UNIVERSITÀ DELLA CALABRIA

Dipartimento di Economia e Statistica
Ponte Pietro Bucci, Cubo 0/C
87036 Arcavacata di Rende (Cosenza)
Italy
http://www.ecostat.unical.it/

Working Paper n. 06 - 2010

MONETARY INCENTIVES AND STUDENT ACHIEVEMENT IN A DEPRESSED LABOUR MARKET: RESULTS FROM A RANDOMIZED EXPERIMENT

Maria De Paola
Dipartimento di Economia e Statistica
Università della Calabria
Ponte Pietro Bucci, Cubo 1/C
Tel.: +39 0984 492459
Fax: +39 0984 492421
e-mail: m.depaola@unical.i

Vincenzo Scoppa
Dipartimento di Economia e Statistica
Università della Calabria
Ponte Pietro Bucci, Cubo 1/C
Tel.: +39 0984 492464
Fax: +39 0984 492421
e-mail: v.scoppa@unical.it

Rosanna Nisticò
Dipartimento di Economia e Statistica
Università della Calabria
Ponte Pietro Bucci, Cubo 1/C
Tel.: +39 0984 492449
Fax: +39 0984 492421
e-mail: r.nistico@unical.it

Aprile 2010

Monetary Incentives and Student Achievement in a Depressed Labour Market: Results from a Randomized Experiment

Maria De Paola, Vincenzo Scoppa, Rosanna Nisticò[*]

Department of Economics and Statistics, University of Calabria

In this paper we carry out an evaluation of the effectiveness of monetary incentives in enhancing student performance using a randomized experiment involving undergraduate students enrolled at a Southern Italian University. Students participating at the experiment were assigned to three different groups: a high reward group, a low reward group and a control group. Rewards were assigned following a tournament rule to the 30 best performing students in each treated group. Findings suggest that financial rewards increase student performance both in terms of number of credits earned and grades obtained at exams. High ability students react strongly while the effect is null for low ability students. When the "intention-to-treat" effects are adjusted for non-participation using IV, it emerges a stronger impact of the treatment on student performance.

Keywords: Education Production Function; Student Effort; Incentives; Merit Scholarship; Higher Education; Randomized Evaluation.

JEL classification: I21; J31; D82

1. Introduction

The use of monetary incentives to improve student performance is a common practice in the educational systems of several countries. For example, in the United States, following the "Helping Outstanding Pupils Educationally" (*HOPE*) program implemented in Georgia, many others States have introduced analogous scholarship programs (Heller, 2006); in the United Kingdom the program Education Maintenance Allowance confers to students from low income families financial rewards based on their school performance (Middleton *et al.*, 2003); in Colombia the program *PACES* (Programa de Ampliaciòn de Cobertura de Educaciòn Secundaria) assigned vouchers to secondary school students in relation to their performance (Angrist *et al.*, 2002).

[*] E-mail: m.depaola@unical.it**;** v.scoppa@unical.it**;** r.nistico@unical.it**.** The project was funded by the Regional Government of Calabria (Assessorato Istruzione, Alta Formazione e Ricerca, Piano Operativo Regionale Calabria, European Social Fund 2007-2013). We would like to thank Alfio Cariola, Domenico Cersosimo, Alfredo Fortunato, Sara Laurita who assisted with the project and the University of Calabria for providing access to the administrative data. We have benefited from helpful comments from Mariarosaria Agostino, Giovanni Anania, Federica Demaria and Michela Ponzo.

Contextually to the diffusion of monetary incentives to highly performing students, a number of works has tried to investigate their effectiveness using experimental methods. Results are not univocal: some studies provide evidence of a positive impact, while others find rather weak effects.

Angrist and Lavy (2009) evaluate the effectiveness of financial rewards on the achievement of Israeli students using a school based randomized experiment offering monetary prizes to students who obtain the matriculation certificate. The authors show that the program has led to relevant effects for girls but not for boys.

Differences among genders emerge also from Angrist, Lang and Oreopoulos (2009), who conducted an experimental evaluation of different strategies aimed at improving student academic performance. A sample of students enrolled at a large Canadian university were randomly assigned to one of three treatment groups: the first group was offered a number of support services; the second was offered financial rewards for good performance; the third was offered a combination of support services and monetary incentives. The results of the experiment show that while males did not react to any of the proposed strategies, females improved their academic performance.

Angrist *et al.* (2002) and Angrist, Bettinger and Kremer (2006) analyse the impact of a large school voucher program in Colombia covering part of the cost of private secondary school. Vouchers were renewable conditional on student performance. From these studies, it emerges a positive impact on student performance which may be related both to the effect of school quality and to the incentives to provide a higher level of effort to get the renewal of the voucher.

Kremer, Miguel and Thornton (2009) examine the effects of a program conferring merit scholarships to girls attending primary school in Kenya. From this study it emerges that merit scholarships have produced a number of positive effects: girls eligible for the scholarship have achieved better results compared to those attending schools in the control group; it has also been observed an improvement in teacher attendance and positive spillovers on ineligible students (males) and on students with a low probability of winning the scholarship.

Leuven, Oosterbeek and van der Klaauw (2009) have conducted an experiment involving students enrolled at the University of Amsterdam. Students were randomly assigned to two treatment groups: a high reward group to whom it was promised a bonus of 681 euro for completing all first year requirements by the start of the following academic year; a low reward group to whom, for the same achievement, it was promised a bonus of 227 euro. From the experiment it emerges that only students with higher abilities and those with a highly educated father showed better academic performances when assigned to the high reward group. On the other hand, for the whole sample there was no significant effect of the incentives on achievement, maybe because the requirements for the rewards were too demanding for the average student.

All in all, results emerging from the existing literature do not lead to clear results about the effects of financial incentives on student performance, probably because the experiments were conducted in very different economic and social contexts.

In this paper we investigate the effects of monetary incentives on student performance in a Southern Italian region (Calabria), an area where – due to a relatively low level of development and to the imperfections characterizing the labour market – individuals may lack adequate incentives to provide effort in order to increase their human capital.

The effort provided by students in learning activities is one of the most important factors in human capital accumulation, presumably more important than teacher quality and school resources (Stinebrickner and Stinebrickner, 2008; Costrell, 1994; Bonesronning, 2004). However, it is a costly activity in terms of disutility and opportunity costs and students may choose to not exert an adequate amount of study effort.

The labour markets in Southern Italian regions are far from being competitive and are afflicted by high level of unemployment. Workers' effective skills tend to be not adequately rewarded because of the importance that family background and social networks play in finding a job (Pistaferri, 1991; Fabbri and Rossi, 1997), the high relevance of public sector employment, in which meritocratic systems for determining pay and promotions are rarely used (see Alesina, Danninger and Rostagno, 2001), and the industrial system being specialized in traditional and technologically poor sectors requiring low skills.

