DOCUMENTO DE TRABAJO N° 287

THE IMPACT OF STUDENT LOANS ON EDUCATIONAL ATTAINMENT: THE CASE OF A PROGRAM AT THE PONTIFICAL CATHOLIC UNIVERSITY OF PERU

Luis García Núñez

DEPARTAMENTO DE ECONOMÍA

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

DOCUMENTO DE ECONOMÍA N° 287

THE IMPACT OF STUDENT LOANS ON EDUCATIONAL ATTAINMENT: THE CASE OF A PROGRAM AT THE PONTIFICAL CATHOLIC UNIVERSITY OF PERU

Luis García Núñez

Agosto, 2010

DEPARTAMENTO
DE **ECONOMÍA**

PONTIFICIA
UNIVERSIDAD
CATÓLICA
DEL PERÚ

DOCUMENTO DE TRABAJO 287
http://www.pucp.edu.pe/departamento/economia/images/documentos/DDD287.pdf

Av. Universitaria 1801, Lima 32 – Perú.
Teléfono: (51-1) 626-2000 anexos 4950 - 4951
Fax: (51-1) 626-2874
econo@pucp.edu.pe
www.pucp.edu.pe/departamento/economia/

Encargada de la Serie: Giovanna Aguilar Andía
Departamento de Economía – Pontificia Universidad Católica del Perú,
gaguila@pucp.edu.pe

Luis García Núñez

THE IMPACT OF STUDENT LOANS ON EDUCATIONAL ATTAINMENT: THE CASE OF A PROGRAM AT THE PONTIFICAL CATHOLIC UNIVERSITY OF PERU / Luis García Núñez
Lima, Departamento de Economía, 2010
(Documento de Trabajo 287)

Student Loans / Matching / Treatment Effect

Hecho el Depósito Legal en la Biblioteca Nacional del Perú Nº 2010-06580
ISSN 2079-8466 (Impresa)
ISSN 2079-8474 (En línea)

Impreso en Cartolan Editora y Comercializadora E.I.R.L.
Pasaje Atlántida 113, Lima 1, Perú.
Tiraje: 100 ejemplares

THE IMPACT OF STUDENT LOANS ON EDUCATIONAL ATTAINMENT: THE CASE OF A PROGRAM AT THE PONTIFICAL CATHOLIC UNIVERSITY OF PERU

Luis Garcia Núñez

Resumen

En las últimas décadas, la Pontificia Universidad Católica del Perú (PUCP) ha otorgado préstamos estudiantiles a algunos estudiantes con rendimiento académico satisfactorio pero que enfrentan problemas económicos que pudieran poner en peligro su continuidad como estudiantes. Aunque el programa fue creado hace más de cuarenta años, sus resultados nunca han sido evaluados en forma rigurosa. Este documento intenta evaluar hasta qué punto el programa ha beneficiado a los estudiantes. Debido a que la información disponible proviene de records académicos, el cumplimiento de este objetivo requiere el uso de técnicas modernas diseñadas para trabajar con datos no experimentales. Estimando por el método de propensity score matching, encuentro un impacto estadísticamente significativo sobre el número de semestres necesarios para culminar los estudios solo cuando el estudiante recibió el préstamo por 6 semestres o más. También se encontró un impacto significativo sobre la probabilidad de concluir los estudios seis años y medio después de haberlos iniciado, para el caso de estudiantes que recibieron el préstamo por 6 semestres o más. Sin embargo, tal efecto es pequeño.

Abstract

During the past decades, the Pontifical Catholic University of Peru (known as PUCP) has been giving student loans to some of its students with satisfactory academic performance but who face certain economic problems which might interrupt their studies. Although this program was created more than forty years ago, its results have not been rigorously evaluated. This document attempts to assess to what extent the program has benefited students. Because the collected data come from academic and social records, the completion of this task requires using modern techniques specifically designed to work with non experimental data. After estimating by propensity score matching with multiple treatments, I find a statistically significant impact of this program on the time a student employs to complete the course of study at PUCP (measured in semesters) only when a student was awarded with a loan for 6 semesters or more. That effect is not significantly different from zero when the loan lasts less than 6 semesters. Similar results were found when I analyzed the impact on the probability of degree completion of student loans, where students with loan were more likely to meet all graduation requirements by 6 years and a half after they start studying at PUCP. Again this effect was significant only when the student participates in the program for six semesters or more. However, the impact on that probability was small.

JEL Classification Codes: C13, C14, C21, I22
Key Words: Student Loans, Matching, Treatment Effect

THE IMPACT OF STUDENT LOANS ON EDUCATIONAL ATTAINMENT: THE CASE OF A PROGRAM AT THE PONTIFICAL CATHOLIC UNIVERSITY OF PERU

Luis García Núñez

1. INTRODUCTION

During the past decades, the Pontifical Catholic University of Peru (*Pontificia Universidad Católica del Perú*, known as PUCP for its initials in Spanish) has been giving student loans to some of its students with satisfactory academic performance but who face certain economic problems which might interrupt their studies.

Although this program was created more than forty years ago, its results have not been rigorously evaluated. This document attempts to assess to what extent the program has benefited students. Because the collected data come from academic and social records, the completion of this task requires using modern techniques specifically designed to work with non experimental data.

To my knowledge, this paper is the first attempt to evaluate a student loan program in Peru. Since student loans are not common in Peruvian society, discovering the effectiveness of such a program could give us insight on what would happen if student loans were offered in this country, and what would be necessary to make them effective.

The paper is organized as follows. Section 2 discusses the importance of student loans and the existence of that kind of programs for higher education in Peru. Section 3 presents the identification strategy I employ to estimate the impact. Section 4 describes the data and variables. Section 5 presents the main results of this research. Finally, section 6 concludes and discusses further topics to be analyzed.

