level of effort exerted. This model also shows how the increase in the unconditional probability of winning may reduce the level of effort exerted. The model is in the spirit of Kandel and Lazear (1992) and Lazear and Rosen (1981) incorporating the fact that incentives are grouped and the probability of winning the tournament depends on the own effort, the effort of the competitors, and the percentage of winners.

The unit of analysis in this model is the school, seen as a group of teachers, instead of individual teachers from a particular school. There is at least two reasons to follow this approach. First, the SNED prize is given to the school and then shared by teachers. Then, it may be more interesting to model the competition among schools instead of focusing on moral hazard on teams.

Second, given that the SIMCE test is taken to just one grade per year (it alternates between fourth, eight and tenth grade), there are very few teachers directly affected by the tournament, at least in the first round of the tournament. Of course four graders math test scores depends, not only on four grade math teachers' effort or performance, but also on previous classes and their respective teachers. However, we are going to focus on the introduction of the tournament or first round and, therefore, only teachers in the particular grade that is being evaluated will be affected by the tournament.

Consider school i facing a probability $P(e_i, e_j, q)$ of winning the SNED bonus m, where e_i is its level of effort exerted, e_j is the effort of competing school j (not observed), and q is a quantile indicating one minus the percentage of winners (or the percentage of losers). $P(e_i, e_j, q)$ is increasing in e_i and decreasing in e_j, q. Let us assume, for now, that there are neither sure winners nor sure losers.[4] It is assumed that $\lim_{e_i \to \infty} P(e_i, e_j, q) = \psi << 1$, then schools can affect the probability of winning but a little.

Consider also that a school face a disutility of working equal to $\phi(e_i)$, an increasing convex functions in e_i. Therefore, schools maximizes the following expected utility,

$$\begin{aligned} \max_{e_i} U(e_i, m) &= P(e_i, e_j, q)m - \phi(e_i) \\ s.t. \quad & e \geq \underline{e} \end{aligned} \tag{1}$$

Assuming an interior solution, the first order condition gives us

$$m = \frac{\phi'(e_i)}{P'(e_i, e_j, q)}$$

where $P'(e_i, e_j, q)$ is the derivative of $P(e_i, e_j, q)$ with respect to e_i. Now, since $\phi(e_i)$ is assumed convex in e_i and $P(e_i, e_j, q)$ concave in e_i (at least in the set $e_i \geq \underline{e}$), we have that the right-hand-side is an increasing function of e_i, say $\phi'(e_i)/P'(e_i, e_j, q) = g(e_i, e_j, q)$. Hence the *reaction function* of school i is given by

$$e_i^* = g^{-1}(m, e_j, q)$$

where g^{-1} is increasing in m by definition of inverse function. Hence, as it can be expected the level of effort of school i increases with the amount of the monetary bonus m. Of course, the level of effort is also going to depend of the probability of winning in a more complicated way. The percentage of winners in the tournament (i.e. 25% in the SNED tournament) affects also this results via q. Note as well that effort of school j, e_j, affect negatively the probability of winning for school i.

Figure 1 depicts the the optimal choice of effort of school i, taken e_j

[4]The presence of sure losers or sure winners breaks up the tournament and will be discussed in the next section. In that case, the probability of winning depends on q and does not depend on effort.

and q as given. The probability of winning is plotted assuming school j effort fixed which makes $P(e_i, e_j, q)$ a cumulative (conditional) probability function which does not necessarily converge to one. The optimal choice occurs where the curves are tangent, which is unique as a consequence of convexity of $\phi(e_i)$ and concavity of $P(e_i, e_j, q)$ in e_i (at least in the set $e_i \geq \underline{e}$).

Now let us discuss the potential effect of an increase in the number of winners in this type of tournaments. This is a sensible exercise since in the 2006-2007 round, the percentage of winners increased from 25% to 35%. This implies that the unconditional probability of winning increases from 25% to 35% (favorable number of cases divided by total number of cases). Therefore, the perceived probability of winning by teachers of a given school necessarily change (We are assuming neither sure losers nor sure winners in the tournament). This will increase the intercept and reduce the slope of the cumulative probability function as shown in Figure 2. The intuition is as follow, the increase in the number of winners increases the probability of winning even with no effort exerted at all. This naturally shifts the intercept of the cumulative probability function up. Now, the probability function is less sensitive to increases in the level of effort since there is less room to improvement because it will integrate eventually to one. In this example, the level of effort decreases from e_1^* to e_2^* and the level of the probability of winning seems unaltered.

In this simple model, we have that the level of effort exerted by school i decreases. This hypothesis can be tested empirically if we have in mind a relationship between teachers effort and students test scores. If we think about a production function that depends on student, teacher, peer, and school effects, it is straightforward to come up with an argument in favor of a positive relationship between students test scores and teacher effort (quality, motivation, etc.).

For instance, $y_i = X\beta + F(e_i) + \epsilon$. Where y_i is the average test score of school i, X are school characteristics and $F(e_i)$ is the level of effort and motivation exerted by teachers in school i. Therefore, we can measure the effect of the SNED on student tests scores before and after the change in the design and test whether the increase in the unconditional probability of winning the award reduces the treatment effect, if at all.

4 Evaluation and Identification Strategy

In order to evaluate the effect of SNED on test scores there is at least three interesting questions to answer. The first question is how competition for the prize increases, if at all, schools' mean test scores. According to the model sketched and 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, increase in participant school's mean test scores.

