Paper prepared for the EALE conference, Prague 2006

Is '3+2' Equal to 4? On the Effects of University Reform in Italy

Massimiliano Bratti *, Chiara Broccolini† and Stefano Staffolani‡

22nd February 2006

Abstract

In 2001 a '3+2' (unitary two-tier) university system was introduced in Italy where a 3-year First Level degree followed by a 2-year Second Level degree replaced a one-tier system where the 'old' degree (*Laurea*) length varied between a minimum of four (e.g. economics) and a maximum of six years (e.g. medicine). In this paper we use individual-level data on graduates from the Economics Faculty of the Marche Polytechnic University to investigate the effects of this reform. In particular, we seek an answer to questions such as: did the characteristics of students entering higher education change after the reform? Did the reform induce a change in the behaviours of students and higher education institutions (e.g. course workloads, grade inflation, etc.)? Did it produce a change in students' performances (e.g. student progression, grades)? Although our paper features a case study and evidence from the Marche Polytechnic University cannot be straightforwardly generalised to the whole Italian University system, our analysis is nonetheless informative given the general lack of evaluations of the '3+2' Italian University reform using micro-level data.

Keywords. Italy, Reform, University
JEL. I21

*DEAS, University of Milan, via Conservatorio 7, I-20122, Milano, Italy, and ChilD (Turin), IZA (Bonn), wTw (Milan). Email: massimiliano.bratti@unimi.it

†(Author for correspondence) Marche Polytechnic University, Piazzale Martelli 8, I-60121 Ancona, Italy. Email: broccolini@dea.unian.it

‡Marche Polytechnic University, Piazzale Martelli 8, I-60121 Ancona, Italy. Email: s.staffolani@univpm.it

1 Introduction

The Italian university system has been characterised for a long time by high drop-out rates and long graduation times compared to other OECD countries. In view of the progressive convergence towards an integrated labour market in the EU, the need to increase the competitiveness of Italian graduates and to harmonise the Italian system to the educational systems of other European countries has emerged in recent years. For these reasons, in 2001 a '3+2' (unitary two-tier) university system was introduced in Italy. A 3-year First Level degree followed by a 2-year Second Level degree replaced a one-tier system where the 'old' degree (*Laurea*) duration varied between a minimum of four (e.g. economics) and a maximum of six years.

About four years since the introduction of the reform, there have been a few attempts to analyse the current situation,[1] but, to the best of our knowledge, there has been no attempt to compare the situation before the reform with the one that emerged after the reform, that is to evaluate the effects of the 2001 reform. This lack of empirical analyses can be ascribed to various reasons. Firstly, individual-level administrative datasets on university students' academic careers are not easily available to researchers (due to privacy reasons) and rarely information is collected on students' family backgrounds. Secondly, an interest in monitoring the university system has developed in Italy only in recent years. In particular, Law n. 370/1999 introduced a system for the evaluation of Italian universities that is formed by one central institution, the *Comitato Nazionale per la Valutazione del Sistema Universitario* (Cnvsu), and several peripheral institutions, one for each university (*Nuclei di valutazione Interna degli Atenei*). Last but not least, the university reform was introduced in 2001, and the first 'post-reform' students obtained their degrees in 2004. This is one of the problems for evaluating the effects of the reform since most surveys conducted by Italian universities, which also gather information on students' family and academic backgrounds, collect data only on graduates rather than on students.

The reform is likely to have produced several effects on the Italian university system. Firstly, the reduction in the number of exams required to get an undegraduate degree, generally corresponding to the reduction of one year in the length of undergraduate studies, had huge effects on the number, and probably the characteristics, of university students. Indeed, Italian universities generally registered a large increase in student numbers after the reform and the reduction in the opportunity costs of studying might have increased the participation in higher education of credit constrained individuals. Secondly, another possible effect of the reform concerns the common perception that, irrespective of the reduced degree duration, the

[1] See for instance Boero et al. (2005), Broccolini and Staffolani (2005), Broccolini (2005).

difficulty of university courses in First level degrees reduced compared to the old undergraduate qualification (the old *Laurea*).

In this paper we mainly focus on this second aspect, trying to assess whether this common perception corresponds to reality. We use a data set collecting information on all 3-year graduates (i.e. students who graduated in the 'new regime') from the Faculty of Economics and Business of the Marche Polytechnic University between 2003 and 2005 (that we call Survey Graduates in Economics of Marche University, SGEMU hereafter). From the population of graduates we select individuals who enrolled at university in a time window centered around the year of the reform (i.e. 1999-2002). This data set contains information, therefore, both on the students who passed exams in the new regime ('3+2') and on those who passed exams in the old regime (when the degree duration was 4 years). Our empirical strategy consists in assessing the differences in course workloads required to pass exams and in student performance indicators (such as grades, probability of passing exams, etc.) in first-year courses between the students enrolled before the reform and those enrolled after the reform. In order to distinguish these differences from the effects produced by changes in students' characteristics after the reform, we use propensity score matching (PSM) and match individuals who enrolled after the reform (*treated*) with those who enrolled before the reform (*control*) with similar characteristics. The SGEMU data set also enables us to compare the differences in the characteristics of 3-year graduates enrolled before the reform with those of graduates enrolled after, which are the result of the cumulative effects of the '3+2' reform on enrollment and drop-out rates.[2] Unfortunately, SGEMU does not gather information on graduates' labour market outcomes and, for this reason, we are not able to investigate the effects of the '3+2' reform on the labour market. However, in the last section of this paper we will put forward some possible labour market implications of our empirical findings, which could be tested when data become available.

Our paper aims only at being a first step towards a more extensive and systematic evaluation of the effects of the '3+2' reform and a monitoring of the Italian university system that goes beyond the diffusion of raw descriptive statistics. Although our paper features a case study, and evidence from the Marche Polytechnic University cannot be straightforwardly generalised to the whole Italian university system, our analysis is nonetheless informative given the general lack of evaluation of the '3+2' Italian university reform using micro-level data. Our study might also be of interest to an international audience since similar university reforms have been implemented in other countries, and it would be interesting to compare evidence across countries.

[2]However, SGEMU does not allow the analysis of changes in the characteristics of entrant students.

The structure of the paper is as follows. The next section summarises the characteristics of the Italian university system before the reform, the main features of the 2001 reform and puts forward some expected effects of the reform. Section 3 describes the methodology used in this paper, i.e. the propensity score matching method, the data set and the estimation sample, and finally reports the results of the empirical analysis. Section 4 concludes.

2 University reform in Italy

In this section we briefly describe the characteristics of the Italian university system before the reform, the main features of the '3+2' reform and its expected effects on the behaviour of institutions and students.

2.1 The Italian university system before the reform

The Italian university system has recently experienced a huge process of normative modification of many of its fundamental elements. Ministerial decree n. 509/1999 introduced a new framework regarding the Italian higher education system that came into effect in the academic year 2001/2002.