2. STUDENTS LOANS AND THE PROGRAM IN PERU

Since the 20th century, economic and social research has recognized that education is an essential path to overcome social differences and poverty (Sen, 2000). Several studies on the returns of education have shown that individuals who have the opportunity to receive an education are able to develop themselves in modern societies as professionals and earn enough income to break the "poverty trap" in which their families have been caught for generations.

The question that governments and educational institutions ask themselves is how to improve access to education for those individuals who otherwise do not have the opportunity to study. In this sense, student loans become a powerful instrument to help those individuals to overcome this difficulty. The main objective of student loans as a social program in less developed countries is to equalize the educational opportunities, giving them the opportunity to earn a degree[1]. Even though social differences in less developed countries may be strong, student loans may help to reduce these differences.

Student loans —in contrast to scholarships— have become a popular financial aid for students in many countries[2] due to their multiple advantages: they relate future earnings to present needs, and they are not restricted to a fixed fund since they are self-financed with the amounts repaid. However, student loan programs also face problems that could make them difficult to implement without essential participation in the public sector. In countries where credit markets are not developed, there still exist students who are not eligible for new loans because of the absence of financial products attractive enough for borrowers and

---

1 In particular, students who come from low-income families or those who come from an environment in which it has been difficult to develop professional skills.

2 See for example Kim (2007) in the United States, Callender and Kemp (2000) for the UK experience, Canton and Blom (2004) for the Mexican case, and Larrain and Zurita (2008) for Chilean programs.

lenders. Another important problem is the absence of a credit culture among borrowers.

Unfortunately, in Peru we observe a very low rate of loans for students in higher education (less than 3%), one of the lowest rates of in the region (IADB-WB-APICE, 2009). The majority of these loans have been granted by the public National Institute of Scholarships and Student Loans (known as INABEC for its initials in Spanish, and recently renamed OBEC), and the rest have been given by non-profit organizations or educative institutions such as universities and technical schools. It is important to note that the private financial institutions (such as banks) have a minimum of participation in these kinds of credit loans. Despite the great expansion of other kinds of credit in our economy in recent years, we do not observe a significant increase in student loans.[3]

In this context of a significant lack of credit for students in Peru, this work focuses on the experience of the Pontifical Catholic University of Peru (PUCP) as a student loan granter and attempts to evaluate is effectiveness through an impact study. Although its scope was limited to PUCP students, we can learn some lessons from this experience in Peru.

In 1967, the PUCP created this new student loan program to assist students who had financial problems that could threaten their continuation as regular students. Two conditions were stipulated to apply for a loan: (a) the family income was not high enough to pay academic fees; and (b) a generally satisfactory academic performance. To evaluate these two conditions, a permanent commission composed of social workers and academic representatives was established. Further regulations specified that condition (a) meant a family income below four

[3] It is not the objective of this work to answer why Peru has such low participation rates. Nevertheless, the problems that may cause this result are well known: the absence of real guarantees and collateral (especially for low income students) and the fact that human capital cannot be taken as collateral; a still insignificant credit culture within Peruvian society; the absence of a national policy to back up the loans with public funds, etc.

times the legal minimum wage, and that condition (b) meant a weighted average grade above 60% of the maximum possible score.[4] A student granted for the loan did not pay tuition and fees for one year (two academic semesters), and in special cases could receive extra funds for lunch and books. These extra funds were not part of the loan. After the student's date of graduation, a grace period of six months was established. After that, students were required to contact the authorities at PUCP and to provide a repayment plan. These rules were maintained for about 30 years with minor changes, but at the beginning of the new century the program was revised with stricter conditions. These new conditions are not presented here because they were applied outside the period of analysis of this paper.

It is important to mention that during the first three decades of the program the repayment rate was very low, which worried the university authorities. The absence of punishment to individuals defaulting on their loans, and the inability of the PUCP to take legal action and retrieve the money, was a serious problem for the continuation of the program. It also caused the program to be perceived among many students as a donation rather than a loan.[5]

For many reasons, the thresholds mentioned above were never strict. It was very common to observe students receiving a loan without fulfilling these conditions. Some of the reasons why this happened were: different schools inside the PUCP had different standards to grade their students (e.g. absolute average grades for engineering and economics were particularly low compared to administration and accountability); and the

[4] The academic grading in Peru employs a scale from 0 to 20 points, where the passing threshold is 11 points. These scores are absolute, not relative to the class. In higher education, a student who gets a score in the range 0-10 in one course must repeat that course. A student who repeats the same course three times is expelled from school. The weighted average score is calculated by semester using the number of credits of each course as weights. The cumulative score is the weighted average score of all courses taken (whether passed or not) since the first semester to the present.

[5] The author, as a former student at PUCP in the late 80's and early 90's and former beneficiary of this program, was a witness to this assertion.

committee took into account other academic factors beyond the cumulative average scores, such as reports on applicant motivation and perception from his teachers. The committee also considered the background of the applicant, with a thorough study of his family's socioeconomic situation. It was also understood by the committee that other personal and qualitative factors (the environment in which the student grew up, his relationship with his parents, the development of skills during childhood, and socialization problems) could affect his performance as a student. Another important matter that characterized this program was that students applied voluntarily for the loans, and the committee decided who would be the beneficiary. Social workers who were former members of this committee state that loans were granted to almost all applicants during those years.[6]

In order to define the amounts the students must pay for tuition and fees, the PUCP has a tiered-fee payment structure which takes into account the income distribution in Peruvian society and seeks to ensure that all students can receive education regardless of financial circumstances. Currently there are five levels, with Level 1 corresponding to the low-income students and Level 5 for those students with higher income. When students are admitted to PUCP, a thorough socioeconomic study is conducted to determine into which scale each student is classified. This study collects socioeconomic information through a socioeconomic application form and house visits to verify the information provided. It also includes interviews with parents and other relatives.