This question is not trivial to answer given the difficulties faced when trying to identifying a causal relationship. The construction of a valid control group given the design of the program is troublesome. Participating schools in the SNED tournament account for 90% of the total number of schools in Chile (being non-eligible the private fee-paying). It is natural to think that pre-treatment characteristics in a control group from the private fee-paying schools are different from the pre-treatment characteristics of subsidized schools. One plausible alternative is to construct a control group by a matching procedure but a difficulty of a difference-in-difference approach is that the design of the tournament necessarily implies that there are sure losers and sure winners, schools that are always *in the money* (top schools that systematically rank in the upper quartile or so) and also schools that are *out of the money*. Then a reduced (and unknown) number of schools in the experimental sample are indeed affected by the tournament. We propose a simple method to identify losers and winners by estimating the probability of winning the 1996-1997 tournament with pre-tournament data.

A second question is related to the ex-post benefits of winning the prize. How winning the award affect schools' mean test scores ex-post. The reason to think that it may be an effect is related to the "gift exchange" or "reciprocal gifts" theory (Akerlof (1982) and Akerlof (1984)) in which awarded teachers would exert more effort after obtaining the prize as a "gift exchange" with the principal or the community.[5]

Contreras, Flores, and Lobato (2005) attempt to answer this second question. They consider a regression analysis and a difference-in-difference estimator and found a small and positive impact of winning the SNED on future test scores. 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 and its results are threaten by several issues such as omitted variable bias, nonlinearities and mean reversion.[6] On the other hand, the difference-in-difference estimator implemented by the authors relies on an artificial control group (identified by matching on the propensity score), which tries to exploit the fact that some winner schools in one homogeneous group could have lost had they belonged to a different group. This is a creative approach that exploits somehow the "randomness" of being above and below the threshold. The hindrance of this method, however, is that it is not clear what effect is identified and depends heavily on the matching procedure chosen. We show that a Regression Discontinuity approach allows us to identify a causal average treatment effect with relatively weak identifying assumptions even in the presence of unobserved heterogeneity.

Finally, a third question to answer is related to the changes made to the SNED program. In the last round of the SNED program (2006/2007), the fraction of winners was severely increased from about 25% to 35%. This increase in the unconditional probability of winning –as suggested by the model we propose– may have decreased the level effort exerted by teachers of participant schools, and therefore, its impact on test scores.

The change introduced suggests that some schools that were not affected by the tournament before now they are. In theory, about 90% of the schools participate on the tournament but in practice, a small fraction of these 90% will be affected by the tournament since there will be sure losers and sure

[5] Even though the prize may not be considered as a gift by teachers, they are receiving a monetary prize for something they are supposed to do anyway.

[6] For the latter, see Chay, McEwan, and Urquiola (2005)

winners that break up the tournament. The quality of schools affected by the tournament will probably decrease with the new rules and so could be the level of effort. It is expectable to find some sort of substitution among good schools that were affected by the tournament in the past but with the new rules became sure winners, and bad schools that were sure losers in the past that became participant schools. Unfortunately we cannot answer this question with the available data at this time and, therefore, we focus on the first two questions.

4.1 Assessing the incentive effect on test scores: Matching

The first approach to shed light on the tournament effect on test scores is to create a control group from the nonparticipant schools (private fee-paying) and we follow the methodology proposed by Abadie and Imbens (2005). The variables considered for the matching procedure are region, urban/rural status, type (boys, girls, and mixed sex), size, number of teachers, and average parents' education.

We implement an exact matching in the first three variables and compare a subset of the treatment group, which is the private subsidized schools, with the created control group from the private fee-paying schools. The reason of doing so, is to ensure the comparability of schools, at least defined by its ownership nature.

Now we briefly describe the methodology. Let $W \in (0,1)$ be the treatment indicator (equal to 1 if it is a participating school: winner or loser) and let the potential outcomes given by

$$Y_i = \begin{cases} Y_i(0), & \text{if } W_i = 0 \\ Y_i(1), & \text{if } W_i = 1 \end{cases}$$

Let m be an integer representing the number of neighbors that will be used to create a match and $j_m(i)$ the index $j \in \{1, 2, ..., N\}$ that solves $W_j = 1 - W_i$ and

$$\sum_{l:W_l=1-W_i} Ind\{||X_l - X_i|| \leq ||X_j - X_i||\} = m$$

where *ind* is an indicator function that is equal to 1 when the argument is true. In words, this result means: choose from the control group the m closest observations to X_i.

Let $\mathcal{J}_M(i)$ denote the set of indices for the first M matches of unit i : $\mathcal{J}_M(i) = \{j_1(i), ..., j_M(i)\}$

The matching estimator proposed by AI is a nearest-neighbor matching with replacement estimator. It is with replacement since one observation can be used more than once in the construction of the counterfactual.

The matching estimator imputes the missing potential outcome as

$$\hat{Y}_i(0) = \begin{cases} Y_i, & \text{if } W_i = 0 \\ \frac{1}{M}\sum_{j \in \mathcal{J}_M(i)} Y_j, & \text{if } W_i = 1 \end{cases}$$

and

$$\hat{Y}_i(1) = \begin{cases} \frac{1}{M}\sum_{j \in \mathcal{J}_M(i)} Y_j, & \text{if } W_i = 0 \\ Y_i, & \text{if } W_i = 1 \end{cases}$$

leading to the following estimator for the average treatment effect

$$\hat{\tau}_M = \frac{1}{N}\sum_{i=1}^{N}(\hat{Y}(1) - \hat{Y}(0))$$

As discussed previously, the tournament breaks up since we have schools that surely win and surely loss. In order to assess the effect of competition into test scores we have to identify those schools that are affected by the tournament. By doing so, it is possible to identify a causal effect of competing for the prize on test scores. Let us start assuming that there is a group of schools affected by the tournament, i.e. schools that may win or lose with probability around 1/2. We implement the matching procedure for all the schools and for a subset which does not include sure losers and sure winners schools. The strategy to exclude sure losers and sure winners is explained below.