The low economic rewards offered to skills may discourage students from studying hard (De Paola and Scoppa, 2007) and may be, at least in part, responsible for the high dropout rates and the excessive duration of student academic careers. Distorted labour markets and low student effort may also have a role in explaining the huge differences between Northern and Southern Italian students in performances in international test scores, such as PISA or TIMSS.

The regional government of Calabria, aware of the low level of basic competences acquired by students living in the area and of the implications this has on social and economic development, with the financial support of the European Union (through the European Social Fund), has planned to introduce a programme awarding scholarships to students with high achievements. The randomized experiment presented in this paper has been funded within this policy intervention, with the purpose to evaluate whether financial incentives and merit scholarships may represent an effective tool to improve student performance.

The experiment has involved 462 first year undergraduate students enrolled in the academic year 2008-2009 in the Business Administration Degree Course (B.A.) offered by the University of Calabria. Students were assigned, through a stratified random procedure, to one control group (157 students) and to two "treatment groups", eligible for the attainment of a small (€250) or a large (€700) monetary prize on the basis of their academic performance (respectively, 150 and 155 students). The experiment is aimed at disentangling the effects of financial incentives on academic performance in a depressed labour market where normal incentives to acquire skills may lack, distinguishing the effect that simply derives from the existence of an incentive from that related to its size. Prizes were assigned to the 30 best performing students in each treated group, in terms of credits earned and grades obtained at examinations.

In order to have information on students' effective awareness and interest in the project, we required them to fill in a participation form. About 90% of students assigned to treated groups have joined the experiment.

From our analysis it emerges that financial rewards have been effective in enhancing student performance. The prospective of obtaining an award of €700 significantly increases student performance (by about 0.19 standard deviation of the target variable). Competing for an award of €250 produces an effect that is slightly below (about 0.16 standard deviation).

We carry out our analysis firstly considering all the students enrolled, independently from the fact that they accepted to participate in the experiment. Therefore, we estimated the so-called "intention-to-treat" effect that, from a policy-point of view, may be particularly interesting. Subsequently, we investigate the impact of the effective participation to the treatment, using as an instrument the random assignment to the groups and we show that adjusting the "intention-to-treat" effect for non-compliance leads to a (stronger) highly significant effect of financial incentives.

The effects are heterogeneous among high and low ability students: high ability students in the treated groups perform significantly better than students of comparable ability in the control group, while low ability students in the treated and control groups do not show any significant difference in performance. This is probably related to the fact that the low ability students, facing a low ex-ante probability of obtaining the prize (only the 30 best students were rewarded out of about 150 students in each group), are discouraged and not sufficiently motivated to provide the sizeable effort needed for them to obtain the reward.

Our results show that treated students have improved their performance both in terms of credits earned and grades obtained (as required by the experimental design), but not at the expense of a worse performance in examinations that were not included among those targeted in the experiment. In fact, in these classes students have obtained similar performances, regardless of the group in which they were assigned. Therefore, the improvement in performance for treated students can be ascribed to a higher level of effort.

The paper is organized as follows. In section 2 the design of the experiment is explained and some information on the Italian University system are provided. Section 3 shows the effects of the financial rewards on credits and grades earned by students. In Section 4 we analyse the treatment effects using student-course level observations. Section 5 examines student passing rates on non-target examinations. In Section 6, using an IV strategy, we investigate the treatment effects on students who have effectively participated to the experiment. Section 7 discusses the answers given by students in a post-experiment survey. Section 8 concludes.

2. Experiment Description and Data

The experiment we have conducted has involved 462 students enrolled at the first year of the First Level Degree Course in Business Administration ("Laurea Triennale in Economia Aziendale") at the University of Calabria in the academic year 2008-2009.[1] The University of Calabria is a middle-sized public university located in the South of Italy. It has currently about 34,000 students enrolled in different Degree Courses and at different levels of the Italian University system, which, since the reform of 2001, is organized around three main levels: First Level Degrees (3 years of legal duration), Second Level Degrees (2 years more) and Ph.D Degree. In order to gain a First Level Degree students have to acquire a total of 180 credits. Students who have acquired a First Level Degree can undertake a Second Level Degree.[2]

The First Level Degree Course in Business Administration is highly demanded and enrolls a large number of students in the first year.[3] The first year includes 6 compulsory courses: 1) Calculus; 2) Accounting; 3) Statistics; 4) Microeconomics (each being a 10 credits class); 5) Public Law and 6) Business Administration (5 credits each). In addition, students have to pass 3 additional examinations (English, French and Computer Sciences), for a total of 10 credits, which we do not include in our target variable (see below). In the first year all the students have to take the same classes and cannot make any choice of courses.

Given the high number of students enrolled, they are assigned, on the basis of the alphabetic order of their surnames, to two different teaching classes (Class 1 (A-L) and 2 (M-Z)) for all the first-year compulsory courses. Students in the two teaching classes attend courses with the same programmes, each offered in the same term.

At the beginning of the academic year, students were informed of the experiment both through presentations during classes and through a letter, sent to all students, explaining the purpose and the format of the experiment.[4]

On the basis of the available administrative information on students' characteristics, we proceeded to the stratification of students according to the following variables: gender, class in which students attend courses (Class 1 and 2), type of High School attended (3 categories: a) Lyceum; b) Technical schools; c) Vocational schools and other types of schools) and grade obtained at Final High School examination (split in 4 categories corresponding to quartiles). Following this procedure we obtained 47 (out of 48) non null groups. Within each group, 1/3 of students was randomly assigned to the large reward group "*Treatment A*" (with a reward of €700), 1/3 to the low reward group

[1] Budget constraints did not allow us to extend the experiment to students enrolled in other Degree Courses.
[2] After having accomplished their Second Level Degree, students can enroll in a Ph.D degree.
[3] Applications exceeded the number of students to be enrolled, although it is the degree course with the largest number of enrolled students at the University of Calabria.
[4] Moreover, a web-page on the university website provided information to students on the experiment.

"*Treatment B*" (with a reward of €250), and 1/3 to the *Control Group*.[5] We ended up with 155 students assigned to *Treatment A*, 150 to *Treatment B* and 157 to the control group.

The random assignment procedure was carried out at the presence of students. Students were also informed by e-mail of their assignment status and the list of students belonging to each group was published on the experiment web-page.

In order to have an indicator of student awareness and interest in the experiment, we asked students in treatment groups to fill in a participation form and we announced that students who did not comply were ineligible for the prizes. Approximately, 91% of students in "*Treatment A*" and 88% of students in "*Treatment B*" filled in their participation form.

Prizes were promised to the 30 best performing students in each treatment group. To avoid dysfunctional responses by the incentive system, deriving from the attempt of students to maximize the targeted measure at the expense of other objectives ("multitasking" in the definition of Holmstrom and Milgrom, 1991), student performance was measured considering both grades obtained at examinations and the number of credits earned. More precisely, our targeted measure of performance "*Total Points*" is equal to the sum of the grades obtained in each class.[6] In this way, we avoided both the risk that students aim to pass many examinations but with low grades and the risk that they seek to obtain high grades but passing only a few examinations.

