Insper IBMEC São Paulo

Free-Rider and Gaming in the Distribution of Revenues across Schools

Eduardo de Carvalho Andrade

Mauricio Bugarin

Bruno Rocha

Fábio Miessi

Insper Working Paper

WPE: 163/2009

Free-Rider and Gaming in the Distribution of Revenues across Schools

Eduardo de Carvalho Andrade[ε]

Mauricio Bugarin[μ]

Bruno Rocha[β]

Fábio Miessi[φ]

November 28th, 2008

Abstract

This paper analyzes teachers' behavior in Brazilian public schools after the introduction of the Fundef in 1998. The model predicts that: (i) teachers engaged in gaming, by adjusting the fail and repetition rates to affect the number of students and, consequently, their wages; (ii) the degree of this opportunistic behavior decreases with the number of schools in a municipality due to the free-rider problem. The empirical investigation corroborates these predictions. In particular, the change in the repetition rate after the Fundef ranged from -12% (1st grade) to +38% (8th grade). Finally, there was a fall in students' proficiency that could be associated with a fall in the educational standard.

[ε]Ibmec São Paulo, Rua Quatá 300, Vila Olímpia, 04546-042, São Paulo-SP, eduardo.andrade@isp.edu.br.
[μ]Ibmec São Paulo, Rua Quatá 300, Vila Olímpia, 04546-042, São Paulo-SP, bugarin@isp.edu.br.
[β]Tendências Consultoria Integrada, Rua Fradique Coutinho 465/36, Pinheiros, 05416-010, São Paulo-SP, brunor@usp.edu.br.
[φ]London School of Economics, Flat 4, The Wheel House, 1 Burrells Wharf Square, E14 3TA London, f.a.sanches@lse.ac.uk.

1) **Introduction**

In 1998, the Brazilian Congress approved an educational funding program called Fundef (Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério). It has been in place since its implementation in 1999. Among others, it has the following main characteristics[1]. First, it stipulates that each Brazilian State forms a fund that incorporates 15% of all tax revenues[2]. These resources have to be invested in primary and secondary educations. Each municipality (or the state government) receives a fraction of this fund based on the number of students enrolled in its schools. Finally, 60% of these resources have to be directed to pay teachers' wages.

Hence, the Fundef indirectly linked teachers' wages to the number of students registered in the schools. Therefore, it may have given the teachers incentives to take actions in order to influence their school's enrollment. This paper's objective is precisely to develop both a theoretic and an empirical analysis to assess this possibility.

Initially, we build a model in order to predict teachers' behavior before and after the introduction of Fundef. We find that, after the Fundef, they are likely to engage in opportunistic behavior or gaming, by adjusting the fail and repetition rates[4], in order to affect the number of students and, consequently, their wages. In fact, we show that there is an optimal fail and repetition rates that maximizes teachers' utility. On the one hand, teachers have incentives to increase the fail and repetition rates in order to maintain a

---

[1] See Semeghini (2001) and Souza (2005) for more details on the Fundef program.

[2] These taxes are state and municipal taxes. Brazil is a federation formed by 27 states plus the Federal District. Each state is divided into municipalities. In each state, both state and municipal schools coexist. Each jurisdiction has its own wages's policy but since 1999 they have to follow the outlines defined in the Fundef.

[4] The repetition rate is the fraction of students who fail the grade and decide to repeat it.

student for a longer period in the school. On the other hand, this strategy has its limitations as a student who considers that his probability of failing a grade is high, may end up abandoning the school, which reduces the number of students, i.e., there is a connection between the fail and the dropout rates[5]. Combining these two effects, we predict that teachers' optimal strategy is to impose a relatively lower fail and repetition rates in the initial grades and a relatively higher ones in the final grades.

Moreover, we show that this opportunistic behavior may vary in intensity depending on the number of schools (or students) in a given location or municipality. The reason is the following. When there are many schools (and students), teachers' behavior in one school has a negligible effect on the amount of resources directed to this location, and consequently to their wages. This is the classical free-rider problem[6]. Hence, our model predicts that this opportunistic behavior is more likely to occur the lower is the number of schools (or students) in one location, say, a municipality.

Next, we perform an empirical investigation to check some theoretical predictions. We employ the Difference-in-Difference method, first described in Card (1990) and analyzed in Angrist and Krueger (1999). The control and treated groups are, respectively, the private and public schools. The results seem to corroborate the predictions of the model. In comparison with the control group, the repetition rate is lower in the initial grades (1st to 4th grades) and higher in the last grades (5th to 8th grades). The magnitude of the effect varies from grade to grade, and it increases monotonically from the first to the last grade of the fundamental education. In proportion to the repetition rate in 1997 (the year before the Fundef), the Fundef's effect ranges from a drop of -11.65% in the first grade to a rise of 37.95% in the eighth grade.

---

[5] Ribeiro (1992) analyses and discusses the relationship between the fail and the dropout rate.

[6] The classic reference on this issue is Olson (1965).

The free-rider effect, however, is not very strong. But we still find that, due to the Fundef's effect, in municipalities with relatively small number of schools, the repetition rate is relatively lower and greater in public schools, respectively, in the first and the last four grades.

Finally, we investigate the more likely mechanism employed by teachers to affect the repetition rate. One possible mechanism is the following. Given the educational standard, an improvement in the quality of the education can reduce the repetition rate. Alternatively, given the quality of the education, the reduction in the standard can also reduce the repetition rate.

The empirical evidence suggests that there was a drop in the students' proficiency in mathematics. It occurred both in the fourth grade (four years after the introduction on the Fundef) and in the eighth grade (eight years after the introduction of the Fundef). Hence, there was no indication of an improvement in the quality of the education due to the Fundef. As an alternative explanation for a fall in the students' performance, one can not discard the possibility that teachers reduce the educational standard in order to adjust the repetition rate.

This paper is related to the literature that investigates how teachers react to the incentives of a new educational policy. In particular, there are many studies analyzing how teachers respond to the implementation of a policy adopted in many states in the US in which they receive bonuses or sanctions based on the performance of their students on standardized tests. These studies evaluate if teachers either increase the fail and dropout rates or send more students to special education placement to avoid the relatively bad students to take the test and then reduce their students' average grades. Some examples are: Jacob (2004) in Chicago, Koretz and Barron (1998) in Kentucky, Haney (2000), Carnoy et. al (2001) and Toenjes and Dworkin (2002) in Texas, Figlio and Getzler (2002) in Florida and Carnoy and Loeb (2002) and Hanushek and Raymond (2004) in the US.

The novelty in our paper is to investigate opportunistic behavior by teachers due to the introduction of Fundef in a context in which there is also a free-rider problem. To our knowledge, there is only one paper that analyzes the effects of Fundef by Menezes-Filho and Pazello (2004). Although they estimate its effects on the wages of the teachers in the public schools and the proficiency of the public school students, they do not deal with these incentive issues.

This paper is divided into six sections, including this introduction. Section 2 develops the basic choice model for the before-Fundef case. Section 3 extends the basic model in order to assess the new incentives induced by Fundef, and derives the basic predictions of the model. Sections 4 and 5 present the empirical strategy and the database to analyze the effects of the Fundef, respectively, on the repetition rate and the students' proficiency level. The last section concludes.

2) The Basic Model Before the Fundef

This section builds a basic model in order to predict teacher's behavior before the introduction of Fundef.

There are *N* equal schools/teachers in location *J*. For our purposes, it is sufficient to analyze the problem of a representative teacher/school in location *J*. Representative teacher *j*'s objective is to maximize his utility function, which depends on his wage $\left(W^j\right)$ and leisure time $\left(L^j\right)$[7]. Under the policy before the introduction of Fundef, teacher's wage was basically a function of his tenure and his highest academic degree.

[7] Throughout the analysis, we use "teacher" and "school" interchangeably. "Teacher" represents the group of teachers and directors in the school. The idea is that this "teacher" decides the amount of effort and its allocation in different activities of all school personnel.

In contrast with the Fundef, the number of students enrolled in the school did not influence teacher's wage. In fact, teacher's behavior or performance in the classroom did not affect his wage. Therefore, we assume that teacher's wage is fixed and is equal to $\overline{W}$ in all schools.

School j offers two grades. In the beginning of the period, there are enrolled $\overline{N}$ new students in the school in the first grade. Teacher j can work to influence the number of students enrolled in his school in the end of the period by adjusting the fail rate. In order to decrease the fail rate, teacher can, for example, decrease the requirements necessary for a student to pass the grade or increase the quality of the education, which make more students able to pass a given standard.

Independently of his choice, we assume that a teacher needs to make an effort to adjust the fail rate. Effort, however, is costly as it reduces his leisure time. Let E_g^j be the amount of effort devoted by teacher j to affect the number of students in grade g ($g = 1,2$) and f_g be the fail rate in grade g (g=1,2). The fail rate function when teacher dedicates effort E_g^j is equal to $f_g\left(\delta; E_g^j\right) = \bar{f} + \delta h\left(E_g^j\right)$. The parameter δ may take one of the two values δ=1 or δ=−1, according to the teacher's goal. If the teacher makes the effort E_g^j in order to increase the fail rate, then δ=1. Conversely, if he makes the effort in order to decrease the fail rate, then δ=−1. The function h is nonnegative, strictly increasing, strictly concave, with $h(0)=0$. Moreover, we assume the following feasibility conditions[9]: $0 \leq \bar{f} \leq 1$ and $h(1) \leq \min\left\{\bar{f}, 1-\bar{f}\right\}$.