After students were classified into one of the levels, they could be reclassified at any time if any change occurred that worsened their financial situation. They could then apply for the loan in that case. In either case, a new socioeconomic study was conducted. In addition, students granted with a loan were permanently monitored in order to

6 All this information was provided to the author during several interviews of current and former authorities at PUCP.

detect any positive change in their financial status. If this occurred (or if their academic performance was not as good as expected —a controversial point as discussed above), they could be expelled from the program. These rules meant that some students benefited over a number of semesters; in some cases students received loans for their whole stay at PUCP, and other students benefited for one semester only.

To have some figures about the program in the 90's, Table 1 shows the stock of students from the 1998-2 semester (fall) to 2002-1 (spring), where the numbers have been broken down into fee levels and student borrower status.

Table 1
Number of Students at PUCP by loan status and tiered fee system

	With Loan		Without Loan						Total
Level	1	2	1	2	3	4	5	6*	
Semester									
1998-2	691	0	3,649	2,556	2,223	1,686	988	1,304	13,097
1999-1	666	0	3,937	2,744	2,332	1,717	970	1,267	13,633
1999-2	729	7	4,086	2,934	2,434	1,671	871	1,194	13,926
2000-1	733	0	4,284	3,241	2,508	1,757	967	1,210	14,700
2000-2	724	2	4,436	3,225	2,531	1,676	896	1,096	14,586
2001-1	665	1	4,298	3,353	2,706	1,813	898	1,104	14,838
2001-2	626	0	4,529	3,475	2,614	1,712	774	1,011	14,741
2002-1	504	1	4,744	3,724	2,791	1,793	1,758*	--	15,315
2002-2	434	0	4,871	3,866	2,747	1,676	1,507	--	15,101

* Since 2002, levels 5 and 6 where unified.
Source: PUCP date bases

As we can see in this table, less than 5% of students apply and are elected for a loan out of the total student population. As expected, the majority of students who applied for and benefited from this program belonged to Stratum 1 (the low-income group), while only a few cases belonged to Stratum 2. Therefore, the participation rate increases when we consider only the group that seems to be the "target population", students at Stratum 1. However, simple calculations show that these rates declined over time.

The expected impact of the program was, as mentioned, to give the opportunity to more students to complete their higher education at PUCP, one of the best universities in Peru. To translate these intentions into quantitative outcomes, I defined two indicators to be evaluated. One is the number of semesters the student needs to complete the course of study. A student with financial problems may choose either to apply for a loan or to work and study. This latter option will affect the time the student dedicates to studies, which may prolong the time required to pass all courses in the program. Besides, since students can choose the number of courses they take each semester, fewer courses could be desirable for working students and it would also mean lower fees (which they could afford). On the other hand, a student who benefits from a loan can take as many courses as he wants (depending only on his time and capabilities), and the program forbids the student to work.

The other indicator to be evaluated is the probability of completing the course of study and getting a bachelor's degree at PUCP.[7] If a student faces financial problems, they could be serious enough to compel the student to quit studying at PUCP (perhaps moving to a cheaper and less time-demanding school, or to the labor market). In contrast, students in the program have more chances to succeed in their studies.

The data available for this study comes from students' records, and may suffer of some data problems. Therefore, I need to define an empirical strategy that will produce reliable statistical results. The next section focus on this empirical strategy.

[7] According to Peruvian laws, the Bachelor's degree is attained automatically once the student completes all required courses and credits that the school defines in the course of study.

3. IDENTIFICATION STRATEGY

One possibility to evaluate the impact of student loans is to model structural equations on the determinants of the endogenous variables and to use regression methods to estimate this impact. However, these methods have been criticized as evaluation method in observational studies due to their inefficacy to overcome the "confounder problem", to suffer of sample selection problems, and because they frequently suffer of endogeneity in the treatment variable. Some studies have shown (see for example, Lalonde, 1986) that OLS results are far from those we could obtain in randomized experiments. Besides, they rely on some strong assumptions on the distribution of error to make statistical inference, being those errors unobservable for the researcher (Angrist, Imbens y Rubin, 1996). In addition, regression analysis depicts associative relationships and any causal inference is based on opinions and beliefs rather than on objective evidence. For these reasons, it will not be appropriate to make causal inference based on associative parameters.

To have a statistically valid measurement of the impact of student loans, I employ the standard methodology which relies on the works of Rubin (1974) and Holland (1985), known as the Neyman-Rubin Model or Roy-Rubin Model. In this framework, the observed outcomes of the program are compared with the potential outcomes in a hypothetical scenario where individuals subject to treatment do not receive it. The well-known fundamental problem of causal inference applies in this case because only one of those two potential outcomes is observable. Dealing with this difficulty, the proposed "statistical solution" requires reconstructing the counterfactual scenario using data from those untreated individuals, which in this study refers to students who did not participate in the student loan program.

However, concerns on the comparability of these two groups arise since we are using observational data, in which differences in outcomes from treated and untreated individuals may be caused not only due to differences in participation in the program (students receiving a loan), but also due to other observable and unobservable characteristics of the students. If we had experimental data (in whose case the treated population would have been chosen randomly), those characteristics would be balanced across those groups and we did not worry on the simple mean comparison of treated and untreated outcomes. Unfortunately, an experiment is not ethically feasible in the case of student loans due to the serious long run consequences that may occur if a student who cannot afford to continue his studies did not receive funds like a loan[8].