In the previous system, university studies were organized around only one level of qualification, the *Laurea* degree, whose legal length varied between 4 and 6 years, depending on the field of study. Even after the introduction of the postgraduate qualification, the Ph.D. programme, in 1980 (lasting three years)[3] and university diplomas (*Diploma Universitario*) in 1990 (2-3 years of length and mainly of a vocational nature)[4], the *Laurea* degree remained the main higher education qualification, both at a social and academic level.[5] Therefore, the architecture of the Italian undergraduate university system before the reform was of a one-level type, although it allowed two parallel academic routes, *Diploma Universitario* and *Laurea*: it was a 'binary one tier' system. The structure of the Italian university system before the 2001 reform is shown in Figure 1. After higher secondary school students could enrol either in degree (*Laurea*) or in diploma (*Diploma Universitario*) courses. After obtaining a university degree students could enrol in Ph.D. or in specialization courses.

Besides, there was a strong centralization of decision-making: the syllabus for each single course was laid down at the national level by the National University Council (*Consiglio Nazionale Universitario).*

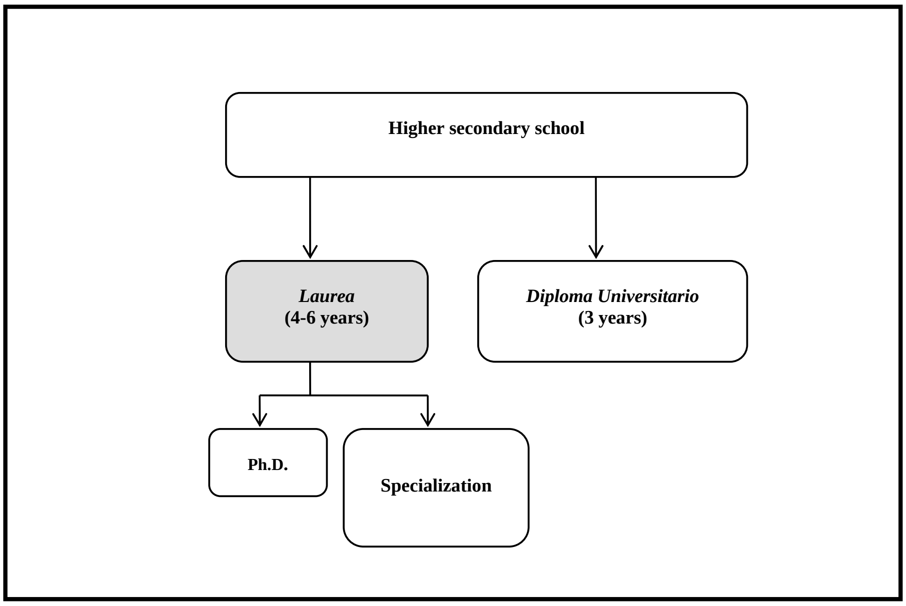
The main goal of the reform was to solve some of the most critical problems of the Italian university. The Italian higher education system has

[3] DPR 382/1980.

[4] Law n. 341/1990.

[5] During 1999, only 6% of the total of university students were studying for a *Diploma Universitario.*

Figure 1: The Italian University system before the 2001 reform ("binary one-tier system")



always been characterized firstly by a remarkable rigidity of curricula, secontly by a mismatch between university education and qualifications and skills demand of the labour market, thirdly by a large number of students who withdrew from their studies and lastly by an actual time of graduation much longer than in most other developed countries.

Furthermore, drop-out rates were particularly high during the first years of degree courses: in the academic year 1999/2000, 20.3% of first year students did not renew their enrolment in the second year.[6] The Italian university drop-out rate was the highest among the European countries: OECD student survival rates[7] in 2000 was the lowest for Italy (42%) compared to an OECD average of 70% (70% for Germany, 59% for France, 83% for the UK).[8]

The percentage of graduates in the population aged 25-34 in Italy was below 10% during 1999-2001, compared with an OECD average of 26% (22% for Germany, 32% for France, 28% for the UK).[9] In 1999, only 6.5% (7.3%

[6]Source: Miur Cnvsu (2005).

[7]Survival rates are calculated as the ratio of the number of students who are awarded a degree to the number of new entrant students n years before, n being the number of years of full-time study required to complete the degree.

[8]Source: OECD (2002).

[9]Source: OECD (2002).

in 2000) of Italian graduates obtained the degree within the legal length of the course, while more than 40.1% (40.2 in 2000) took 8 years or more to graduate.[10]

The gap between actual and legal degree duration was partly due to the high percentage of 'inactive' students - those who did not pass any exam during a given academic year. In the academic year 1999/2000, for instance, 22.8% of Italian students were 'inactive'.[11]

However, it must be noted that these dysfunctions were partly caused by the didactic organization of Italian universities which remained largely untouched by the reform. Indeed, Italian students are free to choose whether to attend lectures and classes or not and when to sit exams. Courses are usually assessed at the end of the teaching periods and exams can be repeated without limitations, in multiple alternative sessions during the same year or the following academic years. Students can 'refuse' a mark in case they are not satisfied with their performance in a specific exam and attempt the exam in subsequent sessions. Exam failures are not usually registered in the students' records. Moreover, there are usually no constraints on the number of exams to be passed in order to enrol in the following year.

2.2 Characteristics of the '3+2' reform

Following the Sorbonne Joint Declaration (Paris, May 25 1998) and the Bologna Declaration (June 19 1999),[12] ministerial decree n. 509/99 produced a radical transformation in the Italian university system, through three main changes.

First of all, the reform has granted universities full teaching autonomy. They can freely decide the names of the degree courses as well as their curricula.

Secondly, the most important innovation was represented by the introduction of a '3+2' scheme that replaced the old *Laurea*. The new academic qualifications are organized around three main levels:

- First Level degrees (3 years of legal duration) that are supposed to provide undergraduate students with adequate knowledge of general scientific principles as well as specific professional skills;

- Second Level degrees (2 years of legal duration) that should provide graduate students with advanced education and training for highly qualified professions in specific sectors;[13]

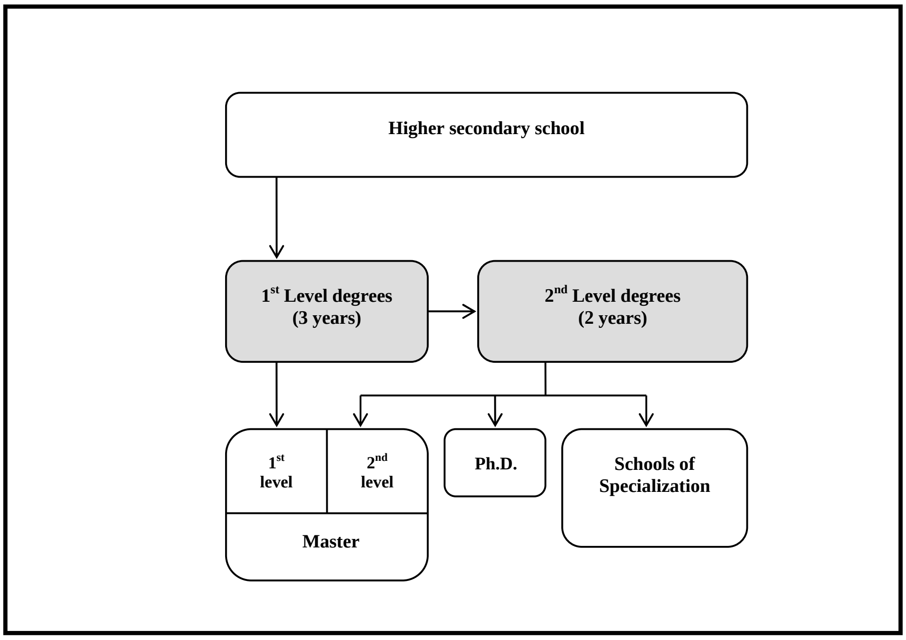
[10]Source: Miur Cnvsu (2005).