It is noteworthy that $f_g(\delta;0) = \bar{f}$, i.e., $\bar{f}$ is the fail rate when the teacher dedicates no effort to affect the fail rate. Therefore, we call $\bar{f}$ the "natural" fail rate.

[9] These conditions ensure that $0 \leq f\left(\delta; E_g^j\right) \leq 1$ for every effort choice.

Note, moreover, that the expression for *f* incorporates the simplifying assumption that the effect of effort on increasing or decreasing the fail rate is symmetric, which clearly needs not be the case. A more general form makes calculations more confusing without adding any insights.

The teacher's goal is to choose the optimal amount of effort. Then teacher j solves the following problem:

$$\max_{\{\delta; E_1^j; E_2^j\}} U\left(\overline{W}; L^j\right),$$

where: $\overline{W}$ is given; $\delta \in \{-1,1\}$; $0 \le E_1^j, E_2^j \le 1$; $E_1^j + E_2^j \le 1$; $E_1^j + E_2^j + L^j = 1$.[10]

The model assumes that U is a strictly increasing, strictly concave, twice continuously differentiable function with strictly positive cross partials: $U_{12} = U_{21} > 0$. This last assumption simply states that the higher your wage, the more you are able to enjoy an extra unit of leisure time, and the more leisure time you have the more you are able to enjoy and extra unit of wage. Note that the model normalizes leisure time to be in the interval [0,1].

Note that, since the teacher's wage $\overline{W}$ is fixed, the only effect of effort in the objective function is to reduce leisure time. Therefore, the problem has a trivial corner solution: $\overline{E}_1^j = \overline{E}_2^j = 0$. Any choice for δ is optimal and ineffective, since the teacher chooses to exert no effort to manipulate the fail rate. The resulting equilibrium fail rate is the natural rate $\bar{f}$.

Note that the solution to the teacher's problem does not depend on the symmetry hypothesis about the effect of effort on the fail rate. That hypothesis will simplify

[10] For simplicity, we assume that a teacher can use his time either to affect the fail rate or to leisure.

calculations in the next section, where the new incentives brought about by the Fundef program is modeled.

3) The Fundef's New Incentives

As discussed in the previous section, before the introduction of Fundef in 1999, teacher's behavior in the classroom did not affect his wages. In contrast, the Fundef added a new feature in the teacher's wage's policy.

First, it established that the amount of resources directed to each location is a function of the number of students enrolled in the schools in this location. Moreover, 60% of all of these resources had necessarily to be directed to pay teacher's wage. Therefore, as the Fundef indirectly linked teacher's wage to the number of students registered in his school, it created an incentive for teachers to affect this number. We then modify the model in the previous section in order to incorporate this new feature in the teacher's problem.

Teacher *j*'s wage $\left(W^{j}\right)$ is now a function of the number of students enrolled in the schools in location *J*. Then, we can write the following wage function $W^{j} = R\left(S^{j}\left(E_1^{j}, E_2^{j}\right) + \sum_{i \neq j} S^{i}\right)$ where: *R* is assumed to be a twice continuously differentiable and strictly concave function; $S^{j}\left(E_1^{j}, E_2^{j}\right)$ is the number of students in school *j*, which depends on the teacher's effort in both grades; and $\sum_{i \neq j} S^{i}$ is the number of students registered in other schools in location *J*.

Note that $\sum_{i \neq j} S^i$ is a given number for teacher *j*, as he can not affect the number of students enrolled in other schools in the same location[11]. Moreover, we can write as: $S^j(E_1^j, E_2^j) = s_1(E_1^j) + s_2(E_1^j, E_2^j)$, where s_g refers to enrollment in grade *g*=1,2.

In addition to the notation defined before, let r_g and d_g be, respectively, the repetition and dropout rates in grade *g*[13]. Then, obviously, $f_g = r_g + d_g$. All these three variables are a function of the efforts employed by teacher *j* in their respective grades. Note that $0 \leq f_g(E_g^j),\ r_g(E_g^j),\ d_g(E_g^j) \leq 1$.

At the end of the period, we have the following situation. The number of students enrolled in grade $g = 1$ is equal to $s_1(E_1^j) = \overline{N} + \overline{N} r_1(E_1^j) = \overline{N}[1 + r_1(E_1^j)]$, that is, the new students who join the school $(\overline{N})$ plus the students who fail the first grade and decide to repeat it. The number of students enrolled in grade $g = 2$ equals $s_2(E_1^j, E_2^j) = \overline{N}(1 - r_1(E_1^j) - d_1(E_1^j)) + \overline{N} r_2(E_2^j) = \overline{N}[1 - r_1(E_1^j) - d_1(E_1^j) + r_2(E_2^j)]$. The first term corresponds to the students admitted into second grade and the second term is the students who fail the second grade and decide to repeat it.

We make two additional simplifying assumptions. The first one is that the repetition rate is a fixed proportion, λ, of the fail rate. Therefore, the dropout rate is also a fixed proportion, (1–λ), of the fail rate, i.e., a fixed fraction $(1 - \lambda)$ of those students who fail a grade decides to abandon school. So that we can write: $r_g(E_g^j) = \lambda f_g(E_g^j)$ and $d_g(E_g^j) = (1 - \lambda) f_g(E_g^j)$. This assumption tries to capture the empirical evidence in which the higher is the fail rate, the greater is the dropout rate[14].

---

[11] Parents enroll their children in the public school closer to their home.
[13] These are the rates in school *j*. The superscript is omitted for simplicity.
[14] See Ribeiro (2000) for empirical evidence supporting that hypothesis.

Using the above assumption, we can define a natural repetition rate ($\bar{r}$). It is the one that prevails when teacher's effort is zero and the fail rate is equal to the natural one. Formally, we have: $\bar{r} = \lambda \bar{f}$.

Also using the above assumption, the total enrollment in school j can be rewritten in the following way:

$$S^j\left(E_1^j, E_2^j\right) = s_1\left(E_1^j\right) + s_2\left(E_1^j, E_2^j\right)$$

$$= \overline{N}\left[1 + \lambda f_1\left(E_1^j\right)\right] + \overline{N}\left[1 - \lambda f_1\left(E_1^j\right) - (1-\lambda) f_1\left(E_1^j\right) + \lambda f_2\left(E_2^j\right)\right]$$

$$= \overline{N}\left[2 - (1-\lambda) f_1\left(E_1^j\right) + \lambda f_2\left(E_2^j\right)\right]$$

The second assumption is that the function that measures the effect of the teacher's effort on the fail rate takes the specific exponential form $h\left(E_g^j\right) = k\left(E_g^j\right)^{\alpha}$, where $\alpha \in [0,1]$ and $k < \max\left\{\bar{f}, 1 - \bar{f}\right\}$ is a nonnegative parameter[15]. Note that under this assumption the fail rate can be written as $f_g\left(\delta_g, E_g^j\right) = \bar{f} + \delta_g h\left(E_g^j\right) = \bar{f} + \delta_g k\left(E_g^j\right)^{\alpha}$. Thus, total school enrollment becomes:

$$S^j\left(E_1^j, E_2^j\right) = s_1\left(E_1^j\right) + s_2\left(E_1^j, E_2^j\right) = \overline{N}\left\{2 - (1-\lambda)\left(\bar{f} + \delta_1 k\left(E_1^j\right)^{\alpha}\right) + \lambda\left(\bar{f} + \delta_2 k\left(E_2^j\right)^{\alpha}\right)\right\}$$

The teacher's objective is, again, to choose the optimal amount of effort. Given the new incentives introduced by Fundef, teacher j solves the following problem:

$$\max_{\left\{\delta_1; \delta_2; E_1^j; E_2^j\right\}} = U\left[R\left(\overline{N}\left\{2 - (1-\lambda)\left(\bar{f} + \delta_1 k\left(E_1^j\right)^{\alpha}\right) + \lambda\left(\bar{f} + \delta_2 k\left(E_2^j\right)^{\alpha}\right)\right\} + \sum_{i \neq j} S^i\right); 1 - E_1^j - E_2^j\right]$$

Note that, as $0 < \lambda < 1$, the optimum fail rate differs across grades. Indeed, the term $-(1-\lambda)\left(\bar{f} + \delta_1 k_1\left(E_1^j\right)^{\alpha}\right)$ in the above objective function is related to the number of

---

[15] This functional form leads to a closed-form solution. In addition, the restriction on k ensures that $0 \le f_g\left(\delta; E_g^j\right) \le 1$ for every effort choice for the present specification of the fail rate.

students in the end of the period that is influenced by the effort in the first grade. Note that the coefficient of the fail rate in that term is negative. Therefore, in order to maximize the objective function, teacher j is better off choosing $\delta_1 = -1$, as R and U are both strictly increasing in their respective arguments. The converse occurs in the second grade, since the term $\lambda\left(\bar{f} + \delta_2 k_2 \left(E_2^j\right)^\alpha\right)$ has a positive coefficient; therefore, his optimal choice is $\delta_2 = 1$. This remark already highlights an asymmetry that was not present in the original model, which is caused by the Fundef program.

Plugging in the optimal choices for δ_1 and δ_2, teacher j's problem in the presence of Fundef becomes:

$$\max_{\{E_1^j;E_2^j\}} = U\left[R\left(\overline{N}\left\{2-(1-\lambda)\left(\bar{f} - k\left(E_1^j\right)^\alpha\right) + \lambda\left(\bar{f} + k\left(E_2^j\right)^\alpha\right)\right\} + \sum_{i \neq j} S^i\right); 1 - E_1^j - E_2^j\right]$$

The propositions below characterize the teacher's optimal effort choice in presence of Fundef. All proofs are presented in the Appendix 1.