Inasmuch as our data available comes from student records, our identification strategy needs to control for those observable covariates which could confound the effect of student loans on the proposed indicators. This also requires certain untestable assumptions on the data generation process which are discussed here.

Before formalizing the identification strategy, it is important to remark that in this program, it is not relevant to describe the program participation by a binary variable because the different intensity of the treatment (measured by the number of semesters the student participated in the student loan program) varies across beneficiary students. As a consequence a multiple treatment analysis is used in this work (Imbens 2000, Lechner, 2001).

Following Lechner (2001) notation, let y_i be the outcome of the program and let $S \in \{0,1,\ldots,M\}$ be the program participation variable who also measures the intensity of the treatment. Let y_i^j be the potential outcome for individual i if he receives treatment $S = j$. After the student graduates

[8] Do not be benefited by the program is harmful for the student because the loose of opportunities. It is not like the placebo used in medicine.

(or quit studying), only one of those potential outcomes is observable for the researcher. Also, let x_i be the set of observed covariates and ε_i the set of unobserved covariates for student i.

I am interested in pair wise comparison of outcomes, where the outcome for $S=0$ is compared with any other result for $S>0$. The average treatment effect on the treated for individuals who received treatment $S=m$ is defined as (omitting subscript i)

$$\theta^{m,0} = E[y^m - y^0 \mid S=m] = E[y^m \mid S=m] - E[y^0 \mid S=m] \tag{1}$$

Clearly the second term in the right-hand side of equation (1) is not observed.

Some assumptions are needed to identify the effect. The conditional independence assumption of potential outcomes states that $y^0, y^1 \ldots \amalg S \mid x$. To our task we do not need this strong assumption, it is only necessary to assume that

$$y^0 \amalg S \mid x, S \in \{0, m\} \tag{2}$$

to identify the effect in equation (1). Under this assumption, $E[y^0 \mid x, S=m] = E[y^0 \mid x, S=0]$. Besides, if we assume that

$$0 < P^j(x) < 1 \quad \forall j = 0, S \tag{3}$$

which is the usual common support assumption where $P^j(x)$ is the probability to participate in $S=j$ for an individual with characteristics x, then it is true that

$$E[y^0 \mid S=m] = E_x[E[y^0 \mid x, S=0] \mid S=m]. \tag{4}$$

The right hand side expression can be estimated by sample analogs.

The well-known problem of dimensionality[9] can be solved by using propensity scores. According to Lechner (2001) who extends Rosenbaum and Rubin (1983) result, if (2) and (3) hold, then $y^0 \coprod S \mid P^{m|m0}(x), S \in \{0, m\}$, where $P^{m|m0}(x)$ is the probability of receiving treatment m on the subpopulation $S \in \{0, m\}$ given observed pre-treatment variables x. Then using propensity scores we can identify $\theta^{m,0}$ and the multiple treatment models have been compressed to several binary treatment models.

Under this setting, we could apply standard matching methods (Rubin, 1973; Heckman, Ichimura and Todd, 1997, 1998; Dehejia and Wahba, 2002; Caliendo and Kopeining, 2005) to estimate the impact. The first step is to define an empirical counterpart of equations (1) and (4). Let T be the set of all treated individuals and let N the set of untreated students; let's define the comparison group for treated student i as $A_i = \left\{k \in N \mid \hat{P}^{m|m0}(x_k) \in v(\hat{P}^{m|m0}(x_i))\right\}$ where $\hat{P}^{m|m0}(x)$ are propensity scores which are estimated by Probit and $v(P^{m|m0}(x_i))$ defines a neighborhood around $P^{m|m0}(x_i)$. This set contains one unit in the case of *nearest neighbor matching* in whose case $A_i = \left\{k \in N \mid \min\left\|\hat{P}_i^{m|m0}(x) - \hat{P}_k^{m|m0}(x)\right\|\right\}$, but it could be empty if we define a minimum distance. In the case of *radius matching*, all individual in a range around $P^{m|m0}(x_i)$ are included, so $A_i = \left\{k \in N \mid \left\|\hat{P}_i^{m|m0}(x) - \hat{P}_k^{m|m0}(x)\right\| < \delta\right\}$. There is another estimator, the *kernel matching estimator* in which the comparison groups contains all the individuals in the untreated group, $A_i = N$. When more than one individual could be used to reconstruct the counterfactual scenario -A_i contains more than one element-, we need to weight the observations with a weighting function $\omega(i,k)$, $0 \le \omega(i,k) \le 1$, and $\sum_{k \in A_i} \omega(i,k) = 1$. For kernel matching, $\omega(i,k)$ is

[9] The problem of dimensionality refers to the fact that if vector X has many variables, it would be almost impossible to match each individual of the treatment group with another individual of the untreated group.

$$\omega(i,k) = \frac{K(\frac{\hat{P}_i - \hat{P}_k}{h})}{\sum_{k \in N} K(\frac{\hat{P}_i - \hat{P}_k}{h})}$$

where $K(\cdot)$ is a kernel and h is the bandwidth.

The standard propensity score matching estimator is

$$\hat{\theta}^{m,0} = \frac{1}{n_T} \sum_{i \in T \subset CS} \left(y_i^m - \sum_{k \in A_i \subset CS} \omega(i,k) y_k^0 \right)$$

where n_T is the number of students in *T*, and *CS* refers to the "common support".

4. THE DATA AND VARIABLES

As usual in this kind of study, the empirical strategy and the data available interact, with one being the result of the other. In the previous section, I presented the methodology for propensity score matching, which requires having two comparison groups: the treated and the untreated students. However, the treated group is disaggregated to have different levels of treatment, generating different treated groups to be compared with the untreated group.