[11]Source: Miur Cnvsu (2005).

[12]which promote the creation of a *European Higher Education Area* (EHEA) through the harmonization of the different European educational systems.

[13]The articulation in a '3+2' system does not apply to Medicine, Veterinary and Architecture for which students directly enrol in Specialist degrees of six and five years, respectively.

Figure 2: The Italian University system after the 2001 reform ("unitary two-tier system")



- Ph.D. degrees (at least three years) that aim at training postgraduates for highly specialized research.

The old university system was therefore transformed into a 'unitary two-tier' system whose architecture is shown in Figure 2. After completing higher secondary school students can now enrol in First Level degrees. After getting a First Level degree students can enrol either in First Level Masters courses or in Second Level degrees. After getting a Second Level Degree students can enrol in Second Level Masters courses, in Ph.D. courses or in Specialization courses.

Finally, the reform introduced a system of university credits.[14] Credits represent the total student workload (including class time, self-study, practical activities, etc.) and they are obtained once a student has passed the assessment for each course. Each credit corresponds to 25 hours of total activities, and the average full-time workload for one academic year is 60 credits.

The main goals of the university reform were to bring the Italian higher education system in line with the European university model and to promote

[14]The Italian university credit system is based on the European Credit Transfer System (ECTS) implemented in the international student exchange programme ERASMUS.

international student mobility. The introduction of a shorter degree course aimed to increase the number of graduates, to lower their average age at graduation and to reduce drop-out rates.

The improvement of universities' 'efficiency' also grants greater financial flows to higher education institutions. Law n. 370/1999 establishes that the amount of public financing to each university is tied to specific parameters of productivity, such as drop-out rates or students' graduation times. Consequently, a potential risk is the distortion of universities' incentives produced by these funding rules. In other words, universities could be interested in finding ways to simplify degree programmes in order to increase their 'efficiency', reducing course workload or increasing grades, rather than improving the quality of the educational services they provide, such as developing new and more effective teaching strategies.[15]

2.3 Expected effects of the reform

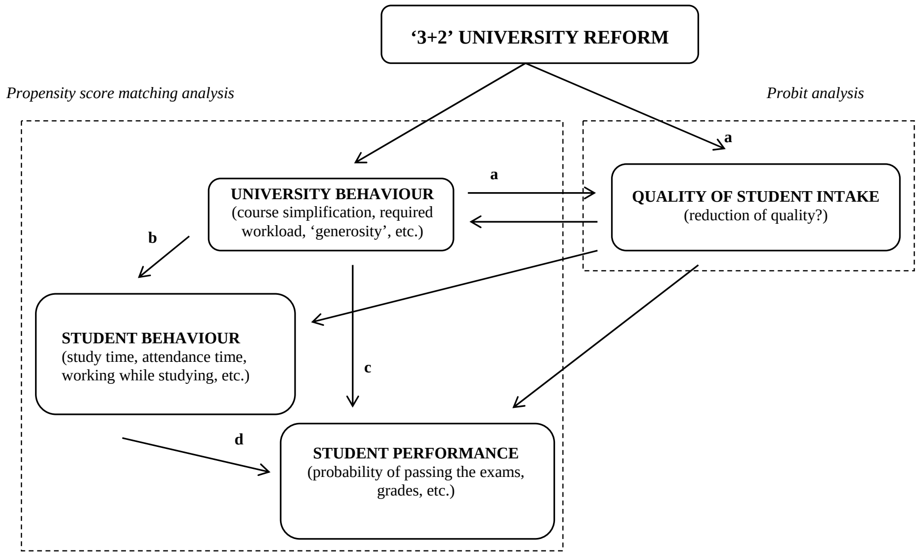
Some possible effects of the '3+2' reform are shown in Figure 3. First Level degrees in Economics and Business Administration, for instance, have a duration of three years while in the old regime the *Laurea* in the same field had a duration of four years. This restructuring of university courses affected the content of course programmes, generally determining shorter and often simplified programmes. This had immediate consequences on the number of students enrolled in first level degrees, which registered a steep increase in the year of the introduction of the reform. In the Faculty of Economics of the Marche Polytechnic University, for instance, the number of first year enrolments rose by 28.3% in 2001/2002, while first year student growth rates were 0.56% in 1999/2000, 4.7% in 2000/2001, 4.3% in 2002/2003 and 4.8% in 2003/2004.[16] We expect that the reduction of course programmes along with their simplification (since the main goal of the reform was to reduce drop-out rates and graduation times) might also have had an effect on the characteristics of university students (Figure 3, channel a). In particular, the reform might have increased the probabilities of enrolment of relatively 'weaker students', in terms of ability and motivation, of those that have less time to devote to study, such as mature students, and of more risk adverse students, such as those coming from lower social class backgrounds. This could have produced further feedbacks in terms of behaviour of HE institutions, which had to meet the needs of new 'types' of students, such as working students or students with weaker academic backgrounds.

The behaviour of universities, in turn, affects student's behaviour (Figure 3, channel b). Shorter and easier courses might have reduced the study

[15]This might have further negative consequences on the amount of skills that graduates acquire through higher education and on the economic returns to university education in the labour market.

[16]Source: www.cnvsu.it.

Figure 3: An analytical framework for the effects of the Italian university reform



and attendance times required of students to pass an exam or to achieve a good performance. The quality of the student intake also indirectly affects the average student's behaviour through 'peer effects'. Study and attendance times in turn affect student performance in terms of probability of passing exams, speed of progression and grades (Figure 3, channel d). The latter might also have been directly affected by the behaviour of universities, i.e. university teachers might have become more generous to students irrespective of the required course workloads (grade inflation), in the attempt to reduce drop-out rates and increase the number of graduates (Figure 3, channel c).

We do not have enough data to investigate all the effects of the reform and we will focus on channels a, b, c and d in Figure 3. In particular, we will attempt to analyse the changes in the characteristics of university students due to the '3+2' reform using a probit analysis. The effects of the reform on student behaviour, institution behaviour and student performance, which represents the main focus of our analysis, will be assessed using propensity score matching methods.

3 Empirical analysis

In this section we describe the methodology used in the empirical analysis, the characteristics of the data set and the estimation sample, and comment on the main results of the propensity score matching analysis.