Passing grades range from 18 to "30 cum laude", which we consider equal to 31. Only exams undertaken within the 31st of July 2009 were taken into account in determining student performance. For examinations based on multiple choice tests and written tests (Calculus, Accounting, Statistics and Microeconomics) in defining our target variable we have also considered grades obtained in failed examinations (below the minimum passing line of 18/30). On the other hand, for examinations for which grades below the minimum passing line were not recorded (Public Law and Business Administration), we have considered a grade equal to zero for fails.

Table 1 provides descriptive statistics for the sample of students. About 55 percent of the students were female. The average age was 20.5. High School grade ranged from 60 (the minimum passing grade) to 100 (the maximum grade), with a mean of 89.4. Students mainly came from Scientific Lyceums (about 33%) and Technical Schools (about 40%). About 21% attended a Vocational School and 6% a Classical Lyceum. 57% of students attended lectures in Class 1 (surnames from A to L). At the end of July students on average earned 18 credits, with an average grade on passed examinations of 22.4. The target variable *Total Points* has a mean of 43.3 with a standard deviation of 32.

---

[5] When the number of students included in each stratified group was not a multiple of 3, one or two students were first assigned randomly to the treatment or control groups.

[6] We have divided by two the grades obtained at classes allowing to acquire 5 credits (Business Administration and Public Law).

Table 1. Descriptive Statistics

Variables	Obs.	Mean	Std. Dev	Min.	Max.
Female	462	0.554	0.498	0	1
Age	462	20.515	2.514	19	51
High School Grade	462	89.435	8.737	60	100
Scientific Lyceum	462	0.325	0.469	0	1
Classical Lyceum	462	0.061	0.239	0	1
Technical Schools	462	0.400	0.491	0	1
Vocational Schools	462	0.214	0.411	0	1
Class 1: A-L	462	0.569	0.496	0	1
Total Points	462	43.286	32.376	0	139.5
Credits	462	18.468	12.768	0	49
Average Grade	383	22.402	2.554	18	31

Notes: Grades in each course ranges from 18 to "30 cum laude" (set equal to 31). High School Grade ranges from 60 to 100.

In the first three columns of Table 2 are reported, by treatment groups, means for a number of individual characteristics. In columns 4 and 5 are reported differences in means respectively between *Treatment A* and Control, and *Treatment B* and Control (standard errors are reported in parentheses). In column 6 we report the *F*-stat (and *p*-value) for a test of equality of variable means across all three groups.

Results show that the randomization has been successful in creating comparable treatment and control groups along the observable characteristics: there are no significant differences by treatment status in students' sex, age, High School grade, type of High School attended and classes in which they attended first year courses.

Table 2. Student characteristics across treatment and control groups

	Means			Differences (s.e.)		*F*-stat (*p*-value)
	Treatment A	Treatment B	Control	Treatment A v. Control	Treatment B v. Control	
Female	0.529	0.580	0.554	-0.025	0.026	0.399
				(0.057)	(0.057)	(0.671)
Age	20.723	20.360	20.459	0.264	-0.099	0.853
				(0.320)	(0.228)	(0.427)
High School Grade	89.432	89.060	89.796	-0.364	-0.736	0.271
				(0.982)	(0.993)	(0.762)
Scientific Lyceum	0.323	0.300	0.350	-0.028	-0.050	0.443
				(0.054)	(0.054)	(0.642)
Classical Lyceum	0.065	0.080	0.038	0.026	0.042	1.206
				(0.025)	(0.027)	(0.300)
Technical Schools	0.400	0.393	0.408	-0.008	-0.014	0.033
				(0.056)	(0.056)	(0.968)
Vocational and other Schools	0.213	0.227	0.204	0.009	0.023	0.119
				(0.046)	(0.047)	(0.887)
"Class 1: A-L"	0.568	0.573	0.567	0.001	0.006	0.008
				(0.056)	(0.057)	(0.992)

Notes: Standard errors are reported in parentheses. In the last column we report the *F*-stat and *p*-value for a test of equality of variable means across all three groups.

3. Empirical Results

In this section in order to analyse the effects of financial incentives on student performance we estimate by OLS the following model:

$$Y_i = \beta_0 + \beta_1 (Treatment\_A_i) + \beta_2 (Treatment\_B_i) + \phi X_i + u_i$$

where Y_i is a measure of performance (alternatively, *Total Points* or *Credits*) of student *i*, X_i is a vector of individual characteristics (gender, age, High School grade, type of High School dummies, assigned teaching class, dummies for province of residence,[7] an interaction term between High School grade and Lyceum); *Treatment* A_i is a dummy variable which takes value of 1 if student *i* has been assigned to the group competing for an award of €700 (and zero otherwise), and *Treatment* B_i is a dummy variable which takes value of 1 if student *i* has been assigned to the group competing for an award of €250 (and zero otherwise); u_i is an error term capturing idiosyncratic shocks or unobserved student characteristics.

The coefficients β_1 and β_2 capture the effects of the financial incentives: β_1 measures the average impact on student performance of competing for an award of €700 (with respect to the control group) and the coefficient β_2 measures the average impact of competing for an award of €250.

We firstly focus on *Total Points*, the target variable on which the merit scholarships were assigned. OLS estimates are reported in Table 3. In all the specifications, standard errors are robust to heteroskedasticity.

Our main specification is reported in column (1). The prospective of obtaining an award of €700 increases student performance, *ceteris paribus*, of 6.02 points, that is, 0.19 standard deviations of *Total Points*. The coefficient is significant at the 5 percent level. Competing for an award of €250 increases performance of 5.35 points, 0.16 standard deviations of *Total Points*, significant at the 10 percent level, slightly below the effect produced by *Treatment A*.

The effects of control variables are consistent with the findings emerging from the existing literature. Students with a higher *High School Grade* obtain a much better academic performance. In addition, students who attended a *Lyceum* perform much better than students who come from *Technical* or *Vocational Schools*. We have also included among our regressors an interaction term between *High School Grade* and *Lyceum* which turns out to be positive and highly significant, implying that the High School grade increases academic performance significantly more for Lyceum students. Older students perform significantly worse. The dummy *Female* is always not statistically significant and there are no significant differences among the two teaching classes in which students attended classes.

[7] Provinces correspond to the Nuts 3 Eurostat classification.

To evaluate the magnitude of the estimated treatment effects it is useful to consider that the effect of being included in *Treatment A* corresponds to the effect produced by an increase in student ability, measured by *High School Grade*, of about 5 points.

Given the structure of the experiment, with a competition among students in which only the best performing ones were rewarded in each treated group, we expect a different effect of the treatment according to student ability. In order to investigate this issue, we estimate separately the effects for students of ability, respectively, above and below the median *High School Grade* of 90 (columns 2 and 3). It emerges that the effect of financial rewards is negligible for lower ability students, whereas both the small and the large reward produce a positive effect for higher ability students. The effect of the large reward is quite large (9.5 points more, an increase of 0.27 standard deviations) significant at the 5 percent level, while the effect of the small reward, equal to about 6 points, is rather imprecisely estimated (*p*-value 0.20).