Proposition 1: Let E_1^*, E_2^* be the solution to the teacher's problem in the presence of Fundef. Then $E_2^* = \left(\frac{\lambda}{1-\lambda}\right)^{\frac{1}{1-\alpha}} E_1^*$. Moreover, since U is a strictly increasing function, $E_1^* > 0$ and $E_2^* > 0$.

Proposition 1 indicates that teacher j makes effort to influence the fail rate, and consequently the repetition and dropout rates, in both grades. In other words, according to the theoretical prediction, the Fundef gives incentives for teachers to engage in opportunistic behavior, through gaming, in order to affect their wages. The way of doing this is by affecting the fail rate and, consequently, the number of students enrolled in their schools, which ultimately determines their wages.

Note that, since δ_1=–1 whereas δ_2=1, the effort in the first period is aimed at decreasing the fail rate (and consequently the repetition rate), whereas the effort in the second period is aimed at increasing the fail rate (and consequently the repetition rate). Therefore, Corollary 2 comes as a direct consequence of the above proposition[16]. It states that, in comparison with the natural rate, the optimum fail (repetition) rate is lower in the first grade and it is higher in the second grade. The intuition is simple. To avoid students to dropout in the first grade and lose students in the second grade, the teacher reduces the fail rate (and consequently the repetition rate) in the first grade relatively to the natural rate. In the second grade, the only way to retain part of the students is by increasing the fail rate (and consequently the repetition rate) as they leave school if they pass the grade.

Corollary 2: The optimum fail (repetition) rate in the first grade, f_1^* (r_1^*) is lower than the natural rate $\bar{f}$ ($\bar{r}$). Conversely, the optimal fail (repetition) rate in the second grade f_2^* (r_2^*) is greater than the natural one. Therefore, $f_1^* < \bar{f} < f_2^*$ and $r_1^* < \bar{r} < r_2^*$.

The result in the next corollary is also a straightforward one. The assumption made previously that a fixed fraction $(1-\lambda)$ of those students who fail a grade decides to abandon school, together with the result in Corollary 2, lead to the obvious conclusion that the dropout rate in the second grade is greater than the one in the first grade.

Corollary 3: The optimal dropout rate in the second grade (d_2^*) is greater than the optimal one in the first grade (d_1^*).

16 The proofs of the Corollaries are straightforward and therefore they are omitted here. They are available upon request to the authors.

The previous results indicate that gaming may be a problem (or a side-effect) caused by the introduction of the Fundef. We now analyze a second possible effect related to the Fundef, that is, the free-rider effect.

We saw above that it is in the teacher's interest to adjust the fail rate in order to affect the number of students enrolled in his school, which ultimately affects his wages. However, this effect may be mitigated if the number of schools (and students) in the location is too big. The reason is that when this number is big, teacher's behavior in one school has a negligible effect on the amount of resources directed to his school's location, and consequently to his wage. This is the classical free-rider problem. Therefore, the opportunistic behavior through gaming is less likely to occur in one location the greater is the number of schools (and students) there.

In fact, the following proposition shows that, indeed, teacher *j* exerts a lower effort the greater is the number of students enrolled in location *J*.

Proposition 4: The greater is the number of students in other schools in the municipality or state j, $\sum_{i \neq j} S^i$, the lower is the optimum effort teacher *j* exerts in each grade, E_1^*, E_2^*, in order to affect the fail (repetition) rate, that is, $\frac{\partial E_g^*}{\partial \sum_{i \neq j} S^i} < 0$, *g*=1,2.

Proposition 4 states that the less important the free-rider effect, i.e., the lower $\sum_{i \neq j} S^i$, the greater the effort a teacher makes in both grades to affect the fail rate (and consequently the repetition rate). Hence, the higher the gap between the natural fail (repetition) rate and the fail (repetition) rates in each grade. As a straightforward consequence, the less important the free rider problem, the greater the difference in the fail, repetition and dropout rates in grades 1 and 2. This is precisely the result in the following two corollaries.

Corollary 5: The less important the free-rider incentive, that is, the lower is $\sum_{i\neq j} S^i$, the greater the difference in the fail (repetition) rate in both grades.

Corollary 6: The less important is the free-rider incentive, that is, the lower is $\sum_{i\neq j} S^i$, the greater is the difference in the dropout rate in both grades.

The above analysis indicates that one should expect teachers to engage in opportunistic behavior after the introduction of the Fundef. The theoretical model makes some predictions about the effects of the Fundef that can be tested empirically. First, it suggests that the fail, repetition and dropout rates should fall in the first grades of the fundamental education – due to the gaming effect. Second, it suggests that the same rates should increase in the last grades of the fundamental education – also due to the gaming effect. Finally, it suggests that these two previous effects should be less important in locations with many schools (and students) – due to the free-rider effect.

The theoretical work assumes a proportional relation between the repetition, the fail and the dropout rates; therefore, anyone of those three rates could be used in our empirical analysis. In the next section, we use the available repetition rate to test the theoretical implications of the model.

4) Empirical Analysis: Repetition Rate

4.1) Empirical Strategy and Database

We employ the Difference-in-Differences (DID) method[17] to test the impacts of the Fundef on the students' repetition rates. The treatment group is formed by the public schools that suffer the impact of Fundef. The control group is formed by the private schools, which are not influenced by the new policy.

[17] See Card (1990) and Angrist and Krueger (1998) for the details of the method.

The basic model is the following:

$$y_{it} = \alpha + \beta_1 d_t + \beta_2 d_{pub} + \delta_1 (d_t d_{pub}) + \beta_3 X_{it} + \varepsilon_{it}\,, \qquad (1)$$

where: the dependent variable, y_{it}, is the repetition rate in school i, year t; d_t is a dummy variable which is equal to one after the introduction of the Fundef and zero otherwise; d_{pub} is a dummy equal to one if the school i is a public school and zero if it is a private school; X_{it} is a vector of exogenous control variables; ε_{it} is the error term; and α, β_i ($i = 1,2,3$), and δ_1 are the coefficients.

The parameter of interest is δ_1, which captures the effect of Fundef on the dependent variable. As it is standard when this method is used, treatment and comparison groups should show the same time trend had the policy change not occurred. In terms of our model, this is equivalent to say that the unobservable characteristics of the public schools vary (before and after the introduction of the Fundef) exactly as the unobservable characteristics of the private schools. To minimize this particular problem, the set of control variables X_{it} are introduced in the model.

In order to capture the existence of the free-rider effect discussed in the theoretical model, we alter the basic empirical model (1). We add a new explanatory variable, which is an interaction between the term $(d_t d_{pub})$ and the number of schools in the municipality where school i is located (n_{it}). This new term captures the different impact of Fundef across bigger and smaller municipalities. Hence, we rewrite model (1) in the following way[18]:

$$y_{it} = \alpha + \beta_1 d_t + \beta_2 d_{pub} + \delta_1 (d_t d_{pub}) + \delta_2 [(d_t d_{pub}) n_{it}] + \beta_3 X_{it} + \varepsilon_{it}\,. \qquad (1')$$

[18] The parameters of this model are different from the ones in model (1). For simplicity, we keep the same notation.

In this new model (1'), the parameter of interest is $(\delta_1 + \delta_2 n_{it})$ and it depends on the number of schools in the municipality where school "*i*" is located. As in model (1), the set of control variables X_{it} are used.

In order to carry out the estimation of equations (1) and (1'), we use data only for the years 1997 (before Fundef) and 1999 (after the introduction of the Fundef). Two data sets are used. From the Brazilian Educational Census database[19], we obtain the following variables for each Brazilian school (public and private): the dependent variable – the repetition rate; control variables related to the schools' infra-structure – number of teachers, number of televisions and VCR, dummy variables indicating if the school have computer labs, sciences labs, sports court, recreation area, library, access to electric energy and water services, and, if the school offers food for the enrolled students; control variables related to the municipality where the school is located – the number of public schools in each municipality, a dummy variable for the state where the school is located and a dummy variable indicating the school location (rural or urban); and a control variable related to the students – number of enrolled students who have previously failed the grade. This last variable is used to control for changes in the quality of the students across the years.

From the Brazilian Bureau of Statistics (IBGE), we obtain additional variables to control for characteristics of the municipalities: the Gini Coefficient, the Human Development Index (HDI), the GDP and the population[20, 21].

---

[19] This dataset is provided by the Brazilian Ministry of Education.
[20] Given that the Gini coefficient and the HDI are available only for 1991 and 2001, we use the values of these variables for 1991. Hence, these variables will be fixed in 1997 and 1999. We expect that these variables capture the "initial socioeconomic conditions" of each municipality.
[21] The GDP variable is available only for 1996 and 1999.

The data on repetition rate is available for each grade[22]. Hence, we are able to check if the effects of the Fundef differ from grade to grade, as predicted by the theoretical model. With this objective in mind, we estimate models (1) and (1') using OLS. First, we combine all grades in the same analysis and then separately for each one of the eight grades that form the fundamental education.