4.1.1 Identifying sure losers and sure winners

As mentioned before the variable W indicates schools that are eligible and influenced by the tournament, i.e. schools whose teachers perceive a positive likelihood of winning and therefore, are willing to exert higher level of efforts to win the prize. Unfortunately we observe just the type of school (private subsidized and private fee-paying) but we do not observe if the schools are truly affected by the tournament. One way of identifying such schools is to estimate the probability of winning with pre-tournament data.

Since we do not know exactly how the SNED index is constructed, we estimate a linear model of the 1996 index on the lagged value of the math test scores and its second difference.

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

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 ingredients of the SNED index we add more controls such geographic region and urban/rural dummies. Then we predict the SNED index and compute in 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. This probability of winning can be used as a weight in the matching algorithm or to select the sample in which the treatment effect will be calculated.

An alternative and complementary strategy way of evaluating the tournament and the presence of sure losers and sure winners is to 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 indicates the presence of sure losers and winners and a tournament effect for at least a sub-population of eligible schools.

In order to do so, 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, parent's 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 losers 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 its prediction 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.

4.2 A RD Approach for the ex-post effect on test scores

In this section we show how the Regression Discontinuity design of the SNED can help us to identify a causal treatment effect of wining the prize on future test scores, what we refer to as a gift-exchange effect. The first known work exploiting this type of discontinuous assignment was Thistlewaite and Campbell (1960) and a growing literature has emerged ever since.[7]

By imposing a relatively weak set of identifying assumptions we show the different causal effects this approach is able to identify. Following Rau (2007) we provide a formal and transparent derivation of the semi-parametric model that has been proposed in the literature to estimate the average treatment effect. (See van der Klaauw (2002), Porter (2003))

[7] See Angrist and Lavy (1999), Chay and Greenstone (2005)

Before proceeding with the derivation remember that the SNED prize assignment follows a discontinuous rule. Top schools in each homogenous group are selected until they account for the 25% of the enrollment in each group. Hence if N_{sg} is the enrollment in school s, in group g, and N_g is the total enrollment in group g, then, there exist a cutoff point of the SNED index in each homogenous such that

$$x_{0g} = argmax_x \left\{ x : \sum_s N_{sg} 1_{\{SNED \geq x\}} \geq 0.25 N_g \right\}$$

and winners in group g are such that $SNED \geq x_{0g}$.

Now, consider the observed average test score y_i for school i

$$y_i = y_{1i} d_i + y_{0i}(1 - d_i)$$

where y_{1i}, y_{0i} are the potential average outcomes, and d_i is an indicator variable for treatment status. Hence, y_{1i} is the outcome when school i receives the SNED award ($d_i = 1$) and y_{0i} is the outcome if not award is received. Let $\alpha_i = y_{1i} - y_{0i}$ be the treatment effect for school i. Rewriting the previous expression we have

$$y_i = y_{0i} + \alpha_i d_i$$

which allows us to rewrite the equation for outcome y_i in a semiparametric representation by taking conditional on $x_i = x$

$$E[y_i | x_i = x] = m(x) + E[\alpha_i d_i | x_i = x]$$

where x represents the SNED score and $E[y_{0i} | x_i = x] = m(x)$. In a *sharp* design, the assignment rule (d_i) depends deterministically on x and is discontinuous at the threshold value x_0. Indeed, the SNED tournament determine the winners by the discontinuous rule: $d_i = 1_{\{x_i > x_0\}}$. Hence, $E(d_i | x_i = x) = \mathrm{P}_r(d_i = 1 | x_i = x)$ will be either 0 or 1.

In a *sharp* design, assuming the common treatment effect ($\alpha_i = \alpha$) it follows that $E[\alpha_i d_i | x_i = x] = \alpha d$. Where we drop the index since d_i is a function of x_i, so d is a function of x. Now, dropping the index for convenience and using $y = E[y|x] + \epsilon$, where $\epsilon = y - E[y|x]$, the following expression is obtained

$$y = m(x) + \alpha d + \epsilon \tag{2}$$

which is the same expression as in van der Klaauw (2002) and Porter (2003). This expression is highly convenient from the econometric point of view since it has been studied since Robinson's (1998) partially linear model.

In equation (2) the parameter of interest is α and not the nonparametric term $m(x)$. Van der Klaauw (2002) refers to $m(x)$ as a *control function*. That might confuse the reader with the notion of control function in endogenous regression. In that case a *control function* transform the problem of endogeneity to a one of omitted variables incorporating a function of residuals from a first stage to the reduced form.

It is important to remark that in this case, $m(x)$ is the conditional expectation of the outcome variable without treatment, y_{0i}, on the selection variable $x_i = x$. But, $m(x)$ is defined in the entire support of x_i, so $m(x)$ includes the counterfactual $E[y_{0i}|x, d = 1]$ since $E[y_{0i}|x] = E[y_{0i}|x, d = 0]\mathrm{P}_r(d = 0|x_i = x) + E[y_{0i}|x, d = 1]\mathrm{P}_r(d = 1|x_i = x)$. In a *sharp* design the probabilities will be either 0 or 1.