In seeking these groups, the first step was to define the target population, according to the original rules of the program. As Table 1 suggests, I focused on students who belong to the socioeconomic stratum (or Level 1), and discarded the other levels from 2 to 6. Besides, since the program also had a merit-based part related to the weighted average score, I focused on students who attained a grade point average of at least 12.[10] Even though some students who received a loan had a GPA below this level, I prefer to use this threshold to define a "potentially

[10] Remember, in the Peruvian grading system, all scores are numbered in the range of 0-20.

eligible student". These two conditions (to belong to Level 1 and to have GPA >=12) were applied to both treated and untreated students, as a first step to ensure comparability between them.

The second criterion to select the data is related to the long-term nature of the student loan program. To assess the impact, we need to give the students enough time to complete their course of study or take the decision of not studying anymore at PUCP. The data covered up to mid-2009, so as a consequence I ruled out all students who were admitted from mid 2003 to later, given that Peruvian laws require that the curricula must take at least five years. Another restriction of the data was that it was very difficult to find good data for years before 1997, since most of the students' information was recorded manually on paper forms, and the archives were periodically destroyed due to maintenance.

For these reasons I was forced to work with students who were admitted to PUCP between the second semester of 1997 and the first semester of 2003. Finally, I randomly selected 1219 students who belong to the subpopulation described above. From them, 350 students received the loan, and the remaining 869 did not receive it (see Table 2).

Due to the small number of observations, only three distinct levels of treatment were considered as suggested by Table 2:

S = 0: The student did not receive treatment;
S = 1: if the student received treatment for 1 to 5 semesters; and
S = 2: if he or she received the loan for 6 or more semesters.

Table 2

Number and Percentage of students in sample, by participation in the program

Type of student	Frequencies	Percentage
Untreated	869	71.29%
Treated	350	28.71%
Received treatment for 1 to 5 semesters	163	13.37%
Received treatment 6 or more semesters	187	15.34%
Total	1,219	100%

Concerning covariates, I obtained abundant information from the Household Socioeconomic Form (*Ficha socioeconómica*) that all students must fill out at the time they are admitted to PUCP. For all untreated students, the data available was for the first semester they studied at PUCP. For all treated students, the data was for the semester they applied for the loan, in which case they needed to update the information provided earlier. As a result, the data from this source was recorded before the treatment. It is an important requirement for good matching to control by pre-treatment covariates, and this data fulfills this requisite.

The household socioeconomic form provided data on individual, family and dwelling characteristics, as well as family income. It is important to mention that we got information on declared and imputed family income. The first item was the income the student's parents declared, and the second was the income which was calculated after a social worker visited the student's home and checked the information provided in the socioeconomic form. Social workers checked bills, payslips and other sources of income. After reviewing the information, a report was written detailing all corrections that were made to declared income.

This data was appended to the academic record of each student. These records included average grade point by semester, the school in which the student was enrolled, the semester each student received a loan, date of graduation, etc.

The following table classifies some of the available variables into 4 categories. There was more information in the household socioeconomic form, but I mainly used these covariates to avoid loss of data due to missing values. The first two are self-explanatory, but the remaining three categories need an explanation. The Cultural Capital category groups variables related to accumulation of education, experience, and moral values of the student. In Peru, the type of high school (private or public) provides information on this capital, due to big disparities in the level of education of each type (BIRF, 2006). Cultural differences by region also make it necessary to include the place of birth as a cultural capital variable. Social Capital is related to the "social environment" to which the student belongs, and it is represented in the list of variables by the income of the district in which the student lives (according to the human development index) and with whom the student resides. The Economic Capital group refers to all the information related to economic assets.[11]

Table 3
Classification of covariates

Groups	Variables
Students' characteristics	Age (at date of admission), Sex, Program or undergraduate School.
Other household characteristics	Number of students' siblings living at home, Number of household members who suffer of chronic diseases[1]
Social and Cultural Capital	Parents' level of education, Type of High School, Place of birth, quintile of district income (according to the human development index), with whom the student resides?
Economic Capital	Income quintile, Other working relatives at home, Housing tenure, Number of floors in dwelling, Access to sewer system.

[1] Includes reports of household members with mental problems, disabilities, etc. Normally only serious diseases were reported in the household socioeconomic form.

Table 4 presents descriptive statistics of those quantitative variables in our sample. The groups "Cultural and Social Capital" and "Economic Capital" were reduced to only one dimension using multiple

[11] I took the idea of grouping the variables in those definitions of capital from CID (2006).

correspondence analyses. Positive values of these indexes are associated were higher needs. The respective indexes are shown in this table.

The same table also shows that there are important differences between the treated and the control group. The standard mean test for equal means shows that treated students come from a subpopulation which has higher needs in comparison with untreated students, because the difference between the Cultural and Social Capital Index and the Economic Capital Index is significantly different from zero. There are also significant differences in the number of offspring living in the household (with greater numbers of offspring in households of treated students), and differences in the number of household members who suffer from chronic diseases (more frequent in treated households). These differences could be in response to either a self-selection process or the by PUCP administrators' criteria to allocate the loan.