3.1 Behavioural and performance outcomes: propensity score matching (PSM) analysis

We have seen in Figure 3 that the introduction of the '3+2' reform might have affected the behaviour of students and higher education institutions. In this section we consider the five courses that are taught to first year undergraduates at the Faculty of Economics of the Marche Polytechnic University (Mathematics, Private Law, Economics, Accounting, Economic History) and analyse the effect of the reform on the following student behavioural and performance outcomes:

- course workloads, defined as the total number of hours that students spent studying and attending lectures and classes in order to pass each of the five first year exams (and the total workload for the five exams);
- grades obtained in each exam (and the average grade);
- probability of failing each first year exam at least once during the degree course (and the total number of failures in first year exams, during the whole degree length);
- probability of passing the exams in the first year (and the number of first year exams passed in the first year).

In section 2.3 we have anticipated some possible effects of the reform. The restructuring of the university courses might have:

- reduced the course workload required to pass the exams;
- increased the grades students receive since courses are now easier and/or teachers are more "generous" to students (grade inflation);
- reduced the number of exams failed and increased the probability that students pass the exams in the first year.

In order to analyse the causal effects of the reform, we use the propensity score matching (PSM, hereafter) method. This method is now very popular, and we give only a brief introduction here. For a recent survey, the interested reader is referred to Caliendo and Kopeinig (2005). Two major advantages of PSM with respect to traditional regression analysis are that, in this specific context, it allows for heterogeneous effects of the university reform (the 'treatment' in which we are interested) on individuals with different observed characteristics[17] and that it easily highlights eventual problems of 'common support'. These problems arise from the fact that the students enrolled before and those enrolled after the reform may systematically differ with respect to observable characteristics. In this case it might be difficult to identify the effects of the reform, which can be confounded with those of the change in students' characteristics. PSM explicitly accounts for this by matching 'similar' individuals (i.e. those with similar observable characteristics) when computing the effect of the reform.

We define as Y_i the outcome of interest (workload, grades, probability of failing, probability to pass the exam in the first year), where $i = 1, ...N$ is the underscript for individuals. We want to estimate the causal effect of a treatment D_i, in our case the fact that a student enrolled after the reform ($D_i = 1$), on the various outcome variables. The treatment effect for an individual i can be defined as $\tau_i = Y_i(1) - Y_i(0)$ where $Y_i(1)$ and $Y_i(0)$ represent the outcomes of individual i when she receives the treatment ($D_i = 1$, 'post-reform' students) and when she does not receive it ($D_i = 0$, 'pre-reform' students), respectively. The problem is that we observe the student i and her outcome only in one of the two possible regimes and we do not have the so called counter factual evidence, i.e. the outcome in the unobserved regime.

We focus our attention here on the so called 'average treatment effect on the treated' (ATT) which is defined as:

$$ATT = E(\tau_i|D_i = 1) = E[Y_i(1)|D_i = 1] - E(Y_i(0)|D_i = 1]. \qquad (1)$$

In this case we do not observe the counter factual $E(Y_i(0)|D_i = 1]$ which is necessary to compute ATT. Under some particular identifying assumptions, we are able to compute ATT using the propensity score matching method.

The first identifying assumption is the so called *Conditional Independence Assumption* (CIA):

$$Y_i(0), Y_i(1) \perp D_i | X_i, \forall X_i. \qquad (2)$$

which implies that selection into the treatment is solely based on observable characteristics and there is no selection on unobservables. This is a very strong assumption and is usually justified in terms of the richness of the available data set. If the above condition holds and we define the propensity score as the probability of receiving the treatment conditional on the observables, i.e. $P(D_i = 1|X_i) = P(X_i)$ then the CIA based on the propensity score (PS) also holds and can be written as:

$$Y_i(0), Y_i(1) \perp D_i | P(X_i), \forall X_i. \qquad (3)$$

[17] This can be obtained also in regression analyses by including interaction effects of the 'treatment' with observable individual characteristics. However, given the very high number of covariates to be included this is possible only when the sample size is large enough.

A further condition (*overlap condition*) requires that the propensity score must be comprised between 0 and 1, i.e. $0 < P(X_i) < 1$, which ensures that individuals with the same X_i have a non zero probability of being both participants and non-participants (i.e. the observables X_i do not perfectly predict the treatment status).

Given these assumptions, ATT can be estimated using different methods of matching treated with untreated individuals. We consider in detail only two specific ways of matching individuals which seem to be more suitable to the characteristics of our data set, radius matching and kernel matching, respectively:

- *radius matching*. Let us define as T the set of treated units and indicate with $C_0(i)$ the set of control units (which may contain one or more individuals) associated with individual i. With radius matching $C(i) = \{P(X_j) | \, \|P(X_j) - P(X_i)\| < r\}$, that is all untreated individuals j with a propensity score falling within a radius r from $P(X_i)$ are matched to the individual i. In this case ATT is computed according to the following formula:

$$ATT_{radius} = \frac{1}{N^T} \sum_{i \in T} \left[Y_i^T - \sum_{j \in C(i)} w_{ij} Y_j^C \right] \tag{4}$$

  where $w_{ij} = 1/N_i^C$ are some weights, N_i^C is the number of control units matched with i, N^T the number of treated units, and Y_i^T and Y_j^C the outcomes of treated and untreated individuals, respectively. In the computation we impose common support by dropping treated individuals whose PS is higher than the maximum or less than the minimum PS of the controls. In our application, we specify quite a small radius ($r = 0.005$) and in such a way, using only a relatively small number of untreated individuals who are 'very close' to the treated ones (i.e. in terms of the PS), we minimise the risk of bias in our estimate of ATT at the cost of a high variance and the risk of obtaining statistically insignificant estimates (see Caliendo and Kopeinig 2005). These results will then be compared with those obtained using kernel matching that reduces the variance of the ATT estimate at the cost of a greater risk of bias;

- *kernel matching*. In this case ATT is computed as:

$$ATT_{kernel} = \frac{1}{N^T} \sum_{i \in T} \left[Y_i^T - \frac{\sum_{j \in C} Y_j^C G(\frac{P(X_j) - P(X_i)}{h_n})}{\sum_{j \in C} G(\frac{P(X_j) - P(X_i)}{h_n})} \right] \tag{5}$$

  where $G(\cdot)$ is a kernel function and h_n a bandwidth parameter. We impose common support and use the Gaussian type of kernel. This type of matching uses all untreated individuals for each treated and therefore uses a higher quantity of information with respect to radius matching, reducing the variance of the estimated ATT but increasing the risk of using untreated individuals who are 'bad matches'. DiNardo and Tobias (2001) show that the choice of kernel type is relatively unimportant for the results which might instead be affected by the bandwidth parameter, where a small bandwidth reduces the risk of bias while increasing the variance of the ATT estimate. We chose optimal bandwidth using cross-validation (see Härdle, 1991). The diagnostic statistics reported for the PSM analysis in Appendix B show that the optimal bandwidth parameter turns out to be quite small (ranging between 0.04 and 0.06). Hence, also in the case of kernel matching our estimates of the ATT should not be subject to a sizeable bias.