Equation (2) is an interesting expression since it links the experimental representation of the response variable (in terms of potential outcomes) with a nonparametric econometric representation. Here, α represents the size of the discontinuity at x_0. A sufficient condition for identification of α, is to assume continuity of $m(x)$ at x_0 and the existence of the limits $\lim_{x\uparrow x_0} E[d_i|x]$ and $\lim_{x\downarrow x_0} E[d_i|x]$. In case of a sharp design, i.e. $\lim_{x\uparrow x_0} E[d|x] = 0$ and $\lim_{x\downarrow x_0} E[d|x] = 1$, it is straightforward to see that α is identified

$$\alpha = \lim_{x\downarrow x_0} E[y_i|x_i = x] - \lim_{x\uparrow x_0} E[y_i|x_i = x] \tag{3}$$

The usual estimators for α have been the Local linear regression (Hahn, Todd, and van der Klaauw (2001)), local polynomials and partially linear models (Porter (2003)) and ordinary polynomials (Lee and DiNardo (2004)).

4.3 Identification of the ATE under Heterogeneity

When the common treatment effect assumption is abandoned, we are still able to identify the average treatment effect (ATE) under the following identifying assumptions. If α_i and d_i are conditionally independent on x, we have that $E[\alpha_i d_i | x_i = x] = E[\alpha_i | x_i = x]d$. Finally, assuming continuity of $E[\alpha_i | x_i]$ at $x_i = x_0$ we have

$$E[\alpha_i | x_i = x_0] = \lim_{x \downarrow x_0} E[y_i | x_i = x] - \lim_{x \uparrow x_0} E[y_i | x_i = x] \tag{4}$$

Note that the conditional independence assumption implies that schools does not self-select into the SNED program due to anticipated gains from it. While this may be an unrealistic assumption since some schools compete to win the bonus, the threshold value of the SNED score is unknown. It is revealed when the tournament ends. Hence, it is less plausible to observe pooling around x_0. Finally, even with prospective gains it is still possible to identify a local average treatment effect (LATE) for schools that their treatment effect changes discontinuously at $x = x_0$ (see Hahn, Todd, and van der Klaauw (2001) for a formal proof.)

4.4 Invariance of the RD estimator under normalization

Since the cutoff points varies among homogeneous groups, we normalize them to 0 in order to get an average treatment effect for the whole subpopulation "around the threshold". It is easy to show that the normalization does not alter the estimation of α.

Consider we want to normalize the cutoff point to 0, hence let $x_i^* = x_i - x_0$. It can be easily proved the invariance of the treatment effect estimator under a RD design. Note that equation (4) is equivalent to

$$\begin{aligned} E[\alpha_i|x_i^* = 0] &= \lim_{x^* \downarrow 0} E[y_i|x_i^* = x^*] - \lim_{x \uparrow 0} E[y_i|x_i^* = x^*] \\ &= \lim_{(x-x_0) \downarrow 0} E[y_i|x_i - x_0 = x - x_0] - \lim_{(x-x_0) \uparrow 0} E[y_i|x_i - x_0 = x - x_0] \\ &= \lim_{x \downarrow x_0} E[y_i|x_i = x] - \lim_{x \uparrow x_0} E[y_i|x_i = x] \\ &= E[\alpha_i|x_i = x_0] \end{aligned}$$

5 Data

This paper uses information from the national SIMCE test. The data sets contain information for the period 1989-2006. Tests are conducted for students attending one of these grades: fourth, eighth or tenth. We have aggregate data, at the school level, from 1989 to 1997. Since 1998 we count with student level data. However, we work with school level data since the tournament is at the school level. SIMCE data sets include information about family and school characteristics.

Table 2 presents the main school characteristics and performance by administrative school dependency: public, private subsidized and private fee-paying. The table summarizes information for the years 1996 and 2006. This table indicates that private fee-paying schools presents higher socioeconomic conditions than private subsidized and public schools. Private fee-paying schools shows the highest household income and parents' education. On the contrary, public schools exhibit the lowest family income and parental education. Consistently, school performance in mathematics and language are lower in public schools with respect to private subsidized and private fee-paying school. Finally, there was a change in the SIMCE scoring scale in 1998 and thereafter. In 1996, the SIMCE test show an average about 70 points with and standard deviation about 10 points. Since 1998, the SIMCE test presents and average of 250 points with a standard deviation of 50 points.

Table 3 summarizes the same variables discussed above for winners and losers. This information is presented for 1996 and 2006. In both years we observe no significant differences in educational performance and socioeconomic characteristics between winner and losers schools. At first sight, it looks like a randomization of the prize 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 the group, differences in performance should be observed between schools in the same homogeneous group. However, by comparing performance between winner and loser 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 never have been awarded with the SNED bonus. Only a small fraction of schools have won the SNED several times. In other words, according to the evidence there are few sure winners, but a significant number of sure losers.[8]

6 Results

In this section we present the results of the evaluation strategies discussed in section 4. We present results for the evaluation of the tournament effect using matching techniques and we present results for the gift-exchange effect implementing a regression discontinuity analysis.

6.1 Tournaments effects

As we mentioned in section 4, in order to evaluate the tournament effect on test scores we create a control group from the nonparticipant schools (private fee-paying) following the methodology proposed by Abadie and Imbens (2005). The variables considered for the matching procedure are region, ru-

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

ral/urban status, type (boys, girls,and mixed sex), size, number of teachers, and average parents' education.

The treatment group was reduced to the private subsidized schools in order to increase comparability between treatment and control groups. We implement an exact matching in the first three variables and compute the average treatment effect (ATE) and the average treatment on the treated effect (ATT).

Tables 5 and 6 present evidence on the tournament effect. The Tables are divided into two panels. The top panel presents the ATE, while the bottom panel summarizes the results of the ATT. The first column, identify the pair of years considered in the difference. The second column, show the ATE coefficient. Columns 3-5 summarize the standard deviation, t-test and number of observation respectively.