Table 4
Descriptive Statistics for Treated and Untreated Students

	Average values			
	Treated	Untreated	Mean test	P-value
Number of semesters	13.08286 (350)	13.14302 (867)	0.3380	0.7354
Did the student complete the curricula by 2009? (1=yes, 0=No)	0.7514286 (350)	0.7139562 (867)	-1.3246	0.1855
1= Male, 0 = Female	0.5057143 (350)	0.5294118 (867)	0.7487	0.4542
Students Age (at the time of admission)	18.06571 (350)	18.08314 (866)	0.1197	0.9048
Economic Capital Index	0.0223341 (339)	-0.0090457 (837)	-4.9184	0.0000
Social and Cultural Capital Index	0.0125998 (310)	-0.0049568 (788)	-2.0517	0.0404
Offspring of parents living in household	2.514451 (346)	2.283879 (856)	-3.1739	0.0015
Number of household members who suffer of chronic diseases	0.8563218 (348)	0.6524249 (866)	-3.6187	0.0003

Note.- Actual number of observations in parenthesis. This number varies due to missing values.
Source: PUCP data bases.

Table 5 presents the comparison of treated versus non-treated students in a sample for qualitative variables. I found significant differences in the quintile of income in the sample. This result shows that, although I limited the student sample to individuals who belong to Level 1, there are income differences within this level, and the treated students seem to be "poorer" than the untreated ones. There are also differences in the "number of floors" of dwelling, with larger dwellings among the untreated. Also, the education of parents is lower for treated individuals.

There are two more important differences in the observed covariates. One is the type of high school in which the individual studied before entering PUCP. The table says that an important percentage of treated students completed high school in a public school, whereas a large percentage of untreated students come from religious and private schools. As I mentioned before, this difference is important, because in Peru public schools are cheaper and quality of education is lower. Finally, many treated students were enrolled (at the time of evaluation) in the School of Sciences and Engineering (44%), while for untreated students that percentage was lower (32%).

Table 5
Comparison of qualitative covariates
(column percentages)

Variable	Categories	Treated	Untreated	Pearson Chi2 Test	P-value
Quintile of income in sample	1 (poorer)	31.61%	15.23	79.09	0.000
	2	26.15%	17.67		
	3	19.54%	20.12		
	4	12.36%	23.37		
	5 (richer)	10.34%	23.6		
Other relatives in household (nuclear family not included)	None	57.31	53.23	6.02	0.110
	1 or 2	34.38	40.99		
	3 or 4	6.02	4.27		
	5 or more	2.29	1.5		
Housing Tenure	Owner-occupancy	48.25	56.09	8.18	0.042
	Tenancy	15.5	13.61		
	Paying by installments	0.29	0.95		
	Others	35.96	29.35		
Number of floors	1	78.86	74.45	5.65	0.059
	2	18.29	19.56		
	3	2.86	5.98		
Access to sewer system	No	0.87	0.12	4.18	0.041
	Yes	99.13	99.88		
Student's father education level	No education	0.00	0.00	17.33	0.015
	Incomplete primary school	0.35	0.29		
	Complete primary school	5.57	3.3		
	Incomplete high school	4.88	3.3		
	Complete high school	31.71	31.66		
	Incomplete technical education	0.00	0.14		
	Complete technical education	9.76	10.32		
	Incomplete higher education	12.54	6.59		
	Complete higher education	35.19	44.41		
Student's mother education level	No education	0.31	0.24	11.23	0.189
	Incomplete primary school	0.92	0.84		
	Complete primary school	8.92	6.25		
	Incomplete high school	4.62	2.52		
	Complete high school	40.92	40.14		
	Incomplete technical education	0.31	0.00		
	Complete technical education	7.38	9.86		
	Incomplete higher education	5.54	4.93		
	Complete higher education	31.08	35.22		
Type of high school before PUCP	Public school	42.00	30.87	14.30	0.014
	Non-religious private school	28.57	34.22		
	Private religious school	28.57	33.53		
	"Fe y Alegria" Program	0.29	0.58		
	Armed Forces	0.29	0.58		

	Parochial	0.29	0.23		
Place of birth	Provinces	18.29	19.79	0.36	0.547
	Lima	81.71	80.21		
With whom do you live?	Parents	77.74	76.59	0.71	0.982
	Other relatives	16.32	16.81		
	Friends	0.59	0.36		
	Boarding house	2.97	3.48		
	Others	0.89	1.08		
	Parents and relatives	1.48	1.68		
Quintile of district income	Poorest	4.05	2.71	6.20	0.185
	Poorer	17.63	14.02		
	Average wealth	28.9	26.86		
	Richer	33.53	37.69		
	Richest	15.9	18.73		
Undergraduate school	Administration and accountability	9.43	14.42	33.66	0.000
	Architecture and urbanism	0.86	1.15		
	Art	2.29	1.73		
	Sciences and engineering	44.57	32.3		
	Social sciences	7.43	4.38		
	Communications arts and sciences	11.14	11.19		
	Law	14.86	19.38		
	Education	3.71	4.73		
	General studies: sciences	2.00	1.61		
	Administration and executive studies	0.29	1.15		
	Liberal arts and humanities	3.43	7.96		

Source: PUCP data bases.

5. ESTIMATION AND RESULTS

In this section, I present the main results of this research. The impact was calculated by two methods: regression analysis and standard propensity score matching.

5.1 Impact on the number of semesters

The first method to measure the impact of student loans is a parametric regression, in which a linear regression model is estimated by ordinary least squares. In this regression, the dependent variable is the number of semesters a student took to attain a degree, and I explore the significance of a dummy variable of participating in the program. As suggested in the empirical strategy section, I evaluate the impact defining subsamples with respect to the intensity of the treatment and I regress a linear model to each subsample. Therefore, the first column of estimates in Table 6 works with the subsample of students who received treatment for fewer than 6 semesters plus those untreated students. The last column presents the estimates for the same regression when the subsample contains all students who received the loan for 6 or more semesters plus the untreated students.