Black and Smith (2004) show that ATT estimates might be sensitive to the covariates used in the estimation of the PS. In particular, if one uses many variables in the computation of PS the CIA is more likely to hold. However, the use of irrelevant variables (i.e. those not affecting outcomes and/or treatment) for the computation of PS, wich will then be used to match treated with control individuals, may produce 'bad matches' and increases the variance of ATT estimates. For this reason, for both radius and kernel matching we reported two sets of results, those produced using a baseline specification including a large set of covariates for the computation of the PS (labelled as 'all'), and those obtained using only significant variables or only variables marginally not significant at 10% statistical level (labeled as 'sig').[18] The baseline specification of the PS was estimated using a probit model including grades obtained in the secondary school final exam, group of surname, if applicable,[19] age group, gender, type of secondary school, residence in the province of Ancona[20] before enrolment, distance from the university during the degree course, working while studying, parents' education and social classes, reason for enrolling at university. Some descriptive statistics are reported in section 3.2.[21]

We have already stated that one of the identifying assumptions of using the PSM method for computing ATT is that the treatment is not endogenous, that is, the assignment to treatment only depends on variables that are observable and that have been included in the estimation of the PS.

[18]For groups of categorical variables, such as parents' education dummies, a joint significance test was performed

[19]Since for some courses classes are split according to students' surnames.

[20]The Marche Polytechnic University is located in the city of Ancona.

[21]We decided to drop out from the analysis the observations with missing values in the explanatory variables which are used to compute the PS in order to avoid matching individuals with missing information in the same variables, as, in reality, they may have very different characteristics.

In order to assess the sensitivity of ATT estimates to this assumption, in particular to the presence of one unobserved confounding variable which affects treatment assignment, it is possible to use the method indicated by Rosembaum (2002). Rosembaum's method enables the researcher to build some bounds for the significance levels of ATT in the case of endogenous self-selection into the treatment status and according to different assumptions about the degree of severity of the hidden bias. More in detail, for a given measure of the bias Γ,[22] which represents the difference in the odds of treatment due to the unobserved variable, the method provides bounds for the p-value of the ATT enabling the researcher to assess "the strength such unmeasured influences would require in order that the estimated treatment effects from propensity score matching would have arisen purely through selection effects" (DiPrete and Gangl 2004, p. 14). It is important to note that Rosembaum's bounds represent worst case scenarios, i.e. they refer to the case in which one unobserved variable not only has an effect on the PS but also on the outcome so big so as to make the estimated effect spurious and only determined by the unobserved heterogeneity.[23] We run a sensitivity analysis by making Γ varying from 1 to 3 with a step of 0.05.[24]

3.2 Data

An electronic questionnaire must be filled out by all the students of the Faculty of Economics Marche Polytechnic University when they apply for graduation. This survey (that we will call SGEMU, Survey Graduates in Economics of Marche University) collects information on student backgrounds (family, previous studies, age at enrolment etc.), student behaviour during studies (course attendance, time devoted to study, failures at exams, etc.) and students' opinions on the different aspects of university life. SGEMU started in 2003 and has collected information on $1,180$ graduates since then. Around 70% of graduates in the 2003-2005 period enrolled between 1999 and 2002. For obvious reasons of comparability, only students enrolled between 1999-2002 will be considered in our analysis.[25]

The SGEMU database collects information only on students who get a First Level degree: drop outs, students who did not finish studies by 2005 and those graduated with the old *Laurea* are excluded. Therefore, our analysis is conditional on graduation with the new regime, i.e. we will

[22] $\Gamma = 1$ is the case of no hidden bias. $\Gamma = 1.5$, for instance, indicates that the unobserved variable is responsible for 50% of the difference in the odds ratio of treatment.

[23] For a detailed description of the method see DiPrete and Gangl (2004).

[24] The p-value reported in the tables indicates the probability value at which the confidence interval for the estimated ATT contains zero, it is p^+ for positive effects and p^- for negative effects (see DiPrete and Gangl 2004).

[25] We do not consider students enrolled before 1999 because, due to the characteristics of the data set, their studies last at least five years while students who enrolled after the reform graduated in less than five years.

Table 1: **Year of enrolment by year of graduation**

	year of graduation			
year of enrolment	2003	2004	2005	Total
1999	109	0	0	109
2000	118	122	0	240
2001	0	123	169	292
2002	0	0	187	187
Total	227	245	356	828

compare the behaviour and performance of 'new graduates' that enrolled in the two different regimes (before and after reform, respectively).[26] The reform might also have had an effect on the probability of drop out, but we cannot investigate these effects with the SGEMU data set.

Table 1 shows the composition of our sample. 828 students will be considered, 349 of them enrolled pre-reform and 479 enrolled after the reform.[27]

Some descriptive statistics of the variables we use to compute propensity scores are presented in Table 2 that shows the percentage composition for categorical variables and the average grade obtained in the secondary school final exam (*Maturità* grade).[28] Some differences in the composition of graduates enrolled before and after the reform seem to emerge: for instance, in academic years 2001 and 2002 women, children of blue collar parents and students coming from schools different from *liceo* showed an increase in their enrolment rates. [29]

[26] In this respect, in the educational production context, our analysis is equivalent of the analyses that use PSM to estimate wage returns to education and focus only on employed individuals (see, for instance, Blundell et al., 2005).

[27] Students enrolled before 2001 could decide either to continue studies with the old regime, based on a legal length of four years, or to switch to the new regime (based on a legal length of three years and requiring a lower number of exams). We analysed the probability to switch to the new regime (see Appendix A), finding that the main explanatory variable is the year of enrolment, so that 'old' students whose academic careers were nearly finished preferred to remain in the 'old' system.

[28] In Italy students at the end of upper secondary education have to pass an exam called 'Esame di Maturità' in which they receive a final grade ranging between 60 and 100. Before the reform of the 'Esame di Maturità' (Law n. 425/1997), the grade ranged between 36 and 60. For reasons of comparability the final grade of 'new' secondary school graduates was converted into the 'old' scale.

[29] In Italy there are various types of higher secondary schools. We grouped under 'technical schools', those schools which mainly give vocational and technical education, *Istituti Tecnici Industriali*, *Istituti Tecnici per Geometri* and *Istituti Professionali*. The group 'accounting' comprises *Istituti Tecnici Commerciali* and offers vocational education mainly in the fields of accounting and business. *Licei* are the schools that give a general type of education and that are usually chosen by individuals who plan to go on to higher education. *Liceo classico* mainly gives 'classical' education (for instance students study ancient Greek and Latin) while in *liceo scientifico* emphasis is on quantitative disciplines. In the last group, 'languages, arts and education' we have grouped *Licei Linguistici*, *Licei Artistici*, *Istituti d'Arte* and *Scuole Magistrali*.