The evidence indicates that tournament effect has a positive and significant effect on the overall performance in education. While the ATE parameter fluctuates between 0.2 and 0.3 standard deviation, the ATT coefficient exhibits and impact above 0.30 standard deviation.

In addition, when sure winner and losers are excluded from the sample (Table 6) the ATE and ATT coefficients remains positive, large and significant. This table shows the ATE and the ATT over a reduced sample of schools: those with probability of winning greater than 0.4 and lower than 0.95.

Related to the probability of winning, in Figure 3 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. Then, the first category includes schools with probability of winning between 0 and 0.05, and so on. [9] Then, it can be seen that the tournament seems to affect schools with probability of winning greater than 0.3 or 0.4. This suggest the existence of sure losers, schools that were not affected by the tournament

[9] 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 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.

according to the prediction error.

Now, in order to see if non-eligible schools show the same pattern as in Figure 3, we repeat the exercise for private schools. Then we predict their SIMCE test score for 2006 using pre-treatment information and then we 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 in each group. Figure 4 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 3 is not observed here.

Finally, the size (number of schools) of the homogeneous group could matter if schools in small groups identify themselves as losers or winners and therefore do not exhibit increase in effort. In Figures 5 and 6 we do the same plots including homogeneous groups with more than 25 schools and 25 or less schools respectively. It is interesting to note that while restricting the data to groups with 25 or less schools the tournament seems to fail. The pattern observed in Figures 3 and 5 is not observed in Figure 6.

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.

6.2 Regression Discontinuity

The theoretical model behind this specification is related to the "gift exchange" or "reciprocal gifts" theory. We are interested in testing if teachers would exert more effort after obtaining the prize as a "gift exchange" with the principal or the community. Thus, using the Regression Discontinuity (RD) design of the SNED may help us to identify a causal treatment effect of winning the bonus on future test scores.

The estimation method we use is the local polynomial approach developed by Porter (2003) in which a weighted polynomial is estimated using a kernel as a weighting scheme centered at the discontinuity point. The bandwidth is chosen using generalized cross validation (GCV), hence the bandwidth chosen is the one that minimizes the GCV score that is a *leave-one-out* mean square error.

The evidence presented in this section is based then, on the classical regression discontinuity approach where schools slightly above and below the threshold has more weight (determined optimally according to the GCV criterion). Since the bonus allocation in this neighborhood is mainly random, then the treated and untreated schools around the threshold might be undistinguishable in educational achievement.

Figure 7 shows that, in general, the RD design was a sharp one. The cutoff point is normalized to 0 and we only observe a very small fraction of slippage in the 1998-1999 tournament. The rest of the tournaments were implemented with a sharp design. The 2004-2005 and 2006-2007 rounds were omitted in this diagram but also show a sharp design.

The results are presented in Tables 7-11 for different years. The evidence suggests that monetary incentives to teachers exhibit a positive, small and no significant effect on student achievement. Both, in years 2004 and 2006 the effects of the SNED incentive is slightly significant (at the 10% level) with a magnitude of 10% of one standard deviation. Overall, the evidence do not support the hypothesis of gift-exchange effect.

7 Conclusion

Although performance-related pay for teachers is being introduced in many developed countries, little evidence has been provided based on measured effects in LDCs. This article contributes with empirical evidence on the effects of performance-related incentive pay for teacher on school academic performance. We examine the effect of a rank-order tournament, the SNED, on standardized test scores distinguishing two types of effects: the tournament effect, the effect of the introduction of the tournament on eligible schools: winners and losers; and the gift-exchange effect, the effect of winning the prize on future test scores.

Matching in characteristics and Regression Discontinuity analysis are used to examine the tournament effect and the gift-exchange hypothesis respectively. We find a positive tournament effect and no evidence of gift-exchange effect. Since the tournament effect evaluated is the effect of introducing the tournament, we cannot extend this result to the following rounds of the SNED. For the ex-post or gift exchange effect we find a small but insignificant effect when analyzed all schools.

The empirical evidence presented in this paper provides support for educational policies oriented towards more differentiation in the salary structure for teachers. In many countries where teachers unions are very important (in particular in LA and LDCs), a wage structure which recognizes pay-for-productivity is theoretically efficient. This paper provides evidence supporting a wage structure for teachers more related to productivity as a mechanism to increase children's achievement. However, this paper also shows that this type of tournaments breaks up given the existence of sure losers. For the case of Chile, practically half of eligible schools have never won the award after eleven years of implementation.

References

Akerlof, G. A. (1982): "Labor Contracts as Partial Gift Exchange," *The Quarterly Journal of Economics*, XCVII(4).

——— (1984): "Gift Exchange and Efficiency-Wage Theory: Four Views," *American Economic Review*, 74(2), 79–83.

Angrist, J., and . Lavy (1999): "Using Maimonides' Rule To Estimate The Effect Of Class Size On Scholastic Achievement," *The Quarterly Journal of Economics*, 114(2), 533–575.

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.

Chay, K., and M. Greenstone (2005): "Does Air Quality Matter? Evidence from the Housing Market," *Journal of Political Economy*, 105(2).

Chay, K. Y., P. J. McEwan, and M. Urquiola (2005): "The Central Role of Noise in Evaluating Interventions That Use Test Scores to Rank Schools," *American Economic Review*, 95(4).

Contreras, D., L. Flores, and F. Lobato (2005): "Monetary incentives for teachers and school performance: Evidence for Chile," *Manuscript. Department of Economics, Universidad de Chile.*

Hahn, J., P. Todd, and W. van der Klaauw (2001): "Identification and Estimation of Treatment Effects with a Regression Discontinuity Design," *Econometrica*, 69(1), 201–209.