In the calculation of the impact on the number of semesters, I use only the sample of students who graduated by the end of 2009. In my opinion, it does not make sense to analyze this kind of impact for students who quit studying and left their programs. Table 6 shows that the impact of student loans is greater as long as the student receives the loan for more semesters. According to the OLS estimates, the impact is not significantly different from zero when a student receives treatment for fewer than 6 semesters. Going over the impact of controls, only students' sex and age seem to have a significant impact. As I mention in section 3, we cannot trust on these OLS results, but they are presented here only as reference values.

Table 6
OLS estimates (restricted to students who graduated by the end 2009)

Dep. Var.: Time a student needed to complete the course of study	Subsamples	
Variables	Treated 1 to 5 semesters vs. untreated	Treated 6 or more semesters vs. untreated
Received Loan=1, Not received = 0	-0.00536	-0.509***
	(0.219)	(0.183)
Sex (1= Male, 0 = Female)	0.276*	0.292*
	(0.161)	(0.151)
Students Age (at the time of admission)	0.172***	0.161***
	(0.0390)	(0.0363)
Economic Capital Index	1.207	0.589
	(0.782)	(0.742)
Social and Cultural Capital Index	0.664	0.308
	(0.607)	(0.565)
Offspring of parents living in household	0.0481	0.0741
	(0.0726)	(0.0692)
Number of household members who suffer of chronic diseases	0.153	0.0647
	(0.0935)	(0.0863)
(dummies for undergraduate schools not shown)		
Constant	9.932***	10.13***
	(0.712)	(0.660)
Observations	657	698
R-squared	0.175	0.181

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Now let's present the result of the non-parametric matching. In Table 7 we see the estimates of the average treatment effect on the treated, with two different levels of treatment, $\theta^{1,0}$ and $\theta^{2,0}$. In the first case (top part of the table), under the three methods (nearest neighbor, radius and kernel) I do not find any significant difference in the number of semester between treated and control groups when the treatment was applied for less than 6 semesters. On the other hand, in the bottom part of the table we see that the impact of the program is significantly different from zero (except when it is calculated with the nearest neighbor method.) The impact is around 0.5 semesters, meaning that students who receive the loan complete their course of study earlier than those students without loan. This result is similar to that found in a similar work in Colombia (see CID, 2006).

Even though this effect may look "small", we should note that the minimum number of semesters to complete the curricula is 10, and in some schools it could take 11 semesters to graduate[12]. If we discount this minimum number of semesters from the calculated averages we see that the amount of the impact is not negligible.

As a final comment, note that the estimates of the average treatment effect on the treated are very close to the program coefficient in the OLS regression.

Table 7
Impact of the program on the number of semesters

	Nearest neighbor	Radius	Kernel
Subsample S = 0 or S = 1			
Treated	13.563	13.563	13.563
Control	13.500	13.595	13.625
$\hat{\theta}^{1,0}$	0.063	-0.032	-0.062
T-stat	0.195	-0.131	-0.241*
Subsample S = 0 or S = 2			
Treated	13.161	13.161	13.161
Control	13.599	13.625	13.634
$\hat{\theta}^{2,0}$	-0.438	-0.465	-0.473
Tstat	-1.611	-2.602	-2.479

* Based on bootstrapped standard errors

5.2 Effect on the probability of degree completion

In this section I explore the effect of the program on the probability of completing the studies at PUCP. As it is known, students with economic difficulties may stop studying for some semesters and may never meet

[12] In the School of Law, the course of study requires at least 12 semesters.

graduation requirements. I calculate the effect by Probit regression analysis and by propensity score matching. Unlike estimation in section 5.1, this time we need all the observations, graduated and non graduated students.

Table 8
Probit Marginal Effects

Dep. Var: 1=Student graduated from college, 0 = Student did not graduated	Subsamples	
Variables	Treated 1 to 5 semesters vs. untreated	Treated 6 or more semesters vs. untreated
Received Loan=1, Not received = 0	-0.0282	0.0794**
	(0.0442)	(0.0346)
Sex (1= Male, 0 = Female)	-0.138***	-0.136***
	(0.0319)	(0.0306)
Students Age (at the time of admission)	-0.0159**	-0.0104
	(0.00694)	(0.00679)
Economic Capital Index	0.122	0.0806
	(0.150)	(0.144)
Social and Cultural Capital Index	-0.311**	-0.271**
	(0.125)	(0.117)
Offspring of parents living in household	-0.00760	-0.0128
	(0.0130)	(0.0128)
Number of household members who suffer of chronic diseases	-0.00696	-0.0175
	(0.0169)	(0.0160)
(dummies for undergraduate schools not shown)		
Observations	889	921

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8 shows the results of the Probit marginal effects for the two subsamples. Again we observe a significant impact when the student received treatment for at least 6 semesters only. It is also interesting to say that results suggests that the probability of degree completion is higher for women than for men, and it is less likely for individual with lower social and cultural capital. As before, the regression results are shown as reference values, and I will not try to explore or correct any econometric problem in the estimation of this equation.

Now, let's see the results of the propensity score matching in Table 9, where the three methods are presented. I could say that these results are similar to those in table 7 because the impact is not significantly different from zero when the intensity of the treatment is "less than six semesters". In contrast, we find significant effects when the intensity is high (six or more semesters), except for the nearest neighbor method. The impact is not also very big: the probability of degree completion increases in about 8% for treated students.

An explanation for this apparent "small" effect relies on the selection of students who participate in the program. Recall that this is also a merit-based program, and in the selection of the sample, I also picked students which could be considered as "good". So, conditional on being a "good" student, the program increases the probability of attaining a bachelor degree in 8%. If the program were based on economic necessities only, it could include some "bad" students, which are not good just because they suffer serious problems that affect their performance as students. If this were the case, I sure the impact would be much larger than it currently is.