Table 2: **Graduates' characteristics before and after the reform (%)**

Variables	pre-post reform		
Age at enrolment	pre-reform	post-reform	Total
less than 19	26.6	28.6	27.8
19	61.3	57.0	58.8
20	7.4	5.2	6.2
21	1.4	3.8	2.8
22 or more	3.2	5.4	4.5
Sex	pre-reform	post-reform	Total
men	37.2	30.3	33.2
women	62.8	69.7	66.8
Father's education	pre-reform	post-reform	Total
no or elementary school	19.2	15.1	16.8
lower secondary school	28.8	31.2	30.2
higher secondary school	38.7	38.6	38.7
degree	13.4	15.1	14.3
Mother's education	pre-reform	post-reform	Total
no or elementary school	19.9	16.4	17.9
lower secondary school	28.1	32.3	30.5
higher secondary school	40.4	37.1	38.5
degree	11.7	14.2	13.2
Father's occupation	pre-reform	post-reform	Total
entrepreneur, manager	12.0	17.7	15.3
independent worker	12.0	13.8	13.0
white collar	22.1	15.0	18.0
blue collar	10.6	16.9	14.3
other	43.3	36.5	39.4
Mother's occupation	pre-reform	post-reform	Total
entrepreneur, manager	4.6	4.0	4.2
independent worker	7.2	6.1	6.5
white collar	31.8	28.4	29.8
blue collar	14.0	17.1	15.8
housewife	27.8	23.0	25.0
other	14.6	21.5	18.6
Type of higher secondary school type	pre-reform	post-reform	Total
technical	6.1	6.7	6.4
accounting	38.3	43.5	41.3
liceo classico	8.1	5.3	6.4
liceo scientifico	41.5	30.7	35.2
languages, arts, education	6.1	13.9	10.6
Residence at less than one hour during studies	pre-reform	post-reform	Total
more than 75%	68.7	79.3	74.1
between 50% and 75%	6.9	7.4	7.2
between 25% and 50%	13.6	2.0	6.9
less than 25%	10.8	11.3	11.1
Work while studying	pre-reform	post-reform	Total
full time, continuously	10.7	3.2	6.4
part time, continuously	11.8	12.7	12.3
seasonally, temporarily	45.0	49.4	47.5
never worked	32.5	34.7	33.8
Resident in Ancona province	pre-reform	post-reform	Total
no	42.7	44.1	43.5
yes	57.3	55.9	56.5
***Maturità* grade (higher secondary school) out of** 60	pre-reform	post-reform	Total
grade	51.6	52.1	51.9

According to Becker's human capital theory (Becker, 1964) individuals self-select into higher education on the basis of their costs and returns, which depend on observed characteristics such as their family background. If none of these observed characteristics has a significant impact on the probability of enrolling at university before vs. after the reform, it is likely that the reform did not change the costs and the returns of higher education and had no effect on the characteristics of universities' student intake. By contrast, if we observe some significant differences in students' characteristics before and after the reform, we think that they are likely to be mainly produced by the reform since other mechanisms, such as changes in the labour market, are likely to operate much more slowly and require more than three years (1999-2001) to affect the costs and the returns of higher education.

Estimating a probit model where the dependent variable is the probability to enrol after the reform, some of the explanatory variables turns out to be statistically significant as Table 3 shows. The probit model of Table 3 includes only significant variables: *Maturità* grade and type of higher secondary school, distance between the student's residence and faculty, father's education and work and mother's education. Results show that students with a higher *Maturità* grade, older students, those coming from *languages, arts and education*, those who do not work full time, those coming from more educated families, children coming from higher social classes, are more likely to have enrolled after the reform.[30] Therefore, graduates from the Faculty of Economics of the Marche Polytechnic University who graduated in three or four years and who enrolled before the reform appear to differ from those who graduated in the same time but enrolled after the reform according to various observed characteristics.

As we said above, SGEMU offers information on student performance and behaviour. In particular, for each course of the first year[31] there are data on:[32]

[30] Since we consider only students who graduated in three or four years before and after reform, it is likely that some of the counter-intuitive effects we find are determined by the differential effect of the reform on the probability of graduation and on graduation times of students with different characteristics.

[31] At the Faculty of Economics of Marche Polytechnic University, first year exams are common to all students. These exams are: Mathematics, Private Law, Economics, Accounting, Economic History. Each course gives 10 credits to students, equivalent to 250 hours of total workload. During the first year, students should also pass two "qualifications" in computer use and the English language, which give five credits each.

[32] All this information is retrospective, and like in any other retrospective survey, data may be affected by memory recollection errors. In Italy university students have a personal university booklet in which the dates and grades of the exams passed are recorded. Therefore, it is likely that students checked their booklets when answering to SGEMU. However, for the first two pieces of information data may be subject to serious measurement errors. In particular, a possible objection is that post-reform students who do not remember workloads tend to declare workloads that corresponds to 'nominal workloads' (i.e. the number of credits multiplied by 25 hours, which is the equivalent of one credit).

Table 3: **Probit estimates of graduates' probability to have enrolled after university reform**

Variables	Marg. Eff.		s.e.
Maturità grade	0.00	*	0.00
Age (less than 19)			
19	-0.04		0.04
20	-0.08		0.08
21	0.19	*	0.10
22 or more	0.18	**	0.08
School (technical)			
accounting	0.06		0.08
liceo classico	-0.13		0.10
liceo scientifico	-0.07		0.08
languages, arts, education	0.15	*	0.08
Time to get to faculty more than 1h			
(more than 75% of the duration)			
between 50% and 75%	0.02		0.07
between 25% and 50%	-0.38	***	0.06
less than 25%	-0.05		0.06
Work while studying (full time, continuously)			
part time, continuously	0.21	**	0.09
seasonally, temporarily	0.24	***	0.08
never worked	0.24	***	0.08
Father's education (no or elementary school)			
lower secondary school	0.08		0.06
higher secondary school	0.11	*	0.06
degree	0.15	*	0.08
Father's occupation (entrepreneur, manager)			
independent worker	-0.01		0.07
white collar	-0.14	**	0.06
blue collar	-0.00		0.07
other	-0.08		0.06
Mother's education (no or elementary school)			
lower secondary school	0.03		0.06
higher secondary school	0.02		0.06
degree	0.07		0.08
No. obs.	728		
Pseudo R^2	0.093		
$Prob > \chi^2$	0.00		

Note. Reference categories for categorical variables are reported in brackets.

Table 4: **Descriptive statistics: student behaviour and performance**

	Overall /	Maths	Accounting	Econ.	Pr. Law	Ec. Hist.
			Workload			
pre-reform	1254.34	267.52	212.66	275.45	321.88	170.15
post-reform	1126.29	246.72	215.06	248.09	245.08	171.97
Total	1181.82	255.76	214.02	260.40	278.56	171.16
			Grade			
pre-reform	24.94	24.26	26.55	24.55	23.98	25.38
post-reform	25.98	24.74	27.41	25.75	25.65	26.36
Total	25.78	24.65	27.24	25.52	25.32	26.16
			No. of failures / prob. of failing			
pre-reform	1.27	0.28	0.10	0.33	0.12	0.02
post-reform	0.73	0.17	0.07	0.25	0.08	0.02
Total	0.96	0.21	0.08	0.29	0.10	0.02
			No. of exams passed in the 1st year / prob. of success in the 1^{st} year			
pre-reform	4.02	0.73	0.86	0.68	0.75	0.95
post-reform	4.32	0.90	0.94	0.83	0.71	0.94
Total	4.19	0.82	0.91	0.76	0.73	0.94

- hours spent attending the courses;
- hours devoted to self-study;
- grades obtained;
- exams failed;
- month and year in which the exam was passed

Table 4 shows these data,[33] distinguishing between students enrolled before and after the university reform.