Holmstrom, B. (1982): "Moral hazard in teams," *Bell Journal of Economics, The Rand Corporation*, 13(2), 324–340.

Kandel, E., and E. Lazear (1992): "Peer Presure and Partnerships," *Journal of Political Economy*, 100, 801–817.

Lazear, E. P., and S. Rosen (1981): "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89(5), 841–864.

Lee, D. S., and J. DiNardo (2004): "Economic Impacts of New Unionization on Private Sector Employers: 1984-2001," *Quarterly Journal of Economics*, 109(4), 1383–1441.

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

Porter, J. (2003): "Estimation in the Regression Discontinuity Model," *Manuscript. Department of Economics. Harvard University.*

Robinson (1998): "Root-N-Consistent Semiparametric Regression," *Econometrica*, 56, 931–954.

Rosen, S. (1986): "Information, Competition, and Markets," *American Economic Review*, 76(4), 701–715.

Thistlewaite, D., and D. Campbell (1960): “Regression-Discontinuity Analysis: An Alternative to the Ex-Post Facto Experiment,” *Journal of Educational Psychology*, 51, 309–317.

van der Klaauw, w. (2002): “Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach,” *International Economic Review*, 43(4).

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 Shooling 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 Shooling of Mothers	-	-	-	4.01	3.20	2.72
	-	-	-	0.18	0.70	0.52
Average Shooling 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)
Shools 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 Shooling of Parents	2.38	2.33	3.04	3.15
	(0.67)	(0.59)	(0.58)	(0.61)
Average Shooling of Mothers			3.03	3.16
			(0.59)	(0.61)
Average Shooling of Fathers			3.05	3.15
			(0.59)	(0.60)
Average household income			232262.00	250254.10
			(159815.90)	(159661.40)
Shools 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, Abadie-Imbens matching

Difference	ATE	SD	T-test	Obs
96-95	0.29	0.06	4.91	1671
97-95	0.31	0.05	6.50	1635

Difference	ATT	SD	T-test	Obs
96-95	0.32	0.07	4.54	1671
97-95	0.32	0.06	5.69	1635

Table 6: Tournament effects, Abadie-Imbens matching*

Difference	ATE	SD	T-test	Obs
96-95	0.18	0.08	2.29	1136
97-95	0.28	0.06	4.38	1114

Difference	ATT	SD	T-test	Obs
96-95	0.30	0.09	3.12	1136
97-95	0.32	0.07	4.38	1114

*No sure winners and no sure losers

Table 7: RD Results: SNED1996/1997 on 1998 Scores

Coefficient	Estimate	Standard error	t-test
intercept	-0.18	0.14	-1.29
α	0.05	0.19	0.24
β_1	2.77	11.13	0.25
β_2	78.77	178.47	0.44
β_3	3814.39	12061.93	0.32
R2	0.014		
Bandwidth	0.037		
GCV	1.7E-05		

Table 8: RD Results: SNED1998/1999 on 2000 Scores

Coefficient	Estimate	Standard error	t-test
intercept	-0.65	0.09	-7.46
α	0.09	0.15	0.63
β_1	-0.21	1.23	-0.17
β_2	-0.06	3.12	-0.02
β_3	45.12	44.40	1.02
R2	0.009		
Bandwidth	0.186		
GCV	9.3E-07		

Table 9: RD Results: SNED 2000/2001 on 2002 Scores

Coefficient	Estimate	Standard error	t-test
intercept	0.02	0.06	0.31
α	0.14	0.10	1.48
β_1	-0.07	1.08	-0.07
β_2	-2.48	3.47	-0.71
β_3	16.43	46.75	0.35
R2	0.010		
Bandwidth	0.186		
GCV	8.10E-07		

Table 10: RD Results: SNED 2002/2003 on 2004 Scores

Coefficient	Estimate	Standard error	t-test
intercept	-0.09	0.08	-1.21
α	0.19	0.11	1.77
β_1	-0.14	0.16	-0.92
β_2	-0.03	0.06	-0.46
β_3	0.07	0.12	0.59
R2	0.005		
Bandwidth	1.414		
GCV	8.38E-07		

Table 11: RD Results: SNED 2004/2005 on 2006 Scores

Coefficient	Estimate	Standard error	t-test
constante	-0.0768	0.07079	-1.08
α	0.17737	0.10634	1.67
β_1	-0.17572	0.15816	-1.11
β_2	-0.00733	0.06616	-0.11
β_3	0.16614	0.11696	1.42
R2	0.007		
Bandwidth	1.414		
GCV	8.72E-07		