Table 9
Impact of the program on the probability of degree completion

	Nearest neighbor	Radius	Kernel
Subsample S = 0 or S = 1			
Treated	0.7007299	0.7007299	0.7007299
Control	0.7737226	0.7408786	0.7284339
$\hat{\theta}^{1,0}$	-0.073	-0.040	-0.028
T-stat	-1.282	-0.943	-0.575*
Subsample S = 0 or S = 2			
Treated	0.8353659	0.8353659	0.8353659
Control	0.7743902	0.7545154	0.7546442
$\hat{\theta}^{2,0}$	0.061	0.081	0.081
Tstat	1.222	2.409	2.183

* Based on bootstrapped standard errors

Again, as in the previous section, the impact is very similar to that in the Probit analysis.

6. CONCLUSIONS AND DISCUSSION

In this paper we have shown that controlling for observables characteristics, we can estimate the impact of student loans on the time a student need to complete his studies at PUCP and the probability of degree completion.

An important conclusion of this study is that in both cases, the impact is important only when a student receives the loan for 6 or more semesters. Students who received loans for less than 6 semesters do not show on average different outcomes than a control group with similar observable characteristics.

Nevertheless, this methodology does not control for unobserved covariates which may invalid this procedure. In addition, due to the nature of this program, which is a long run program, it is reasonable to think that some unobservable covariates may have changed along the time.

Further analysis should explore the accomplishment of assumption (2). It is difficult to test it, but the knowledge of the program rules suggests that there could exist “selection on unobservables” because the participation in the program depends not only on financial needs but also on (a) student preferences for “working and studying” and “only studying”, (b) student motivation on studies as a personal effort, (c) other psychological aspects of students personalities. This work cannot control these effects because we do not have the data to correct this possible source of bias.

REFERENCES

Angrist, Joshua, Guido W. Imbens y Donald B. Rubin. (1996). "Identification of causal effects using instrumental variables." *Journal of the American Statistical Association*, Vol. 91, N° 434 (Junio 1996), pp. 444-455.

BIRF – Banco Internacional de Reconstrucción y Fomento. (2006). *Por una educación de calidad para el Perú: estándares, rendición de cuentas y fortalecimiento de capacidades*. Lima: Banco Mundial.

CID - Centro de Investigaciones para el Desarrollo (2006). *Medición de impacto del crédito educativo Acces en la cobertura, la permanencia y el desempeño académico en la educación superior en Colombia*. Bogotá: Universidad Nacional - ICETEX - MEN.

Caliendo, Marco y Sabine Kopeinig. (2005). "Some Practical Guidance for the implementation of propensity score matching." IZA Discussion Paper N° 1588. The Institute for the Study of Labor.

Callender, Claire and Martin Kemp. (2000). *Changing Student Finances: Income, expenditure and the Take-up of Student Loans Among Full- and Part-time Higher Education Students in 1998/9*. Research Report N°213. Department for Education and Employment, UK.

Canton, Erik and Andreas Blom. (2004). "Can Student Loans Improve Accessibility to Higher Education and Student Performance?" World Bank Policy Research Working Paper N° 3425.

Dehejia, Rajeev H. y Sadek Wahba. (2002). "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *The Review of Economics and Statistics*, Vol. 84, No. 1 (Feb., 2002), pp. 151-161

Holland, Paul W. (1985). "Statistics and causal inference." *Journal of the American Statistical Association*. Vol.81, N° 396, Theory and Methods, pp. 945-960.

Heckman, James J., Hidehiko Ichimura y Petra E. Todd. (1997). "Matching as an econometric evaluation estimator: evidence from evaluating a job training programme." *The Review of Economic Studies*, Vol. 64, No. 4, Special Issue: Evaluation of Training and Other Social Programmes (Octubre), pp. 605-654.

Heckman, James J., Hidehiko Ichimura y Petra E. Todd. (1998). "Matching as an econometric evaluation estimator." *The Review of Economic Studies*, Vol. 65, No. 2 (Abril), pp. 261-294.

IADB-WB-APICE. (2009). *Crédito Educativo: experiencias internacionales y desafios futuros*. Lima: Universia.

Imbens, Guido. (2000). “The role of the propensity score in estimating dose-response functions.” *Biometrika*, Vol. 87, No. 3 (Sep., 2000), pp. 706-710.

Kim, Dongbin. (2007) “The effect of loans on students’ degree attainment: differences by student and institutional characteristics”. *Harvard Educational Review*, Spring, 77, 1, pp. 64-100.

Lalonde, Robert J. (1986). “Evaluating the econometric evaluations of training programs with experimental data.” *The American Economic Review*, Vol. 76, No. 4 (Sep), pp. 604-620.

Larraín, Christian and Salvador Zurita. (2008). “The new student loan system in Chile’s higher education.” *Higher Education*, Vol. 55, N°6, pp. 683-702.

Lechner, Michael. (2001). “Identification and estimation of causal effects of multiple treatment under the conditional independence assumption.” In Michael Lechner and Friedhelm Pfeiffer (Eds.). *Econometric Evaluation of Labour Markets Policies*. Heidelberg: Physica-Verlag.

Rubin, Donald B. (1973). “Matching to remove bias in observational studies.” *Biometrics*, 29: 159-183.

Rubin, Donald B. (1974). “Estimating causal effects of treatments in randomized and nonrandomized experiments”, *Journal of Educational Psychology* 66, pp.688-701.

Rosenbaum P.R. y Donald B. Rubin. (1983). “The central role of the propensity score in observational studies for causal effects.” *Biometrika*, Vol. 70 N°1, pp.41-55

Sen, Amartya. (2000). *Desarrollo y libertad*. Buenos Aires: Planeta, 2000.