It clearly emerges that the first year became easier (second column): workload reduced, the average grade increased, the number of exams failed reduced and the number of exams passed during the first year increased.

There was a reduction in student workload in Maths, Economics and Private Law in the post-reform period and it is now at around 250 hours.[34] In Maths, Accounting and Economic History grades seem to have increased after the reform. For all courses, except Economic History, students show a lower probability of failing and a higher probability of passing the exams in the first year. In particular, a higher probability of passing first year

If students behave in this way, we should consistently observe similar workloads in all exams (as each exam gives 10 credits and data refer to the same graduates). However, the reader can check from the column 'treated avg' in Tables 6-11 that post-reform students' workloads differ across exams. Moreover, also pre-reform students might behave in the same way, but this is not the case as suggested by the colunmn 'untr. avg' in Tables 6-11.

[33]The variable *workload* is obtained by summing the time spent attending the courses and the time spent studying.

[34]According to the Italian legislation, 250 hours is exactly the amount of workload for each exam. Therefore, the five exams presented in Table 4 should require a total workload of 1250 hours.

exams during the first year should imply, *ceteris paribus*, a lower probability of dropping out and a shorter time needed to get a degree, in line with the aims of the reform.

3.3 The sample

Our aim is to evaluate the effects of the '3+2' university reform on student behaviour and performance. In section 3.2, we defined as being 'pre-reform' (i.e. non-treated) those students who enrolled before 2001 and 'post-reform' those students who enrolled in 2001 or later. Some problems related to this definition emerge in the empirical analysis.

A first problem arises for courses taught by different professors before and after the reform. In this case, the effects of the reform are indistinguishable from the effects arising from the changing of professors. Thus, in the propensity scores matching analysis, we preferred not to consider students who, in a given course, had a professor who taught only before or only after the reform. Therefore the 828 students[35] in Table 1 will be used for analysis concerning the 'overall' first academic year, whereas our sample will be composed of the observations in the third column of Table 5 when individually analysing the various exams.

A second problem arises for 'pre-reform' students who passed some of the exams in the post-reform regime. These students probably attended pre-reform courses but passed the exam in the post-reform regime, so that there is no clear way of defining the treatment. To overcome this problem we use two different samples in the estimation:

- we first analyse all students described in the third column of Table 5, defining the treatment status according to students' years of enrolment;

- we then analyse only those students who are in the 'pure' pre-reform or the 'pure' post-reform regime, that is only those individuals who enrolled before 2001 and passed the exam in the first year and those who enrolled after 2001 and passed the exam in the first year, respectively. In this case, the number of observations in our sample is strongly reduced as indicated in the fourth column of Table 5.

3.4 Results of PSM analysis

Tables 6-11 report the results of the PSM analysis both for single first year exams and for the whole first year.

Mathematics represents, in the opinion of students, one of the most difficult first year exams. Our analysis suggests a reduction in the workloads

[35] Those who completed their studies in 3 or 4 years.

Table 5: **Number of observations in the different samples used for propensity score matching**

Exams	All	Same teacher pre and post reform	Exams passed during 1st year
Mathematics	828	705	360
Private Law	828	712	287
Economics	828	547	335
Accounting	828	767	378
Economic History	828	828	379

(study and attendance times) required of students to pass the exam. This finding is robust across different methods of matching treated with control individuals and different specifications of the PS. In particular, when using radius matching and the complete specification of the PS, the reduction in workload is about 51 hours (-17%), significant at 5%, and rises to 55 hours (-18%) when only significant variables are included in the PS, significant at 1% level. ATT estimated using kernel matching varies between -15% in the full specification and -17% in the specification including only significant variables and turns out to be significant at least at 5% level. There is no effect of the reform on the average grade in Mathematics. The effects on the probability of having failed the exam at least once during the whole degree course and of passing the exam in the first year are more substantial. Reduction in the first student performance indicator is always significant at the 5% level and ranges between -39% and -45% while the increase in the second indicator is always significant at the 1% level and ranges between 21% and 28%. Therefore, the effect of the reform was particularly strong in Mathematics, one of the first year exams in which many students experience major difficulties: the hours workload reduced, the probability of failing the exam at least once dropped and that of passing the exam in the first year increased. Since the workload reduced, we cannot judge if the increase in all student performance indicators was mainly due to the reduced effort required of students or by a more generous attitude of teachers who inflated grades after the reform, or due to a mix of the two. However, the fact that the average grade did not rise after the reform make us to propend for the first explanation. One thing is clear, since by using PSM we estimate the effect of the reform matching very similar individuals, accounting therefore for differences in observable student characteristics across the two regimes, such as ability (*Maturità* grade) and academic readiness (type of secondary school), it is likely that the reduction in workloads leads into a reduction of the amount of 'knowledge' (i.e. human capital) that the students possess.

Private Law was one of the exams with the biggest workload before the reform. Table 7 shows a significant reduction in the workload. The estimates are surprisingly robust across methods and specifications of the PS and range between -17% (53 hours) and -18% (58 hours) with respect

Table 6: **Propensity score matching - Mathematics**

methods/ variables	ATT		s.e.$^{(a)}$	treated avg$^{(b)}$	untr. avg$^{(c)}$	n. obs.	Opt. bandw.$^{(d)}$	R-bounds$^{(e)}$ Γ	p-value
Workload									
radius/all	-50.7	**	22.9	251	302	518	-	2.10	0.12
kernel/all	-44.8	**	19.5	246	291	518	0.06	2.80	0.11
radius/sig	-55.1	***	21.3	246	302	525	-	2.25	0.12
kernel/sig	-51.3	***	18.8	245	297	525	0.06	(§)	(§)
Grade									
radius/all	-0.19		0.54	24.73	24.92	520	-	-	-
kernel/all	0.48		0.44	24.86	24.37	520	0.06	-	-
radius/sig	0.28		0.51	24.71	24.43	527	-	-	-
kernel/sig	0.39		0.42	24.84	24.44	527	0.06	-	-
Probability of one or more failures									
radius/all	-0.11	**	0.06	0.16	0.28	606	-	2.15	0.12
kernel/all	-0.11	**	0.05	0.18	0.28	606	0.06	1.90	0.13
radius/sig	-0.13	**	0.06	0.16	0.30	617	-	2.35	0.12
kernel/sig	-0.13	**	0.05	0.17	0.30	617	0.06	1.95	0.14
Probability of passing in the first year									
radius/all	0.16	**	0.06	0.90	0.74	525	-	(§)	(§)
kernel/all	0.19	***	0.06	0.90	0.72	525	0.06	(§)	(§)
radius/sig	0.20	***	0.06	0.92	0.72	532	-	(§)	(§)
kernel/sig	0.20	***	0.05	0.90	0.70	532	0.06	(§)	(§)