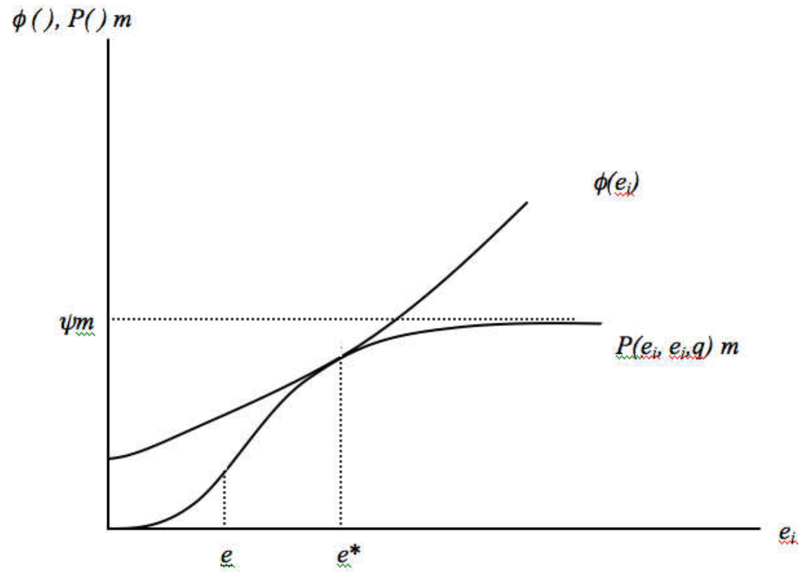

Figure 1: Equilibrium in the Model

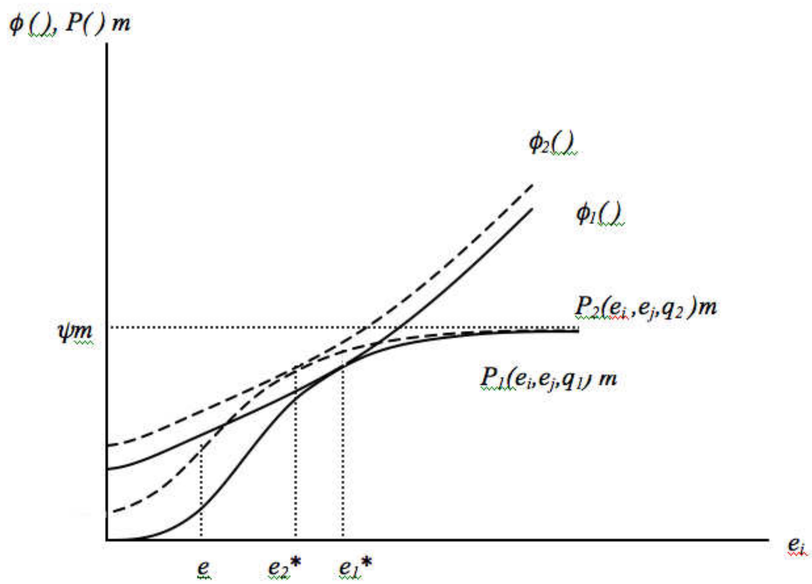

Figure 2: Effect of an increase in the unconditional probability of winning

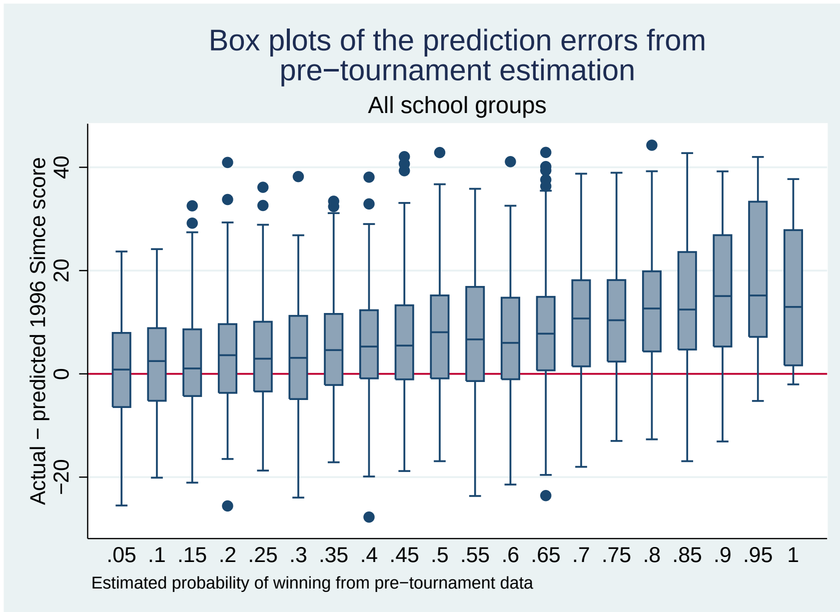

Figure 3: Box plots of the test score prediction errors across probability of winning groups: All eligible schools

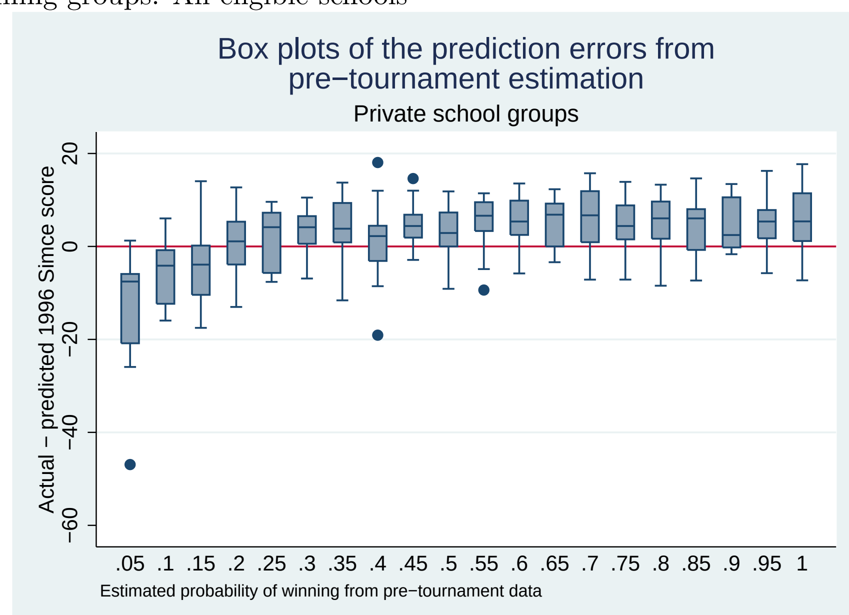

Figure 4: Box plots of the test score prediction errors across probability of winning groups: Non-eligible Schools