Therefore, high ability students in the large reward group perform significantly better than students of comparable ability in the control group. On the other hand, low ability students in the treated and control groups do not show any significant difference in performance. This is probably related to the fact that these students, facing a low ex-ante probability of obtaining the prize, are typically discouraged to provide the significant amount of effort needed for them to perform well. However, treatment coefficients are never negative, implying that even for the low ability students the negative effects that financial incentives might produce on the "intrinsic motivation" – related to the desire to reach the objective for itself rather than for a financial reward – are not relevant or are not strong enough to counterbalance the positive effects that these incentives produce on the "extrinsic motivation".

Table 3. The Impact of the Incentives on Student Performance (measured by *Total Points*). OLS regressions

	(1) All	(2) Top	(3) Bottom	(4) Males	(5) Females
Treatment A (€700)	6.023**	9.488**	0.408	5.390	5.841
	(3.059)	(4.710)	(3.612)	(4.615)	(4.061)
Treatment B (€250)	5.350*	5.889	4.017	2.354	6.157
	(3.164)	(4.608)	(4.118)	(4.877)	(4.207)
Female	-0.531	-3.344	-1.919		
	(2.644)	(4.102)	(3.088)		
Age	-1.217**	-2.981***	-0.618	-0.388	-0.865
	(0.516)	(0.882)	(0.538)	(0.754)	(1.108)
High School Grade	1.224***	2.689***	0.878**	1.105***	1.398***
	(0.203)	(0.618)	(0.427)	(0.297)	(0.262)
Scientific Lyceum	31.815***	29.718***	35.185***	39.358***	26.287***
	(3.550)	(9.057)	(7.011)	(5.419)	(4.890)
Classical Lyceum	13.424**	9.308	15.199	-1.556	13.422*
	(6.522)	(11.346)	(12.107)	(13.145)	(7.764)
Technical Schools	15.429***	22.311***	8.603**	22.129***	9.487*
	(3.398)	(4.982)	(4.179)	(4.835)	(4.940)
Class 1: A-L	1.630	1.324	1.259	-1.370	3.968
	(2.585)	(3.760)	(3.525)	(4.171)	(3.405)
Lyceum* (High School Grade)	0.845***	1.689	1.550***	0.965**	0.773*
	(0.290)	(1.040)	(0.590)	(0.487)	(0.400)
Constant	-65.112***	-175.087***	-40.860	-75.677**	-84.283**
	(20.464)	(60.892)	(33.152)	(30.190)	(36.264)
Observations	462	236	226	206	256
R-squared	0.327	0.356	0.252	0.380	0.317

Notes: The Table reports OLS estimates. The dependent variable is *Total Points*. Standard errors (reported in parentheses) are corrected for heteroskedasticity. In all the regressions we control for province of residence dummies (6 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

In columns (4) and (5) we report separate estimates for males and females respectively. The effects of the incentives turn out to be more pronounced for females, particularly for the low reward group. However, the treatment effects are rather imprecisely estimated (for females the *p*-values of the treatment effects are about 0.14).

In Table 4 are reported estimates considering as dependent variable the number of credits earned by students, *Credits*. Results are very similar to those obtained measuring student performance using *Total Points*. The average treatment effects of both the large and the small financial reward are positive and statistically significant (at the 5 and 10 percent level, respectively). The treatment effects are equal to about 0.17-0.18 in terms of standard deviations of the dependent variable. The magnitude of the two treatment effects is quite similar, implying that the size of the prize is not particularly relevant for the average student in the population. However, when we estimate separately according to student ability, it emerges that students with higher ability strongly react to the large reward (earning 3.8 credits more than the control group), while the effect of the small reward is weaker. Finally, as shown in columns (4) and (5), females are more reactive to financial incentives than males.

Table 4. The Impact of the Incentives on *Credits* Earned. OLS regressions

	(1) All	(2) Top	(3) Bottom	(4) Males	(5) Females
Treatment A (€700)	2.335**	3.827**	0.065	1.759	2.490*
	(1.197)	(1.833)	(1.532)	(1.854)	(1.518)
Treatment B (€250)	2.194*	2.529	1.616	0.714	2.766*
	(1.266)	(1.804)	(1.733)	(1.970)	(1.655)
Female	-0.245	-1.436	-0.592		
	(1.069)	(1.623)	(1.319)		
Age	-0.554***	-1.182***	-0.342	-0.189	-0.476
	(0.205)	(0.329)	(0.231)	(0.308)	(0.453)
High School Grade	0.471***	1.063***	0.382**	0.453***	0.517***
	(0.082)	(0.246)	(0.184)	(0.120)	(0.107)
Scientific Lyceum	12.566***	12.930***	13.884***	15.387***	10.367***
	(1.408)	(3.687)	(2.901)	(2.108)	(1.972)
Classical Lyceum	5.393**	5.843	5.191	-1.173	5.164
	(2.681)	(4.634)	(4.887)	(5.546)	(3.158)
Technical Schools	6.194***	8.666***	3.727**	9.293***	3.304
	(1.393)	(2.030)	(1.759)	(1.987)	(2.013)
Class 1: A-L	0.749	0.422	0.727	-0.737	2.005
	(1.035)	(1.471)	(1.472)	(1.700)	(1.334)
Lyceum* (High School Grade)	0.251**	0.394	0.511**	0.261	0.240
	(0.112)	(0.401)	(0.248)	(0.185)	(0.155)
Constant	-21.962***	-67.989***	-15.896	-29.225**	-26.257*
	(8.235)	(24.244)	(14.124)	(12.181)	(14.867)
Observations	462	236	226	206	256
R-squared	0.314	0.346	0.246	0.376	0.301

Notes: The Table reports OLS estimates. The dependent variable is *Credits*, the number of credits earned. Standard errors (reported in parentheses) are corrected for heteroskedasticity. In all the regressions we control for province of residence dummies (6 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

To further investigate the differences existing between high and low ability students, instead of using the *High School Grade* as a measure of ability, we follow the approach of Angrist and Lavy (2009) and use the predicted performance of students (both in terms of *Total Points* and *Credits*) in order to obtain a composite measure of student ability and to divide students into two equal-size groups.

We first estimate a model for student performance considering exclusively students in the control group, using as explanatory variables individual characteristics such as gender, age, type of high school attended, high school grade. The *High School Grade*, used above as an alternative variable to identify students of different ability, turns out to be a key determinant of student performance. The estimated coefficient of this model are then used for predicting the performance of students in treated and control groups, on the basis of their effective characteristics.

In Table 5 are reported separately the effects for students with a predicted ability, respectively, above and below the median. In columns (1) and (2) the dependent variable is *Total Points*, while in columns (3) and (4) student performance is measured considering *Credits*. It emerges that this classification scheme performs very well in identifying students who are more likely to be affected by the treatment. Estimation results show that both the small and the large financial reward produce a strong positive effect on the academic performance of high ability students (both when considering *Total Points* and *Credits*): the effects are around 0.33 standard deviations of the dependent variable.