To perform the estimations, some adjustments in the database are necessary. First, some schools in the sample are classified as "Paralyzed" or "Extinguished", which, according to the Census Dictionary, means that the activities in these schools are temporarily or definitely suspended. Second, some schools are not in the database in both years (1997 and 1999) which can be due either to sampling error or the creation of new school units in 1999 (or both). Therefore, we exclude from the sample the "Paralyzed" and "Extinguished" schools and, to balance the panel, the schools which are not in the database in both years 1997 and 1999.[23] The total number of observations is equal to 318,206.

4.2) Empirical Results

Initially, we show in Table 1 (all empirical results are in appendix 2) the "raw" results (difference of means) and make simple comparisons between the average repetition rates in public and private schools, before (in the year 1997) and after (in 1999) the introduction of Fundef.

The numbers suggest the following. First, the repetition rates in public schools are on average substantially higher than in private schools in both years. For example, in 1999, the average repetition rate in public schools is 15%, well above the 3.6% in

[22] In the 90's, there was eight grades in the Brazilian fundamental education.
[23] The qualitative results are basically the same with the balanced and unbalanced panels. For simplicity, we only present the results obtained with the balanced data.

private schools. Second, after the introduction of Fundef, the difference in the repetition rate between public and private schools decreased from 12.8% in 1997 to 11.4% in 1999, as shown in the last column. The Difference-in-Difference parameter (in bold in the last column) is equal to -1.4%.

Table 2 presents the same "raw" results but now separated for each of the eight grades of the fundamental education. It is interesting to notice that the result varies across grades. After the introduction of Fundef, in comparison with private schools, there was a reduction in the repetition rate for the lower grades – more precisely, first, second and third grades – and an increase almost monotonically from the fourth grade up to the eighth grade. In other words, the Difference-in-Difference parameter (in bold in Table 2) is negative for grades 1 to 3 and positive for the others. This pattern is illustrated with the help of Figure 1. These numbers suggest that the effects of Fundef may be different across grades.

So far, our analysis did not consider any explanatory variable controlling for the potential (unobservable) differences in the trends of the treatment (public schools) and control (private schools) group[24]. In order to take this into consideration, we now estimate model (1) using OLS using data for the years 1999 (the year after the introduction of Fundef) and 1997 (the year before the introduction of Fundef).

The results are shown in Table 3. The explanatory variables are the same in all regressions. They are the variables explained in the previous subsection with controls related to the schools' infrastructure, the students' characteristic and the municipalities where the schools are located. In addition, there are three dummy variables: D1999 –

[24] The inclusion of control variables in the right-hand side of model (1) – and subsequently in model (1') – attempts to assure that the identification hypothesis of the model is indeed valid. Intuitively, we expect that when controlled for characteristics of the school, students and region – in which the school is located – the unobservable differences (across the time) between public and private schools vanish.

equal to one for the year 1999 and zero otherwise; Dpublic – equal to one for the public schools and zero otherwise; (D1999*Dpublic) – dummy variable that combines the year 1999 and the public schools. The parameter of interest is the one related to this last variable.

The first column presents the results when the dependent variable is the average repetition rate of each school (taking the average repetition of all grades in the school). The coefficient of the variable (D1999*Dpublic), reported in the third line in Table 3, is negative and significantly different from zero. It suggests that, due to the Fundef's effect, there was an overall decrease in the repetition rate in public schools by 0.87 percentage points in comparison with the private schools.

The other columns in Table 3 also present the results of the estimation of model (1) using OLS. The difference is that now the regressions are done separately for each of the eight grades of the fundamental education. Column 2 corresponds to the regression for the first grade up to column 9 and the regression for the eighth grade.

There are interesting results with respect to the coefficient of variable (D1999*Dpublic). With the exception of the regression for the second grade, the coefficients are statistically significantly different from zero. It is negative for the first three grades and positive to the others. It also increases almost monotonically from grade 4 to 8.

These results are in line with the theoretical results obtained previously in this paper. We saw that, due to the gaming effect, the theory predicted relatively lower and higher repetition rates, respectively, in the first and last grades. For example, in grade 1, the repetition rate in public schools, due to the Fundef effect, is 2.76 percentage points lower vis-à-vis the private schools. In grade 8, it is 1.67 percentage points higher. In other words, the empirical results seem to corroborate the predictions of the theoretical model with respect to the gaming effect.

Table 3A gives information about the order of magnitude of Fundef's effect. Column 1 replicates the coefficient of the variable (D1999*Dpublic) in the regressions in Table 3, on average and for each of the eight grades. Recall that this coefficient measures the impact of the Fundef. Column 2 presents the repetition rate in public schools in 1997, also on average and for each grade. The last column shows the ratio of columns 1 and 2.

On average, the magnitude of the Fundef's effect corresponds to only -4.94% of the repetition rate in 1997 (line 1 in Table 3A). However, its importance varies significantly from grade to grade. It increases monotonically from the first to last grade, ranging from -11.65% to 37.95%. Although its impact is negligible, for example, in the second grade (-1.09%), it reaches a considerable magnitude and its peak in the eighth grade. In this last grade, the interpretation is that, due to the Fundef, there was an increase in the repetition rate of almost 38%! In fact, the order of magnitude of the Fundef's impact is more important in the last four grades of the fundamental education.

We now estimate model (1') using OLS. The results are presented in Table 4. In comparison with Table 3, Table 4 adds one explanatory variable in the regressions, the interaction variable (D1999*Dpublic*n). It is the dummy variable (D1999*Dpublic) used in model (1) multiplied by "n", the number of public schools in the municipality where the public school is located. The idea is that the parameter related to this variable should capture the free-rider effect discussed in the theoretical section.

In Table 4, the first column presents the results when the dependent variable is the average repetition rate of each school. In the other columns (from 2 to 9), the regressions are separated for each of the eight grades of the fundamental education.

Initially, it is worth noting that the coefficients of both dummy variables of interest ("D1999*Dpublic*n" and "D1999*Dpublic") are significantly different from zero at 5% in all regressions, with the exception of the coefficients related to the variable "D1999*Dpublic" in the regressions for the second and third grades[25].

There is evidence of the gaming effect. The coefficient of the variable "D1999*Dpublic" is significant and negative for the first grade and positive from grade four to eight. In other words, there is relatively a lower repetition rate in the first grade and higher in the last ones. This result is similar to the ones reported in Table 3 when the variable "D1999*Dpublic*n" was not included in the regression.

The sign of the interaction "D1999*Dpublic*n" is always negative. It means that, for the average and for each grade, municipalities with more (less) schools had a lower (higher) repetition rate. This result is only in part in line with the theoretical results in this paper.

From the theoretical section, we obtained the prediction that the free-rider effect would mitigate the gaming effect. In other words, the free-rider effect should increase the repetition rate in the first grades and decrease it in the last ones. Although we do see empirical evidence corroborating the latter effect (with the negative sign of "D1999*Dpublic*n" in the last grades), there is no evidence of the former (the sign of the "D1999*Dpublic*n" is also negative in the first grades and the theory predicted it to be positive).

It is interesting to evaluate the free-rider and gaming effects combined. This is exactly what the parameter $(\delta_1 + \delta_2 n_{it})$ in model (1') captures. Recall that δ_1 and δ_2 are,

[25] They are significant at 1%.

respectively, the parameters associated with the variables "D1999*Dpublic" and "D1999*Dpublic*n".

In the first grade, the combined effect of Fundef led to a reduction in the repetition rate, as δ_1 and δ_2 are negative. Note that this effect is stronger in municipalities with a higher number of public schools. For the second and third grades, although the parameter δ_1 is not statistically differently from zero, the combined effect is qualitatively the same, as δ_2 is negative.

For the other grades, the combined effect of Fundef can be either positive or negative, depending on the number of schools in the municipality, as δ_1 is positive and δ_2 is negative. For municipalities with relatively few public schools, the combined Fundef effect is distributed across grades according to the following pattern: relatively lower and higher repetition rate, respectively, in the first and last grades.[26] Figure 2 illustrates this pattern for a municipality where ten public schools are located, that is, n=10.

The same pattern is not observed in municipalities with a large number of public schools. In those, the repetition rate is relatively lower for all grades[27]. Figure 3 illustrates this pattern for a municipality with a large number of public schools, n=1000.

In conclusion, the combined effect of Fundef (free-rider and gaming effects) seems to corroborate the view that teachers may have adjusted the repetition rate in order to maximize their income. For the municipalities with relatively low number of public schools, the pattern empirically identified is characterized by the reduction of this rate in the initial grades followed by an increase of the repetition rates in the last grades of the

[26] Formally, $(\delta_1 + \delta_2 n_{it})$ is negative and positive, respectively, for the initial and last grades.

[27] Formally, $(\delta_1 + \delta_2 n_{it})$ is negative for all grades.

fundamental education. This is exactly the optimum strategy suggested by the theoretical model. Moreover, the same pattern is not observed in municipalities with relatively high number of public schools.

4.3) Robustness

Approximately at the same time the Fundef was introduced, there was also an important change in the educational policy in some Brazilian States. They adopted the social promotion policy. For example, the state of São Paulo adopted a system of two cycles. The first and the second encompass, respectively, the four initial and the last four grades. Accordingly to this, students in the first, second, third, fifth, sixth and seventh grades pass automatically to the next grade. Students can only fail in the fourth and eighth grades. Other states adopt cycles with different number of grades and others more than two cycles. Obviously, these changes affect the actual repetition rates.

As a consequence, the results obtained in the previous subsection could in principle be related to the introduction of the social promotion policy and not due to the Fundef effect. Therefore, it is necessary to check the robustness of the results.