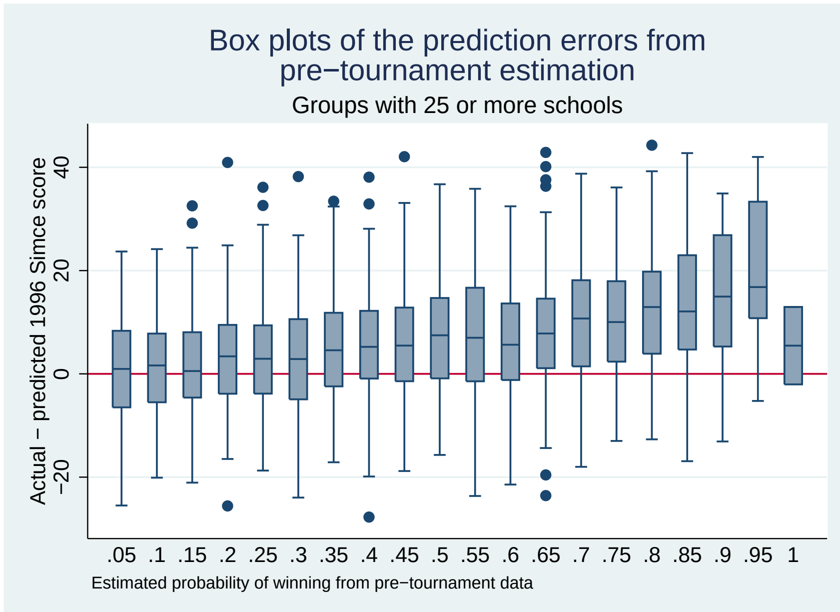

Figure 5: Box plots of the test score prediction errors across probability of winning groups

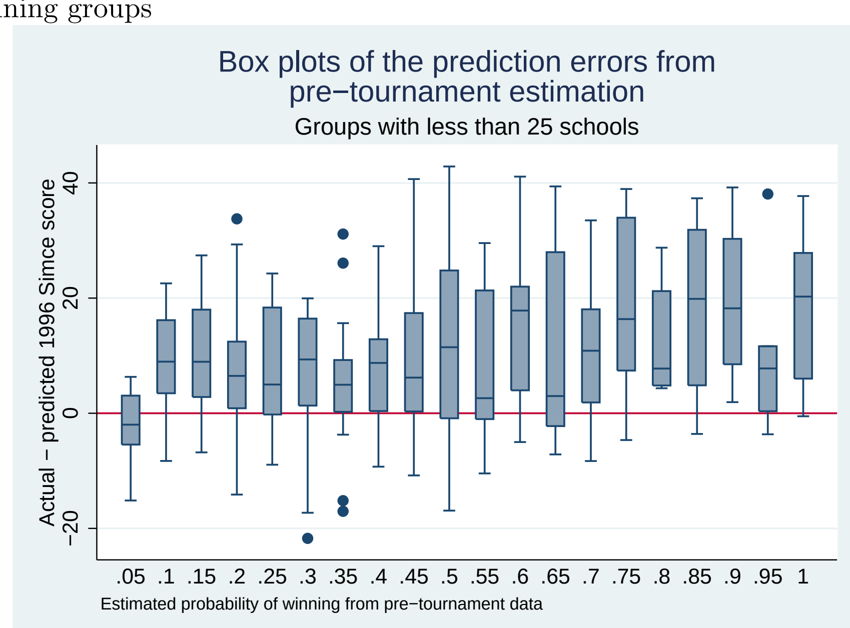

Figure 6: Box plots of the test score prediction errors across probability of winning groups

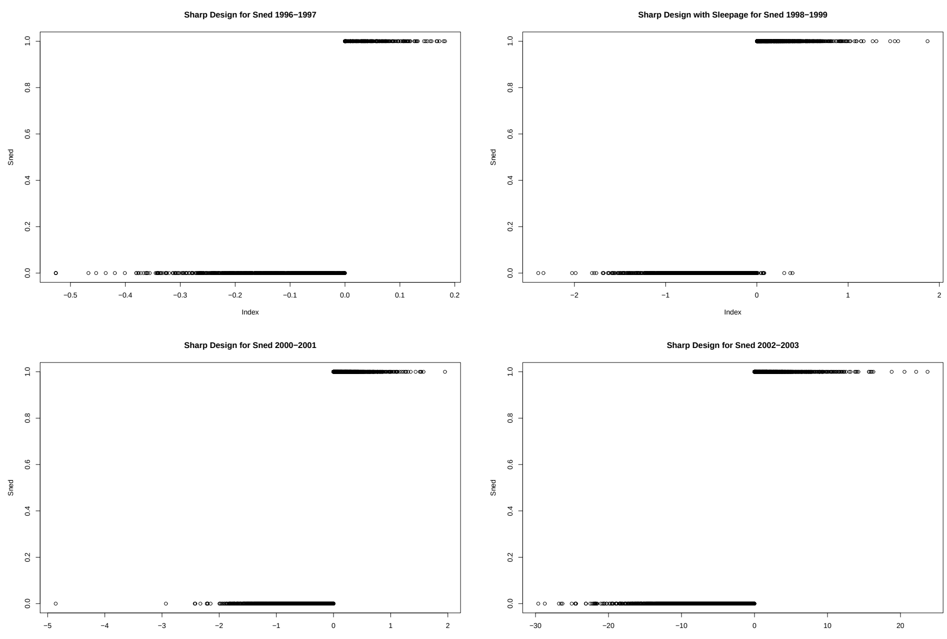

Figure 7: Regression Discontinuity Design for SNED