The effect of the large reward is almost the same than that produced by the small reward. On the other hand, low ability students in the treated and control groups do not show any significant difference in performance.

Table 5. Differences between High and Low Ability Students (distinguished by predicted student performance). OLS regressions

	Total Points		Credits	
	(1) Top	(2) Bottom	(3) Top	(4) Bottom
Treatment A (€700)	10.783**	1.435	4.139**	0.625
	(4.760)	(3.819)	(1.806)	(1.627)
Treatment B (€250)	10.697**	-0.088	4.393**	-0.058
	(4.557)	(4.146)	(1.752)	(1.761)
Observations	231	231	231	231
R-squared	0.252	0.136	0.243	0.132

Notes: The Table reports OLS estimates. Standard errors (reported in parentheses) are corrected for heteroskedasticity. In all the regressions we control for gender, age, High School grade, type of High School dummies, assigned teaching class, dummies for province of residence, an interaction term between High School grade and Lyceum. The symbol ** indicates that coefficients are statistically significant at the 5 percent level.

4. The Effects of the Incentives Using Course-Level Observations

In this section we stack data with the aim to use student-course level observations (462 students * 6 courses) to estimate the treatment effects on the student probability of passing an exam and on the grade obtained at each exam.

In Table 6 we report estimates of a Linear Probability Model in which the dependent variable, *Pass*, is a dummy equal to 1 when the student passed a given exam, and 0 otherwise. The mean value of *Pass* is 0.38 and its standard deviation is 0.48. In addition to the usual individual characteristics, we control for course dummies (Calculus is the reference category) to take into account heterogeneous levels of difficulty among courses and/or different grading standards among instructors. Standard errors are robust to heteroskedasticity and are clustered at the student level.

Students in the treatment groups show a higher probability of passing exams: the incentives generate an increase of about 5 percentage points. The effect is statistically significant at the 5 and 10 percent level for the large and the small reward, respectively.

When we run separate regressions for students of different ability, it emerges that only students with a level of ability above the average react to the financial incentives (column 2), while incentives do not have any impact on low ability students (column 3). Results in columns 4 and 5 confirm that females are more reactive than males to the prospective of obtaining a monetary reward.

Very similar results are obtained using a probit estimator instead of a linear probability model (results are available upon request).

Table 6. Incentives and the Probability of Passing an Exam. Linear Probability Model

	(1) All	(2) Top	(3) Bottom	(4) Males	(5) Females
Treatment A (€700)	0.051**	0.087**	-0.002	0.033	0.059*
	(0.025)	(0.036)	(0.031)	(0.037)	(0.032)
Treatment B (€250)	0.048*	0.065*	0.027	0.018	0.055*
	(0.026)	(0.036)	(0.036)	(0.041)	(0.033)
Accounting	0.500***	0.432***	0.571***	0.451***	0.539***
	(0.024)	(0.034)	(0.034)	(0.036)	(0.033)
Microeconomics	0.043*	0.042	0.044	0.010	0.070**
	(0.023)	(0.031)	(0.034)	(0.036)	(0.029)
Statistics	-0.251***	-0.326***	-0.173***	-0.223***	-0.273***
	(0.021)	(0.033)	(0.027)	(0.032)	(0.029)
Business Administration	0.063**	0.004	0.124***	0.019	0.098**
	(0.027)	(0.042)	(0.035)	(0.040)	(0.038)
Public Law	0.069**	0.064	0.075**	0.049	0.086**
	(0.028)	(0.043)	(0.037)	(0.041)	(0.039)
Female	-0.003	-0.011	-0.020		
	(0.022)	(0.032)	(0.027)		
Age	-0.012***	-0.024***	-0.007	-0.003	-0.009
	(0.004)	(0.006)	(0.005)	(0.006)	(0.010)
High School Grade	0.010***	0.021***	0.007*	0.008***	0.012***
	(0.002)	(0.005)	(0.004)	(0.002)	(0.002)
Technical Schools	0.130***	0.180***	0.075**	0.184***	0.078*
	(0.028)	(0.041)	(0.036)	(0.039)	(0.041)
Scientific Lyceum	0.254***	0.259***	0.286***	0.326***	0.201***
	(0.029)	(0.074)	(0.059)	(0.042)	(0.039)
Classical Lyceum	0.133**	0.147	0.134	-0.008	0.131**
	(0.056)	(0.093)	(0.102)	(0.117)	(0.066)
Class 1: A-L	0.006	0.002	0.005	-0.026	0.028
	(0.021)	(0.029)	(0.031)	(0.034)	(0.027)
Lyceum* (High School Grade)	0.005**	0.007	0.011**	0.007*	0.004
	(0.002)	(0.008)	(0.005)	(0.004)	(0.003)
Constant	-0.519***	-1.380***	-0.333	-0.575**	-0.722**
	(0.166)	(0.481)	(0.296)	(0.240)	(0.303)
Observations	2772	1416	1356	1236	1536
R-squared	0.298	0.296	0.304	0.302	0.314

Notes: The dependent variable is the dummy *Pass*. Standard errors (reported in parentheses) are corrected for heteroskedasticity and clustered at the student level. In all the regressions we control for province of residence dummies (6 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

In Table 7 we investigate, using the same specifications as in Table 6, the effects of financial incentives on the grades obtained by students at exams. Results from OLS estimates suggest that the incentives are effective not only in enhancing the student probability of passing exams, but also in increasing the grades they obtain. Being included in the large reward group increases student grades by about 1.3 points (statistically significant at the 5 percent level). Once again, treatment effects turn out to be significant for high ability students but not for low ability students. Females react more than males, but treatment effects are imprecisely estimated.