However, there is an important difficulty in separating the Fundef effect from the social promotion policy one on the repetition rate. The reason is that there is no information about which and when the schools actually started adopting the social promotion policy. For instance, there are States that adopt the social promotion policy to only part of the schools and maintained the old system in the rest.

In order to circumvent this lack of information problem and be able to check the robustness of our results, we adopt the following strategy. The Brazilian Ministry of Education collected information in the year of 1999 of the percentage of the schools in each Brazilian State that adopted the social promotion policy. Although it is not clear when this policy was implemented in the States, there is no evidence that the States that had adopted it had moved back to the previous policy (without social promotion) in the years before 1999. Based on this information, we re-estimate the same models (1) and (1') using OLS, but using a sub-sample that comprises only the States in which the social promotion policy was not introduced in at least 80% of its schools[28].

We expect that changes in the repetition rate in these States are then due to the introduction of the Fundef and can not possibly be due to the introduction of the social promotion policy.

The results are qualitatively and quantitatively basically the same as in the previous subsection, with the unrestricted sample[29]. It reinforces the conclusion obtained earlier that the Fundef seems to have given incentives to teachers to manage the repetition rate in order to maximize their income. In schools located in municipalities with relatively low number of public schools, the optimal pattern empirically identified is characterized by the reduction of this rate in the initial grades followed by an increase of the repetition rate in the last grades of the fundamental education.

5) Empirical Analysis: Proficiency Level

5.1) Empirical Strategy and Database

The results of the previous section that teachers seem to have received incentives to affect the repetition rate raise another important issue. That is, which are the mechanisms employed by teachers to affect this rate. There are some alternatives related with changes either in the educational standard or the quality of the education.

---

[28] Brazil is a federation with 26 States plus the Federal District. This sub-sample is formed by 18 States.
[29] The results are available under request from the authors. They are not included due to space limitations.

Given the educational standard, additional efforts to improve the quality of the education can reduce the repetition rate. Alternatively, given the quality of education, the reduction in educational standard can also reduce the repetition rate. In the former case, one should observe an increase in the students' proficiency level. In the latter case, a fall in the students' proficiency level is more likely to occur.

In order to evaluate these issues, we investigate the effects of the Fundef on the students' proficiency level. As in the analysis of the previous section, we employ the Difference-in-Difference (DID) method and estimate model (1) using OLS, but we use a different database.

We use the data of the National Basic Education Evaluation System (SAEB), provided by the Brazilian Ministry of Education. The SAEB is a governmental program aimed at evaluating the quality of the Brazilian basic education. This program consists of biennial proficiency tests of Mathematics and language (Portuguese) applied to a sample of students enrolled in 4th and 8th grades of fundamental education and 3rd grade of secondary education in public and private schools.

From the SAEB, we obtain the following variables for a sample of Brazilian schools for the years of interest, necessary to estimate model (1): the dependent variable – the math test school average; explanatory control variables related to the quality of schools' infra-structure – dummy variables indicating if the school has television *(TV)*, VCR/DVD *(VCR/DVD)*, library *(Library)*, access to electric energy *(Electricity)* and water services *(Water)*; explanatory control variables related to the school location – a dummy variable indicating the school location (rural or urban) *(Location)* and the State in the Brazilian federation where it is located; explanatory control variables related to teachers' characteristics – dummy variable indicating teacher's gender *(Man)* and teacher's education level *(Education level)*, years of experience *(Experience)* and wage *(Wage)*; and explanatory control variables related to students' characteristics – dummy variables indicating student's race *(White)* and gender *(Man)*, if he has not failed any grade before *(Failed)*, if he had already studied in private school *(Previously Private School)*, and father's *(Father's education)* and mother's *(Mother's education)* education. These control variables are the explanatory variables (X_{it}) in model (1).

We evaluate the effects of the Fundef on the proficiency level in the fourth and eighth grades. Menezes-Filho and Pazello (2004) also make this evaluation using the same methodology[30]. However, in contrast with their paper, we introduce a modification in the empirical strategy. They used data for the years 1997 (before Fundef) and 1999 and 2001 (after the introduction of the Fundef) to evaluate potential differences in performance of the public schools' students vis-à-vis the private school ones. The problem with this time frame is that, for example, a public school student registered in the 8th grade in 2001 started the fundamental education cycle no later than 1994. Hence, he studied at least during four years in the public school (from 1994 up to 1998 when the Fundef was introduced) that had not yet received the Fundef's influence. The same type of problem occurs in the evaluation of the effects of the Fundef on the performance of the fourth grade students.

To deal with this problem, we use the following strategy. In order to evaluate the performance of the eighth grade students, we use data for the years 1997 (the year before the introduction of the Fundef) and 2005 (eight years after). In a similar way, in order to evaluate the performance of the fourth grade students, we use data for the years 1997 (the year before the introduction of the Fundef) and 2001 (four years after).

[30] They found no evidence that the Fundef improved the proficiency level of students in the eighth grade.

5.2) Empirical Results

We now turn to the empirical results and analyze the effects of Fundef on the proficiency level. Table 7 presents the results.

In the first column, we present the results of the estimation of model (1) using OLS and data for the years 2001 and 1997 to check the Fundef's effects on the proficiency level on the fourth grade. We use an unbalanced panel of schools. In other words, all schools that are in the SAEB samples in 1997 and 2001 are included, independently if the school is included in both samples or not.

The dependent variable is the math test school average on the fourth grade. Besides the control variables discussed in the previous subsection, there are three other explanatory variables: the dummy variable (D2001) – equal to one for the year 2001 and zero otherwise; the dummy variable (Dpublic) – equal to one if the school is public and zero otherwise; and the dummy variable (D2001*Dpublic) – the interaction of the two previous dummy variables.

The parameter of interest is the one related to the dummy variable (D2001*Dpublic). It is negative and significantly different from zero. It suggests that the Fundef led to a reduction in the level of Mathematics proficiency by the fourth grade students in public schools. This result could be interpreted as an indication that there was a fall in the quality of education in the first four years of the fundamental education after the introduction of Fundef.

From the previous section, we obtained the result that there was a reduction in the repetition rate in the initial grades. It occurred simultaneously with a fall in the quality of education. One could interpret that these two results combined could be an indication that there was a reduction in the educational standard. That is, the requirements necessary for a student to pass to an advanced grade in the initial grades were reduced.

We re-do the estimation of model (1) using OLS and data for the years 2001 and 1997, but using the balanced panel[31]. That is, we incorporate in the regression only schools that are included in both SAEB samples, in 1997 and 2001. The results are reported in the second column in Table 7. They are qualitatively equal to the results for the unbalanced panel.

We now check the effects of Fundef on the proficiency level on the eighth grade. In Table 7, columns 3 and 4 report the results of the estimation of model (1) using OLS and data for the years 1997 and 2005 using, respectively, the unbalanced and balanced panel.

The dependent variable is the math test school average on the eighth grade. Besides the control variables discussed in the previous subsection[32], there are three other explanatory variables: the dummy variable (D2005) – equal to one for the year 2005 and zero otherwise; the dummy variable (Dpublic) – equal to one if the school is public and zero otherwise; and the dummy variable (D2005*Dpublic) – the interaction of the two previous dummy variables.

The parameter of interest is the one related to the dummy variable (D2005*Dpublic). The results are qualitatively the same in both panels (unbalanced and balanced). It is negative and significantly different from zero. As in the analysis of the

[31] Note that the number of observations in the balanced panel is significantly smaller. It is equal to 270 in contrast with the unbalanced panel with 4994 observations.
[32] In the regressions related to the eighth grade, the control variable *(Localization)* is not included because it is not part of 2005 Saeb database. Moreover, in the balanced panel, the control variable *(Electricity)* is also not included because all schools in this panel have electricity.

fourth grade, it also suggests that the Fundef led to a reduction in the level of Mathematics proficiency by the eighth grade students in public schools. Again, it signals a fall in the quality of education.

From the previous section, we obtained the evidence that there was an increase in the repetition rate in the final grades. This result could be explained either by a rise in the educational standard or a fall in the quality of education. However, the first explanation should have led to a rise in the students' proficiency, which does not seem to have occurred based on the evidence presented above. Hence, the second explanation is more likely to have occurred.

In conclusion, there is no empirical evidence corroborating the view that the Fundef led to an increase in the quality of the education in Brazilian public schools. The results obtained above suggest that it is more likely to have occurred a fall in the standard level in the initial grades, and a fall in the quality of education in the last grades.

6) Conclusion

We investigated possible effects associated with the introduction of the Brazilian program called Fundef. In particular, we checked if one important feature of this program – the fact that teachers' wage become indirectly linked to the number of students enrolled in their schools – affected their behavior. We tested the predictions of the model developed in this paper and found empirical evidence supporting them.

On the one hand, teachers seem to have engaged in opportunistic behavior or gaming. They adjusted the fail and repetition rates in order to maximize the number of students and, consequently, receive higher wages. The optimal strategy was to reduce the fail and repetition rates in the first grades – to avoid students to drop out from school

– and to increase them in the final grades – to maintain students longer periods in school. The magnitude of this effect was significant. For example, there was a fall of almost 12% in the repetition rate in the first grade vis-à-vis the pre-Fundef period. In the last grade of the fundamental education, there was a rise of almost 38%.

On the other hand, the optimal strategy indicated above occurs in the schools located in the municipalities with a small number of public schools. However, the same pattern is not observed in municipalities with relatively high number of schools. These are the results when we take into consideration the combined effects of the Fundef, that is, the gaming and free-rider effects.

The results obtained in this paper are additional evidence that teachers seem to respond to incentives. When designing educational programs, policymakers should take into consideration these possible side-effects and, if possible, adopt measures to mitigate them.

References

- Angrist, J.; Krueger, A. B. (1999). "Empirical Strategies in Labor Economics" in Handbook of Labor Economics (Vol. 3), Ashenfelter, O. and Card, D. (editors), North-Holland.
- Card, D. (1990). "The Impact of the Mariel Boatlift on the Miami Labor Market"; Industrial and Labor Relations Review, Vol. 43, No.2, pp.245-257.
- Carnoy, M.; Loeb, S.; Smith, T. (2001). "Do higher state test scores in Texas make for better high school outcomes?" Research Report Series, RR-047, Consortium for Policy Research in Education.

- Carnoy, M. e Loeb, S. (2002). “Does external accountability affect student outcomes? A cross-state analysis”. Educational Evaluation and Policy Analysis, vol. 24 no. 4.
- Figlio, D. e Getzler, L. (2002). “Accountability, ability and disability: Gaming the system?” National Bureau of Economic Research, WP 9307.
- Haney, W. (2000). “The myth of the Texas miracle in education.” Education Policy Analysis Archives 8, no. 41.
- Hanushek, E. and Raymond, M. (2204). “The Effect of School Accountability Systems on the Level and Distribution of Student Achievment”; Journal of the European Economic Association; April-May.
- Jacob, B. (2004). “Accountability, incentives and behavior: Evidence from school reform in Chicago.” National Bureau of Economic Research, WP 8968.
- Koretz, D. e Barron, S. (1998). “The validity of gains in scores on the Kentucky Instructional Results Information System (KIRIS)”. Santa Monica, CA: RAND Corporation.
- Menezes-Filho, N. e Pazello, E. (2004). “Evaluating the effects of FUNDEF on wages and test scores in Brazil”. Mimeo.
- Olson, M. (1965). “The logic of collective action”. Cambridge, MA: Harvard University Press.
- Ribeiro, S. (1992). “A pedagogia da repetência”. Estudos Avançados, IEA/USP.
- Semeghini, U. (2001). “FUNDEF: Uma revolução silenciosa”. Mimeo.
- Souza, P. (2005). “A revolução gerenciada: educação no Brasil, 1995-2002”. Prentice Hall.
- Toenjes, L. e Dworkin, G. (2002). “Are increasing test scores in Texas really a myth, or is Haney´s myth a myth?” Education Policy Analysis Archives 10, no. 17.

Appendix 1

Proof of Proposition 1.

Recall the teacher's maximization problem.

$$\max_{\{E_1^j;E_2^j\}} = U\left[R\left(\bar{N}\left\{2-(1-\lambda)\left(\bar{f}-k\left(E_1^j\right)^{\alpha}\right)+\lambda\left(\bar{f}+k\left(E_2^j\right)^{\alpha}\right)\right\}+\sum_{i\neq j}S^i\right);1-E_1^j-E_2^j\right]$$

Since the objective function is concave, the first order conditions yield the optimal choice for the teacher.

Denote by *A* the expression $\bar{N}\left\{2-(1-\lambda)\left(\bar{f}-k\left(E_1^j\right)^{\alpha}\right)+\lambda\left(\bar{f}+k\left(E_2^j\right)^{\alpha}\right)\right\}$ and by Σ the sum $\sum_{i\neq j}S^i$. Then, the first order conditions can be written as below.

$$U_1\left[R(A+\Sigma);1-E_1^j-E_2^j\right]\cdot R'(A+\Sigma)\cdot\bar{N}\cdot\alpha(1-\lambda)k(E_1)^{\alpha-1}=U_2\left[R(A+\Sigma);1-E_1^j-E_2^j\right].$$

$$U_1\left[R(A+\Sigma);1-E_1^j-E_2^j\right]\cdot R'(A+\Sigma)\cdot\bar{N}\cdot\alpha\lambda k(E_2)^{\alpha-1}=U_2\left[R(A+\Sigma);1-E_1^j-E_2^j\right].$$

Combining the two first order conditions yield $(1-\lambda)(E_1)^{\alpha-1}=\lambda(E_2)^{\alpha-1}$, hence $E_2=\left(\frac{\lambda}{1-\lambda}\right)^{\frac{1}{1-\alpha}}E_1$. Note that, since *U* is strictly increasing, it must be the case that $U_2\left(R(A+\Sigma);1-E_1^j-E_2^j\right)>0$. But then, the first condition above requires that $E_1^*>0$ and the second condition requires that $E_2^*>0$.

Proof of Proposition 4.

Recall the notation $\Sigma=\sum_{i\neq j}S^i$. We will determine the effect of a change in Σ on the solution E_1^*,E_2^*. From the previous proof of Proposition 1, we can write

$E_1^* = E(\Sigma)$, $E_2^* = lE(\Sigma)$ where $l = \left(\frac{\lambda}{1-\lambda}\right)^{\frac{1}{1-\alpha}}$. Let $A(\Sigma)$ represent the expression $\overline{N}\{2-(1-\lambda)(\bar{f} - k(E(\Sigma))^{\alpha}) + \lambda(\bar{f} + kl^{\alpha}(E(\Sigma))^{\alpha})\}$; then, the first order conditions in equilibrium reduce to:

$$U_1[R(A(\Sigma)+\Sigma);1-(1+l)E(\Sigma)] \cdot R'(A(\Sigma)+\Sigma) \cdot \overline{N} \cdot \alpha(1-\lambda)k(E(\Sigma))^{\alpha-1} = U_2[R(A(\Sigma)+\Sigma);1-(1+l)E(\Sigma)]$$

Suppose Σ increases. We will consider three possible effects on $E(\Sigma)$.

Suppose, first, that $E(\Sigma)$ does not change. In that case, the right hand side of the first order condition does not change. On the other hand, since *R* is strictly increasing and strictly concave, and since *U* is strictly concave, the left hand side decreases. But this yields a contradiction. Therefore, $E(\Sigma)$ has to change.

Suppose, second, that $E(\Sigma)$ increases. In that case $R(A(\Sigma)+\Sigma)$ increases. Since $U_{21}>0$ and $U_{22}<0$ (*U* is strictly concave), it must be the case that the right hand side of the first order condition increases. On the other hand, $R'(A(\Sigma)+\Sigma)$ and $E(\Sigma)^{\alpha-1}$ decrease. Moreover, since $U_{12}>0$ and $U_{11}<0$ (*U* is strictly concave), then $U_1[R(A(\Sigma)+\Sigma);1-(1+l)E(\Sigma)]$ also decreases. Hence, the left hand side of the first order condition decreases. But this yields again a contradiction.

Therefore, it must be the case that $E(\Sigma)$ decreases. Hence, as the number of schools in a municipality or state increase, the incentives for teacher to make an effort in order to affect the fail rate decrease.

Appendix 2

Table 1: DID – Average Repetition Rate

	Public	Private	Difference
1999	0.150	0.036	0.114

1997	0.176	0.048	0.128
Difference	-0.027	-0.013	**-0.014**

Table 2: DID – Repetition Rate for Grades 1-8

	Public	Private	Difference	Grade
1999	0.203	0.029	0.174	
1997	0.237	0.041	0.196	**S=1**
Difference	-0.034	-0.012	**-0.022**	
1999	0.150	0.024	0.126	
1997	0.165	0.035	0.130	**S=2**
Difference	-0.016	-0.011	**-0.005**	
1999	0.101	0.026	0.075	
1997	0.117	0.037	0.080	**S=3**
Difference	-0.016	-0.012	**-0.005**	
1999	0.079	0.024	0.055	
1997	0.084	0.032	0.051	**S=4**
Difference	-0.004	-0.008	**0.004**	
1999	0.105	0.049	0.057	
1997	0.118	0.071	0.047	**S=5**
Difference	-0.013	-0.022	**0.009**	
1999	0.091	0.049	0.042	
1997	0.094	0.064	0.030	**S=6**
Difference	-0.003	-0.016	**0.013**	
1999	0.070	0.045	0.025	
1997	0.070	0.056	0.013	**S=7**
Difference	0.000	-0.011	**0.011**	
1999	0.054	0.037	0.017	
1997	0.044	0.043	0.001	**S=8**
Difference	0.009	-0.006	**0.016**	

Figure 1: DID – Average and Grades 1-8

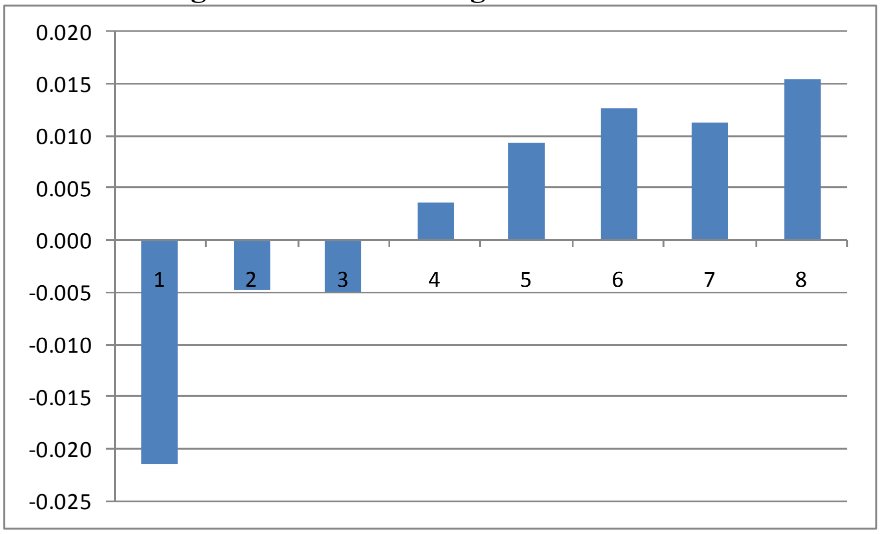

Table 3: OLS Results Equation (1) – Average and Grades 1-8 (*) (**)