Table 7. Incentives and Grades Obtained at Examinations. OLS regressions

	(1) All	(2) Top	(3) Bottom	(4) Males	(5) Females
Treatment A (€700)	1.327**	2.171**	-0.095	0.993	1.283
	(0.597)	(0.887)	(0.757)	(0.888)	(0.796)
Treatment B (€250)	0.930	1.615*	0.734	0.525	1.197
	(0.630)	(0.912)	(0.851)	(0.977)	(0.826)
Accounting	11.227***	10.822***	12.757***	10.587***	12.719***
	(0.420)	(0.730)	(0.709)	(0.753)	(0.692)
Microeconomics	4.221***	1.843***	1.558**	0.932	2.324***
	(0.407)	(0.684)	(0.725)	(0.788)	(0.636)
Statistics	-5.719***	-7.000***	-3.442***	-4.602***	-5.789***
	(0.428)	(0.717)	(0.568)	(0.705)	(0.624)
Business Administration	0.268	1.089	3.496***	1.112	3.195***
	(0.591)	(0.965)	(0.781)	(0.893)	(0.870)
Public Law	-0.199	1.763*	1.836**	1.126	2.340***
	(0.586)	(0.967)	(0.780)	(0.859)	(0.890)
Female	0.488	0.110	-0.659		
	(0.519)	(0.819)	(0.644)		
Age	-0.350***	-0.605***	-0.115	-0.071	-0.150
	(0.103)	(0.169)	(0.089)	(0.149)	(0.230)
High School Grade	0.223***	0.506***	0.138***	0.198***	0.298***
	(0.039)	(0.128)	(0.053)	(0.056)	(0.052)
Technical Schools	3.174***	4.348***	1.658*	4.253***	2.062**
	(0.672)	(1.037)	(0.843)	(0.919)	(0.983)
Scientific Lyceum	6.012***	5.803***	7.309***	8.133***	5.085***
	(0.713)	(1.768)	(1.307)	(1.052)	(0.959)
Classical Lyceum	3.063**	2.475	3.687	0.079	3.322**
	(1.364)	(2.122)	(2.377)	(2.714)	(1.602)
Class 1: A-L	-0.344	0.255	-0.044	-0.509	0.595
	(0.512)	(0.745)	(0.676)	(0.811)	(0.672)
Lyceum* (High School Grade)	0.140**	0.316	0.338***	0.219**	0.136*
	(0.059)	(0.206)	(0.109)	(0.094)	(0.080)
Constant	-8.757**	-33.364***	-7.272	-14.798**	-21.316***
	(4.124)	(12.511)	(5.241)	(5.740)	(7.158)
Observations	2772	1416	1356	1236	1536
R-squared	0.330	0.293	0.294	0.310	0.315

Notes: The dependent variable is *Grade*. Standard errors (reported in parentheses) are corrected for heteroskedasticity and clustered at the student level. In all the regressions we control for province of residence dummies (6 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

In the estimates shown in Table 7 we have considered, when possible, also grades obtained in failed examinations (below the minimum passing line of 18/30), while for those examinations for which grades below the minimum passing line were not recorded (Public Law and Business Administration), the grade was set equal to zero. As an alternative strategy, we have considered all grades below the minimum passing line as censored and we have estimated a Tobit model considering 17 as the lower limit for left censoring and 31 as the upper limit for right censoring. Results (not reported) are quite similar to those presented in Table 7. In particular, on the basis of the Tobit estimations, we are able to determine the marginal effect of the treatment on the probability that a student passed an exam (that is, the effect on the probability that the grade is greater than 17): this effect is equal to 0.06. Moreover, given that a student has passed the exam (i.e., he/she is not censored), we can calculate the marginal effect of the treatment on the expected grade: being included in the large reward group increases the expected grade of 0.29.

5. Higher Effort or Substitution Effects?

Cornwell, Lee and Mustard (2005) and Binder, Ganderton, and Hutchens (2002) find that scholarships have been effective in increasing student grades, but have also induced students to take less ambitious course loads. In contrast to these works, our findings show that financial rewards did not lead to adverse changes in student behaviour. Our sample students have improved both grades and credits earned, suggesting that the improvement in their academic performance can be attributed to a higher level of effort rather than to the attempt of "gaming" the incentive system. Given the experimental design adopted, in which we define the target both in terms of grades and in terms of number of credits earned, this is somehow an expected result. Students had low chances of gaining the monetary rewards by substituting effort from non targeted tasks to targeted ones, with the aim of keeping their effort level unchanged.

To further investigate this issue we analyse student passing rates on a number of examinations that were not defined as a target for the experiment. As mentioned in Section 2, students enrolled at the first year of the Degree in Business Administration are required to take some additional examinations (English, French and Computer Sciences), which were not included in the target variable *Total Points*. The choice to exclude these exams was made since for them no grade is awarded and only a pass or fail result is observed.[8]

Table 8 reports Linear Probability Model estimates of student passing rates in these examinations using student-course level observations (462*3=1386 observations). We do not observe any significant difference among students in treated and control groups in the performance on non-target examinations.

This result has two important implications. Firstly, without financial incentives students in the treated groups do not outperform those in the control group, confirming the effectiveness of monetary incentives in enhancing student performance. Secondly, it suggests that "multitasking" problems are not a major concern in our experiment: students eligible for the financial rewards – while increasing effort to improve their "measurable" performance – did not reduce their effort in other studying activities not directly relevant for obtaining the prize.

Therefore, we can conclude that the improvement in student achievements discussed in the previous sections has been determined by a higher provision of effort in studying activities.

[8] Moreover, students may have very different initial levels of unobservable competence in these specific subjects.

Table 8. Probability of Passing the Exams not Included in the Target Variable. Linear Probability Model

	(1) All	(2) Top	(3) Bottom	(4) Males	(5) Females
Treatment A (€700)	0.043	0.031	0.030	0.053	0.034
	(0.032)	(0.044)	(0.045)	(0.053)	(0.038)
Treatment B (€250)	-0.026	-0.019	-0.024	-0.062	-0.010
	(0.034)	(0.049)	(0.044)	(0.053)	(0.044)
Observations	1386	708	678	618	768
R-squared	0.246	0.244	0.296	0.275	0.209

Notes: The Table reports OLS estimates. The dependent variable is *Pass* in English, French and Computer Sciences examinations. Standard errors (reported in parentheses) are corrected for heteroskedasticity and clustered at the student level. In all the regressions we control for gender, age, High School grade, type of High School dummies, assigned teaching class, dummies for province of residence, an interaction term between High School grade and Lyceum.

We have also tried to assess whether the increased level of student effort has taken the form of a higher attendance to teaching classes. Students can improve their performance devoting more time to studying activity at home, attending a higher number of hours in classes or participating more regularly at tutorials with the teaching assistants. We do not have information on time spent in studying activities, but we have collected data on student attendance in five randomly chosen dates through unannounced checks. We have used these data to build a variable *Attendance* equal to the number of times the student was found attending a class. As shown by OLS estimates, reported in Table 9, in which *Attendance* is regressed on the treatment dummies and on the usual controls, the financial reward does not increase in a significant way student physical presence in classes. We find a positive and weakly significant effect (*p*-value=0.25) for *Treatment A* and a null effect for *Treatment B*. These findings suggest that the better performance of treated students is not related to higher classes attendance.

Table 9. The impact of financial incentives on class attendance. OLS regressions

	(1) All	(2) Top	(3) Bottom
Treatment A (€700)	0.162	0.009	0.278
	(0.149)	(0.230)	(0.189)
Treatment B (€250)	-0.069	-0.178	0.050
	(0.145)	(0.213)	(0.190)
Observations	462	236	226
R-squared	0.144	0.190	0.120

Notes: The Table reports OLS estimates. The dependent variable is *Attendance.* Standard errors (reported in parentheses) are corrected for heteroskedasticity. In all the regressions we control for gender, age, High School grade, type of High School dummies, assigned teaching class, dummies for province of residence, an interaction term between High School grade and Lyceum.