	Average	G=1	G=2	G=3	G=4	G=5	G=6	G=7	G=8
D1999	-0.0141	-0.0089	-0.0123	-0.0113	-0.0136	-0.0201	-0.0129	-0.0100	-0.0070
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Dpublic	0.0609	0.0840	0.0840	0.0526	0.0409	0.0286	0.0231	0.0126	0.0074
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
D1999*Dpublic	-0.0078	-0.0266	-0.0012	-0.0044	0.0082	0.0156	0.0158	0.0132	0.0167
	0.000	0.000	0.384	0.002	0.000	0.000	0.000	0.000	0.000
School/Students Controls									
Fail Rate Stock	1.9E-01	2.1E-01	1.1E-01	7.2E-02	7.4E-02	1.9E-01	1.5E-01	1.3E-01	1.3E-01
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Television (Number)	-0.0011	-0.0004	-0.0008	-0.0006	0.0000	-0.0013	-0.0010	-0.0008	-0.0009
	0.000	0.012	0.000	0.000	0.915	0.000	0.000	0.000	0.000
VCR/DVD (Number)	-2.4E-06	-6.1E-06	4.5E-06	-1.7E-06	-3.1E-06	4.9E-05	-2.5E-04	-2.0E-04	4.0E-06
	0.331	0.098	0.309	0.566	0.213	0.775	0.134	0.046	0.980
Computer Lab	0.0107	0.0115	0.0002	0.0032	-0.0003	-0.0066	-0.0025	-0.0006	0.0032
	0.000	0.000	0.878	0.001	0.765	0.000	0.016	0.519	0.002
Sciences Lab	0.0081	0.0056	0.0021	-0.0010	-0.0021	-0.0017	-0.0028	-0.0006	-0.0008
	0.000	0.000	0.020	0.226	0.018	0.070	0.002	0.509	0.317
Sports Court	-0.0018	-0.0058	-0.0059	-0.0032	0.0018	0.0041	0.0037	0.0035	0.0042
	0.000	0.000	0.000	0.000	0.009	0.000	0.000	0.000	0.000
Recreation Area	-0.0020	0.0000	0.0012	0.0002	0.0003	-0.0012	-0.0016	-0.0011	-0.0021
	0.000	0.995	0.167	0.784	0.699	0.146	0.048	0.136	0.008
Library	-0.0082	-0.0140	-0.0129	-0.0057	0.0029	0.0090	0.0047	0.0037	0.0030
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003
Electricity	-0.0180	-0.0216	0.0022	0.0029	-0.0011	0.0086	0.0047	-0.0010	-0.0043
	0.000	0.000	0.069	0.011	0.313	0.021	0.279	0.807	0.330
Water	-0.0086	-0.0067	-0.0025	-0.0023	-0.0012	0.0090	0.0071	0.0001	-0.0040
	0.000	0.006	0.176	0.203	0.480	0.031	0.055	0.969	0.356
Food	0.0023	0.0042	0.0055	0.0072	0.0067	0.0085	0.0020	0.0011	-0.0031
	0.044	0.049	0.002	0.000	0.000	0.000	0.096	0.325	0.059
Municipality Controls									
Number of Teachers	-0.0001	-0.0003	-0.0001	0.0000	0.0002	0.0002	0.0003	0.0003	0.0004
	0.000	0.000	0.000	0.396	0.000	0.000	0.000	0.000	0.000
Number of Public Schools	1.4E-04	1.8E-04	9.0E-05	3.4E-05	-1.8E-05	-5.0E-05	-3.2E-05	-1.1E-06	-3.7E-05
	0.000	0.000	0.000	0.001	0.077	0.000	0.001	0.907	0.000
Number of Private Schools	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	0.000	0.000	0.024	0.050	0.000	0.000	0.000	0.000	0.000
HDI	-0.2172	-0.2709	-0.1429	-0.1218	-0.1074	0.1339	0.1562	0.1399	0.0753
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Gini	0.0221	0.0444	0.0345	0.0091	0.0265	0.0119	-0.0030	-0.0062	0.0081
	0.001	0.000	0.000	0.274	0.001	0.151	0.714	0.390	0.285
GDP	-1.7E-09	-7.1E-10	-1.7E-09	-7.1E-10	-2.3E-09	-1.3E-09	-5.5E-10	-6.0E-10	-6.2E-10
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Population	1.0E-08	-9.6E-09	1.1E-08	8.1E-09	3.7E-08	3.0E-08	1.8E-08	1.2E-08	1.7E-08
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Area (A)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.002	0.001	0.072	0.919	0.142	0.029	0.159	0.019	0.607
Dlocalization	-0.0155	-0.0331	-0.0089	-0.0038	0.0063	0.0169	0.0168	0.0154	0.0171
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Observations	318206	294622	286160	271839	245512	81040	74222	70219	66396

(*) In all regressions, the equations include dummies for each state and a constant. For simplicity, we excluded these variables from the table.
(**) In all regressions, P-values based on White standard-errors are in italic.

TABLE 3A

	FUNDEF'S EFFECT (%) (1)	REPETITION RATE IN 1997 (%) (2)	(1)/(2)
AVERAGE	-0.78	17.6	-4.43
G1	-2.66	23.7	-11.22
G2	-0.12	16.5	-0.73
G3	-0.44	11.7	-3.76
G4	0.82	8.4	9.76
G5	1.56	11.8	13.22
G6	1.58	9.4	16.81
G7	1.32	7	18.86
G8	1.67	4.4	37.95

Table 4: OLS Results Equation (1') – Average and Grades 1-8

	Average	G=1	G=2	G=3	G=4	G=5	G=6	G=7	G=8
D1999	**-0.0175**	**-0.0118**	**-0.0154**	**-0.0143**	**-0.0155**	**-0.0241**	**-0.0174**	**-0.0130**	**-0.0096**
	0.000	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***
Dpublic	**0.0639**	**0.0868**	**0.0870**	**0.0555**	**0.0428**	**0.0316**	**0.0264**	**0.0147**	**0.0092**
	0.000	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***
D1999*Dpublic	-0.0004	**-0.0198**	**0.0059**	0.0026	**0.0125**	**0.0243**	**0.0259**	**0.0197**	**0.0224**
	0.733	***0.000***	***0.000***	*0.070*	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***
D1999*Dpublic*(N)	**-5.1E-05**	**-5.0E-05**	**-5.3E-05**	**-5.1E-05**	**-3.1E-05**	**-4.9E-05**	**-5.5E-05**	**-3.5E-05**	**-3.0E-05**
	0.000	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***
School/Students Controls									
Fail Rate Stock	**0.1849**	**0.2052**	**0.1107**	**0.0725**	**0.0739**	**0.1862**	**0.1524**	**0.1336**	**0.1242**
	0.000	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***
Television (Number)	**-0.0012**	**-0.0005**	**-0.0009**	**-0.0007**	0.0000	**-0.0014**	**-0.0011**	**-0.0009**	**-0.0009**
	0.000	***0.002***	***0.000***	***0.000***	*0.644*	***0.000***	***0.000***	***0.000***	***0.000***
VCR/DVD (Number)	-2.3E-06	-6.0E-06	4.6E-06	-1.6E-06	-3.1E-06	2.4E-05	-2.8E-04	**-2.2E-04**	-1.0E-05
	0.352	*0.111*	*0.290*	*0.590*	*0.218*	*0.889*	*0.109*	***0.037***	*0.949*
Computer Lab	**0.0108**	**0.0114**	0.0000	**0.0031**	-0.0004	**-0.0061**	-0.0019	-0.0002	**0.0035**
	0.000	***0.000***	*0.987*	***0.002***	*0.716*	***0.000***	*0.065*	*0.805*	***0.001***
Sciences Lab	**0.0079**	**0.0052**	0.0017	-0.0014	**-0.0023**	-0.0017	**-0.0028**	-0.0005	-0.0008
	0.000	***0.000***	*0.061*	*0.094*	***0.009***	*0.075*	***0.002***	*0.542*	*0.343*
Sports Court	**-0.0017**	**-0.0057**	**-0.0057**	**-0.0030**	**0.0020**	**0.0042**	**0.0038**	**0.0036**	**0.0043**
	0.001	***0.000***	***0.000***	***0.000***	***0.006***	***0.000***	***0.000***	***0.000***	***0.000***
Recreation Area	**-2.0E-03**	2.7E-05	1.2E-03	2.4E-04	3.0E-04	-1.1E-03	-1.5E-03	-1.0E-03	**-2.0E-03**
	0.000	*0.979*	*0.159*	*0.761*	*0.685*	*0.175*	*0.063*	*0.167*	***0.011***
Library	**-0.0078**	**-0.0135**	**-0.0125**	**-0.0053**	**0.0032**	**0.0091**	**0.0048**	**0.0038**	**0.0031**
	0.000	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.002***
Electricity	**-1.8E-02**	**-2.2E-02**	1.9E-03	**2.6E-03**	-1.2E-03	**8.3E-03**	4.4E-03	-1.3E-03	-4.6E-03
	0.000	***0.000***	*0.109*	***0.020***	*0.246*	***0.026***	*0.313*	*0.768*	*0.304*
Water	**-0.0086**	**-0.0067**	-0.0025	-0.0023	-0.0013	**0.0087**	0.0067	-0.0002	-0.0042
	0.000	***0.005***	*0.169*	*0.196*	*0.467*	***0.037***	*0.074*	*0.963*	*0.328*
Food	0.0016	0.0034	**0.0046**	**0.0064**	**0.0061**	**0.0070**	0.0002	0.0000	**-0.0041**
	0.155	*0.107*	***0.008***	***0.000***	***0.000***	***0.000***	*0.841*	*0.986*	***0.017***
Municipality Controls									
Number of Teachers	**-6.6E-05**	**-2.7E-04**	**-1.0E-04**	1.9E-05	**1.7E-04**	**1.9E-04**	**2.6E-04**	**3.3E-04**	**4.4E-04**
	0.000	***0.000***	***0.000***	*0.274*	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***
Number of Public Schools (N)	**1.6E-04**	**2.0E-04**	**1.1E-04**	**5.4E-05**	-5.4E-06	**-3.2E-05**	-1.2E-05	1.1E-05	**-2.7E-05**
	0.000	***0.000***	***0.000***	***0.000***	*0.599*	***0.001***	*0.203*	*0.233*	***0.003***
Number of Private Schools	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***
	0.000	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***
HDI	**-0.2239**	**-0.2759**	**-0.1483**	**-0.1273**	**-0.1112**	**0.1235**	**0.1440**	**0.1319**	**0.0681**
	0.000	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***
Gini	**0.0212**	**0.0436**	**0.0336**	0.0081	**0.0258**	0.0113	-0.0035	-0.0065	0.0078
	0.001	***0.000***	***0.000***	*0.327*	***0.001***	*0.174*	*0.673*	*0.363*	*0.304*
GDP	**-2.8E-09**	**-1.6E-09**	**-2.7E-09**	**-1.7E-09**	**-2.9E-09**	**-2.4E-09**	**-1.8E-09**	**-1.4E-09**	**-1.3E-09**
	0.000	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***
Population	**2.6E-08**	3.8E-09	**2.5E-08**	**2.2E-08**	**4.6E-08**	**4.7E-08**	**3.7E-08**	**2.4E-08**	**2.7E-08**
	0.000	*0.105*	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***
Area (A)	**0.0000**	**0.0000**	0.0000	0.0000	0.0000	**0.0000**	0.0000	**0.0000**	0.0000
	0.003	***0.001***	*0.093*	*0.826*	*0.170*	***0.011***	*0.074*	***0.009***	*0.462*
Dlocalization	**-0.0154**	**-0.0329**	**-0.0087**	**-0.0036**	**0.0064**	**0.0171**	**0.0170**	**0.0156**	**0.0173**