6. Dealing with Partial Compliance using an IV Estimation Strategy

In the previous sections we have analysed intention-to-treat effects, since we have considered the whole population of students enrolled at the first year of the Business Administration Degree Course, although not all students signed up to show their interest and awareness of the experiment. Therefore, the estimated effects are diluted by the fact that some treated students may actually not have been informed of the experiment or may not been interested in it (see Angrist and Pischke 2009; Bloom, 1984).

In this section we analyse the impact of the financial incentives on the students who have effectively participated to the treatment, that is, those students who filled in the participation form. To deal with endogeneity problems that may affect participation, following the literature we instrument the effective participation in each group with the two random assigned treatments (*Treatment A* and *Treatment B*).

In Table 10 are reported Two-Stage-Least Squares (*TSLS*) estimates using as dependent variable *Total Points* (columns 1-3) and *Credits* (columns 4-6). Results show that adjusting intention-to-treat effects for non-participation leads to a stronger impact of the treatment on student performance measured both as *Total Points* and *Credits*. The increase in the effect is proportional to the reciprocal of the participation rate.

The average treatment effect of the €700 reward is an increase of 14% in *Total Points* (with respect to the sample mean), while the effect of the small reward is of 13%. The effects are higher when focusing on high ability students who, when competing for obtaining the prize of €700, get about 10 points more, that is, an increase of 19% with respect to the average *Total Points* realized by students of comparable ability.

Table 10. TSLS estimates of the effects of effective participation to the treatment

	Total Points			Credits		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Top	Bottom	All	Top	Bottom
Treatment A Effective	6.610**	9.747**	0.485	2.562*	3.932**	0.078
	(3.335)	(4.820)	(4.274)	(1.327)	(1.878)	(1.815)
Treatment B Effective	6.029*	6.510	4.590	2.473*	2.796	1.847
	(3.545)	(5.055)	(4.670)	(1.417)	(1.977)	(1.965)
Female	-0.831	-3.547	-2.142	-0.368	-1.527	-0.681
	(2.611)	(4.053)	(3.072)	(1.054)	(1.600)	(1.316)
Age	-1.248**	-3.012***	-0.615	-0.565***	-1.195***	-0.339
	(0.515)	(0.898)	(0.543)	(0.205)	(0.334)	(0.234)
High School Grade	1.184***	2.616***	0.817*	0.455***	1.032***	0.358*
	(0.206)	(0.616)	(0.440)	(0.083)	(0.244)	(0.190)
Scientific Lyceum	31.372***	29.117***	35.243***	12.386***	12.672***	13.911***
	(3.498)	(8.974)	(6.988)	(1.388)	(3.640)	(2.896)
Classical Lyceum	12.844**	8.439	15.351	5.159*	5.475	5.255
	(6.491)	(11.396)	(11.911)	(2.666)	(4.640)	(4.807)
Technical Schools	14.909***	21.903***	8.203*	5.982***	8.489***	3.568**
	(3.380)	(4.956)	(4.162)	(1.387)	(2.017)	(1.756)
Class 1: A-L	1.745	1.499	1.335	0.795	0.497	0.757
	(2.567)	(3.749)	(3.500)	(1.027)	(1.464)	(1.463)
Lyceum* (High School Grade)	0.864***	1.736*	1.600***	0.259**	0.414	0.531**
	(0.290)	(1.034)	(0.590)	(0.111)	(0.397)	(0.249)
Constant	-60.419***	-166.991***	-35.689	-20.103**	-64.542***	-13.902
	(20.231)	(60.627)	(33.463)	(8.126)	(23.928)	(14.296)
Observations	462	236	226	462	236	226

Note. In the First Stage dummies for the assigned treatment are used as instruments for "Treatment A Effective" and "Treatment B Effective". Standard errors (reported in parentheses) are corrected for heteroskedasticity. In all the regressions we control for province of residence dummies (6 categories, not reported). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

The same analysis has been conducted using student-course level observations on the probability of passing examinations and on grades obtained. Results are similar to those discussed above (not reported).

7. Some Evidence from a Post-Experiment Survey

The results shown in our analysis might be related, a part from the financial incentives offered in our project, to a number of other factors taking place simultaneously with the experiment and correlated to the treatment. In order to investigate these aspects, we have conducted a post-experiment survey among our sample students. About 80% of students have answered to the questionnaire.

Firstly, peer effects or other externalities deriving from better performing students by affecting positively performance of students in the control group, may potentially produce a downward bias in our estimation results. To try to tackle this issue, we have asked students the names of other students (maximum 4) they study with the most often. Then, we have built a variable, *Treated Peers*, measuring for each student the number of treated students among his/her peers (without distinguishing between treatment A and B) and run a regression of both *Total Points* and *Credits* on this variable and the usual controls (as in Table 3, column 1). From the estimation results (not reported) the number of treated peers does not seem to affect student performance. The coefficient is far from being statistically significant (p-value=0.79). However, these results are only suggestive since students participating to the survey might not be a random sample of the population of interest.

Secondly, the treatment in our experiment may have been modified by the intervention of other subjects, such as parents, who may have promised other financial on in kind rewards. We have asked students whether parents or other subjects have promised rewards additional to those established by the experiment. From students' answers it clearly emerges that supplementary rewards were rather uncommon. Only 2% of students report that their parents have promised them a monetary prize in case of good performance at their first year examinations.

Thirdly, it could be that the effect we have found has been negatively influenced by poor cooperation among peers induced by the design of the experiment. The experiment had a tournament structure that, as it is well known in the agency literature (see Prendergast, 1999), tends to discourage cooperation. In our post-experiment survey we have asked students whether the experiment has led them to be less cooperative with their peers and whether they have obtained or not cooperation in terms of course materials, information sharing and support in studying activities by their colleagues. About 3% of students admits to have been scarcely cooperative with their peers as a result of the experiment, while about 15% of students says that their colleagues have been scarcely cooperative in studying activities. Students included in the treated groups are both less incline to offer cooperation and less likely to receive cooperation from their peers, giving support to the idea that the tournament structure of the experiment may have, at least in part, reduced cooperation. Therefore, it is likely that the treatment effects we find could have been larger if we had used a reward system based on an absolute, rather than on a relative, measure of performance.

Finally, one might speculate that the instructors might have changed their grading standards according to the student treatment status. We do not think this is a serious concern because most of the examinations we consider is based on multiple choice tests. Moreover, we presented the experiment to the instructors, explaining the aims of the project and asking them complete impartiality in grading policy (they were not informed of students' treatment status).

8. Concluding Remarks

Merit scholarships may constitute an incentive for students to provide more effort and they may help them to improve their attitudes toward studying activities. This is especially true in those labour markets characterized by a number of distortions determining relatively low expected returns for human capital investments and discouraging students to provide high effort in studying activities.

The empirical investigations trying to shed light on the effects of financial incentives for students have reached heterogeneous results, suggesting the need to expand the research in this field in order to better understand which factors are more likely to affect the success of this type of policy.