	0.000	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***
Observations	318206	294622	286160	271839	245512	81040	74222	70219	66396

Figure 3: DID – Average and Grades 1-8, N=10 Schools

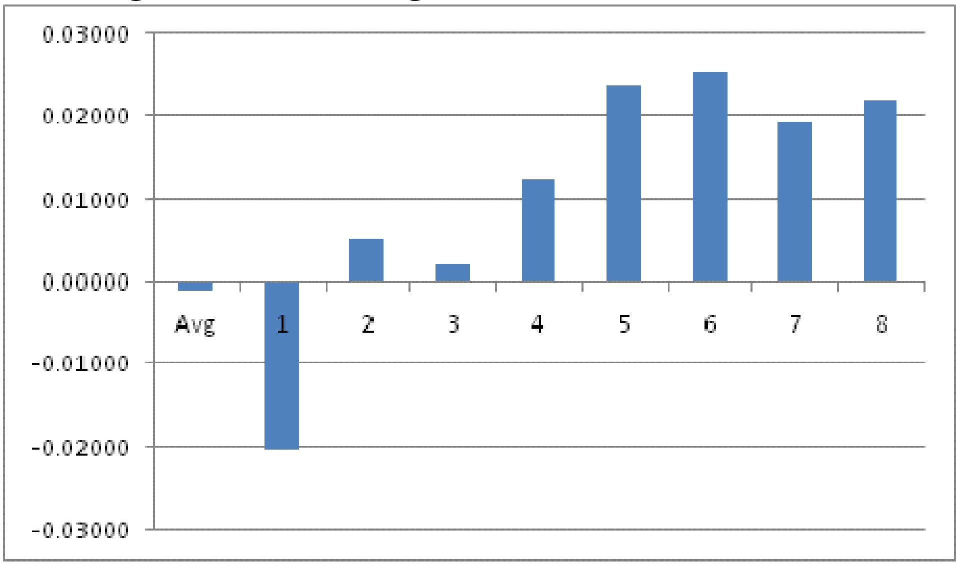

Figure 4: DID – Average and Grades 1-8, N=1000 Schools

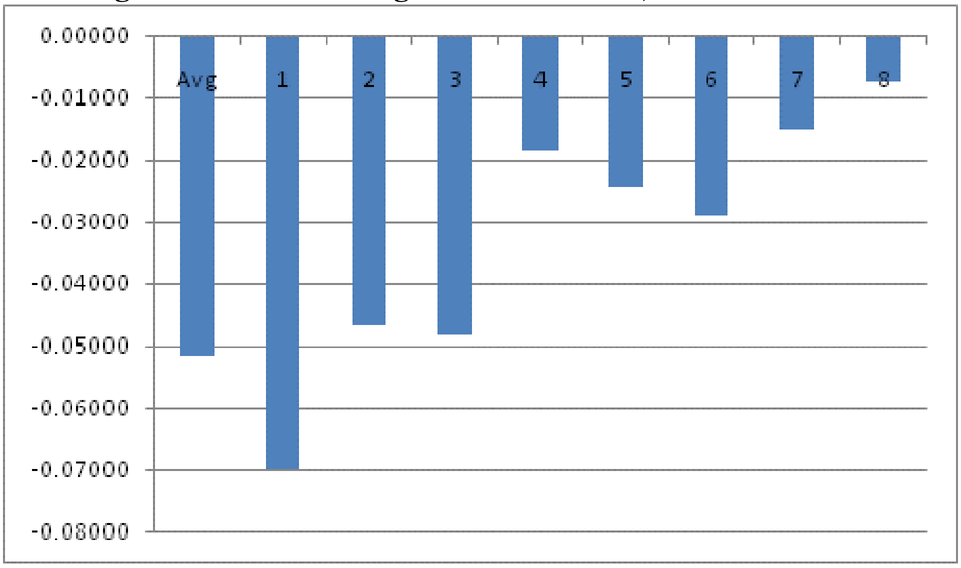

Table 7: OLS Results Equation (1') – Average and Grades 4th and 8th

4th year			8th year		
	Unbalanced	Balanced		Unbalanced	Balanced
D2001	4.7630	8.5407	D2005	-6.1869	-1.5700
	0.000	0.015		0.000	0.763
Dpublic	-16.7466	-17.6638	Dpublic	1.8038	13.9907
	0.000	0.041		0.508	0.317
D2001*Dpublic	-8.4527	-10.6191	D2005*Dpublic	-5.2951	-17.0311
	0.000	0.013		0.002	0.0030
Students Controls			Students Controls		
Men	-4.6161	-11.6825	Men	7.2576	21.0084
	0.024	0.259		0.017	0.070
White	3.8258	11.2983	White	16.4065	2.4353
	0.034	0.244		0.000	0.838
Father's education	7.5262	1.7053	Father's education	8.0401	7.1758
	0.000	0.674		0.000	0.284
Mother's education	8.8561	14.5321	Mother's education	13.7558	6.4454
	0.000	0.000		0.000	0.248
Previous private school	-0.9862	-0.7976	Previous private school	13.6065	18.8509
	0.448	0.895		0.000	0.043
Failed	16.9084	18.7950	Failed	25.7560	29.2164
	0.000	0.011		0.000	0.000
Teachers Controls			Teachers Controls		
Men	-0.6732	2.3641	Men	-1.0045	-1.8681
	0.530	0.509		0.409	0.635
Education level	-0.1443	1.4093	Education level	2.7839	7.9472
	0.633	0.355		0.079	0.058
Experience	0.8693	-1.2438	Experience	1.1033	-1.4907
	0.311	0.748		0.354	0.774
Wage	3.4028	0.2597	Wage	1.9467	3.4371
	0.000	0.859		0.004	0.208
Schools Controls			Schools Controls		
Localization	6.2796	7.5569	Localization	-	-
	0.000	0.1640		-	-
Water	-4.3761	-8.0962	Water	10.2786	11.5209
	0.028	0.3150		0.111	0.2020
Electricity	-0.2834	-5.3236	Electricity	-23.6435	-
	0.924	0.3340		0.008	-
Library	0.7014	2.3812	Library	1.2888	3.0801
	0.283	0.3100		0.229	0.3730
Television	0.7417	1.0270	Television	2.2636	-2.5040
	0.676	0.8680		0.412	0.6190
VCR/DVD	0.5820	2.1806	VCR/DVD	2.1315	1.4751
	0.712	0.6490		0.214	0.8010
Observations	4994	270	Observations	2595	165