In this paper we have provided some additional evidence on the effects of financial rewards on student academic performance, reporting the results of an experiment conducted among a sample of undergraduate students enrolled at a Southern Italian public University. The project was funded by the regional government of Calabria with the purpose to evaluate whether financial incentives may represent an effective tool to improve the performance of students living in the area. Southern Italian students, according to the results obtained at international tests, such as PISA or TIMSS, show a very low level of basic competences compared to other students living in the Centre and North of Italy.

From our analysis it emerges that financial incentives have significantly improved student performance. The students included in the treated groups obtain better grades and earned a greater number of credits. Consistently with findings emerging from other works, we also find that gains from the treatment are mainly concentrated among students with higher ability. When included in the treated groups, these students perform significantly better than students of comparable ability in the control group, while the effect of treatment on low ability students is not statistically significant. This is probably related to the fact that the low ability students, facing a low ex-ante probability of obtaining the prize, are typically discouraged from providing higher effort. Therefore, to permit to the low ability students to benefit more from this type of programs it would be necessary either to restrict competition only to them, or to take into account the gains in performance in relation to those expected based on the student initial ability.

Differently from other studies which find that financial incentives may lead to adverse changes in student behaviour, we do not find evidence of any dysfunctional responses. The improvement in student performance detected in our research seems to be imputable to a higher level of effort in studying activities: students in the treatment groups have obtained better grades and earned a higher number of credits, without showing worse results in examinations that were excluded from the experiment target. Furthermore, we do not find any evidence of adverse effects of financial rewards on intrinsic motivation. Treatment coefficients are never negative, implying that even for the low ability students the negative effects that financial incentives may produce on intrinsic motivation are not particularly relevant.

These positive findings may be related to the economic and social context in which our sample students live: in depressed labour markets, such as those characterizing the South of Italy, students may be more reactive to financial incentives since they lack the normal incentives deriving from a well functioning labour market that offers high rewards to skills in terms of better employment opportunities and higher wages.

All in all, the financial incentives given to students have been quite effective in improving student performance, with negligible negative side-effects. In terms of standard deviations, student performance has increased of about 0.18. Compared to other educational policies, the cost of the project – a total outlay of €28,500 in prizes for students – appears also quite small. Performing a "back of the envelope" calculation and considering as an alternative strategy for increasing student performance the reduction of the class size, it emerges that monetary incentives are a cost-effective policy. In fact, if we take for granted Krueger's (1999) estimates showing that a reduction of class size of one third (from 22 to 15) increases student performance by about 0.2 standard deviations and suppose these estimates are valid for our sample of students, a reduction of class size from 230 (the current size) to 155 (organizing an additional teaching class) would have almost the same effect on performance as the monetary incentives used in our experiment. However, the cost of organizing an additional teaching class to bring the number of classes from the current 2 to 3, considering only additional expenses for instructors' remunerations, would imply a cost of at least €40,000, a figure well above the cost of the incentives given to students.

References

Alesina A., Danninger S. Rostagno M., (2001), "Redistribution Through Public Employment. The Case of Italy", *IMF Staff Paper*, 48 (3), pp. 447-473.

Angrist J., Bettinger E., Kremer M., (2006), "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Record", *American Economic Review*, 96(3), pp. 847-862 .

Angrist, J., Bettinger, E., Bloom, E., King, E., Kremer, M. (2002), "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment", *American Economic Review*, 92(5), pp. 1535–58.

Angrist, J., Lang, D., Oreopoulos, P., (2009), "Incentives and Services for College Achievement: Evidence from a Randomized Trial", *American Economic Journal: Applied Economics*, 1(1), pp. 136–163.

Angrist, J., Pischke S., (2009), *Mostly Harmless Econometrics*, Princeton University Press, Princeton.

Angrist, J.,Lavy, V. (2009), "The Effect of High School Matriculation Awards: Evidence from Randomized Trias", *American Economic Review*, 99 (4), pp. 1384–1414.

Binder, M., P.T. Ganderton, K. Hutchens (2002), "Incentive Effects of New Mexico's Merit-Based State Scholarship Program: Who Responds and How?" in Heller, D.E. & P. Marin (eds.), *Who Should We Help? The Negative Social Consequences of Merit Scholarships*, pp. 43-56, Cambridge, MA. Harvard University.

Bloom, H., (1984), "Accounting for No-shows in Experimental Evaluation Designs", *Evaluation Review*, 8, pp.225-246.

Bonesronning, H., (2004), "Do the Teachers' Grading Practices Affect Student Achievement?", *Educational Economics*, 12 (2), pp. 151-167.

Cornwell, C., Lee K., Mustard, D.,(2005) "Student Responses to Merit Scholarship Retention Rules", *Journal of Human Resources*, XL(4), pp. 895-917.

Costrell, R. (1994), "A Simple Model of Educational Standards", *American Economic Review*, Vol. 84, N. 4, pp. 956-971.

De Paola, M., Scoppa, V., (2007), "Returns to skills, incentives to study and optimal educational standards", *Journal of Economics*, 92 (3), pp. 229–262.

Fabbri F. and Rossi N., (1997), "Caste, non classi. Una società immobile", *Il Mulino*, 1, pp. 110-116.

Heller, H., (2006), "Merit Aid and College Access", Symposium on the Consequences of Merit-Based Student Aid, Wisconsin Center for the Advancement of Postsecondary Education.

Holmstrom, B. and Milgrom, P., (1991), "Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership, and Job Design", *Journal of Law, Economics, and Organization*, vol. 7, pp. 24-52.

Kremer, M., Miguel, E.,Thornton, R., (2009), "Incentives to Learn", *Review of Economics and Statistics*, 91(3), pp. 437–456.

Krueger, A., (1999), "Experimental Estimates of Education Production Functions", 114, 2, pp. 497-532.

Leuven, E., Oosterbeek, H. and van der Klaauw B., (2009), "The Effect of Financial Rewards on Students' Achievements: Evidence from a Randomized Experiment", *Journal of the European Economic Association*, forthcoming.

Middleton, S., Maguire, S., Ashworth, K., Legge, K., Allen, T., Perren, K., Battistin, E., Dearden, L., Emmerson, C., Fitzsimons, E., Meghir, C. (2003), The Evaluation of Education Maintenance Allowance Pilots: Three Years Evidence. A Quantitative Evaluation, Department for Education and Skills Research Report No. 499. Nottingham: DfES Publications.

Pistaferri, L., (1999), "Informal Networks in the Italian Labor Market". *Giornale degli Economisti e Annali di Economia*, 58, pp. 355-376.

Prendergast, C., (1999), "The Provision of Incentives in Firms", *Journal of Economic Literature*, vol. 37, n. 1, pp. 7-63.

Schultz, T. P. (2004), "School subsidies for the poor: evaluating Mexico's Progresa poverty program", Journal of Development Economics, 37, pp.199-250.

Stinebrickner, R. Stinebrickner, T. (2008), "The Causal Effect of Studying on Academic Performance," *The B.E. Journal of Economic Analysis & Policy*: Vol. 8: Iss. 1 (Frontiers).