Note. The number of observations may change across analyses due to the different number of missing values in the outcome variables.
$^{(a)}$ Bootstrapped standard errors (1,000 replications).
$^{(b)}$ Averages of outcome variables for treated individuals.
$^{(c)}$ Averages of outcome variables for untreated individuals.
$^{(d)}$ The optimal bandwidth for the Gaussian kernel matching was selected using cross-validation (see Härdle, 1991).
$^{(e)}$ The Γ for the Rosembaum's bounds refers to the first value for which the estimated ATT is not significant at the 10% statistical level while the p-value reports the corresponding probability value. Γ and p-value are reported only for ATT estimates significant at least at the 1′% level.
(§) ATT estimate turns out to be significant at the 10% statistical level also for Γ = 3.
* significant at 10%; ** significant at 5%; *** significant at 10%.

to the period before the reform, generally significant at the 1% level. Our estimates show no effect on the average grade, except in the case of kernel matching with only significant variables, in which case ATT shows a 3.2% increase in the average mark. The reform had no effect on the probability of experiencing more than one fail and had a counter-intuitive negative effect on the likelihood of passing the exam in the first year, ranging between -12% and -14%. Given the evidence on the likelihood of exam failures after the reform, which did not rise, this might perhaps be due to the different strategies followed by students who after the reform might have been relatively more likely to attempt this exam after the first year.

Economics,[36] like Mathematics, is an exam with a quantitative content and reputed as quite difficult by students. Our estimates generally suggest a reduction in workload of the magnitude of 11-14%. ATT estimates including only significant variables in the PS are usually more significant. The reform

[36]The first year Economics exam corresponds to Microeconomics.

Table 7: **Propensity score matching - Private Law**

methods/ variables	ATT		s.e.[a]	treated avg[b]	untr. avg[c]	n. obs.	Opt. bandw.[d]	R-bounds[e] Γ	p-value
				Workload					
radius/all	-53.4	**	22.0	265	318	540	-	2.10	0.12
kernel/all	-57.8	***	17.8	260	317	540	0.05	2.70	0.12
radius/sig	-56.9	***	19.9	261	318	547	-	2.30	0.12
kernel/sig	-55.4	***	16.3	260	316	547	0.04	2.70	0.11
				Grade					
radius/all	0.32		0.44	25.47	25.15	543	-	-	-
kernel/all	0.59		0.36	25.47	24.87	543	0.05	-	-
radius/sig	0.54		0.42	25.42	24.89	550	-	-	-
kernel/sig	0.78	**	0.34	25.44	24.68	550	0.04	1.50	0.11
				Probability of one or more failures					
radius/all	0.01		0.03	0.10	0.09	621	-	-	-
kernel/all	-0.00		0.03	0.09	0.09	621	0.05	-	-
radius/sig	-0.05		0.04	0.09	0.14	632	-	-	-
kernel/sig	-0.02		0.03	0.09	0.11	632	0.04	-	-
				Probability of passing in the first year					
radius/all	-0.11	*	0.06	0.64	0.74	549	-	1.05	0.16
kernel/all	-0.10	**	0.05	0.64	0.74	549	0.05	1.25	0.16
radius/sig	-0.08		0.05	0.64	0.72	556	-	-	-
kernel/sig	-0.11	**	0.04	0.64	0.74	556	0.04	1.25	0.13

Note. See Table 6.

had no effect on the average grade. As in Mathematics, we register a sizeable and generally statistically significant reduction in the probability of at least one fail, which ranges between 25-35% and a remarkable increase in the probability of passing the exam in the first year ranging between 15-23%. For Economics, as for Mathematics, it is not possible to say whether the increase in student performance was mainly driven by a reduction in the amount of material to be studied (workload) or by a reduction in the difficulty of the exam irrespective of workloads, i.e. by an increase in teachers' generosity. However, also in this case, as for Mathematics, since average grades did not change after the reform we prefer the first explanation.

Accounting and Economic History were the exams with the lowest levels of workloads before the reform and the highest probability of passing at the first attempt. For such exams, we would expect a lower impact of the reform whose main aims were, as we said, to reduce drop-out rates and graduation times. Indeed, in *Accounting* we do not register any effect on workloads that remained the same after the reform. However, grades increased between 1.4 and 2 per cent (generally significant at the 5% level) while the effect on the likelihood of passing the exam in the first year is positive but generally not statistically significant. In this case a possible interpretation of our results is that after the reform teachers were more generous to students, mainly inflating grades.

Also in *Economic History* the reform produced no effect on workloads, on the probability of one or more failures and on that of passing the exam in the first year. It is interesting to observe that in this course failure was pratically

Table 8: **Propensity score matching - Economics**

methods/ variables	ATT		s.e.$^{(a)}$	treated avg$^{(b)}$	untr. avg$^{(c)}$	n. obs.	Opt. bandw.$^{(d)}$	R-bounds$^{(e)}$ Γ	p-value
				Workload					
radius/all	-31.0		21.5	256	287	425	-	-	-
kernel/all	-30.7	*	16.0	250	280	425	0.06	2.15	0.12
radius/sig	-39.0	**	19.9	246	285	434	-	1.80	0.11
kernel/sig	-32.1	**	14.2	249	281	434	0.06	2.25	0.12
				Grade					
radius/all	-0.08		0.58	25.84	25.92	433	-	-	-
kernel/all	0.28		0.42	25.93	25.65	433	0.06	-	-
radius/sig	0.05		0.53	25.83	25.78	442	-	-	-
kernel/sig	0.37		0.37	25.96	25.60	442	0.06	-	-
				Probability of one or more failures					
radius/all	-0.12	*	0.07	0.24	0.37	475	-	1.55	0.12
kernel/all	-0.13	**	0.05	0.23	0.35	475	0.06	1.25	0.13
radius/sig	-0.09		0.07	0.25	0.34	486	-	-	-
kernel/sig	-0.12	**	0.05	0.23	0.35	486	0.05	1.25	0.12
				Probability of passing in the first year					
radius/all	0.14	*	0.07	0.84	0.70	430	-	1.90	0.11
kernel/all	0.16	***	0.06	0.85	0.69	430	0.06	2.10	0.13
radius/sig	0.11		0.07	0.85	0.74	439	-	-	-
kernel/sig	0.14	***	0.05	0.85	0.70	439	0.05	2.15	0.13

Note. See Table 6.

inexistent before the reform and that most students passed the exam in the first year. The only effect produced by the reform was an increase in the average grades received by students of 1.9-2.9% (generally significant at 5% level). In this case, therefore, as in Accounting, the reform seems to have mainly produced grade inflation, without affecting the likelihood of failing or the probability of passing the exam in the first year, since most students used to pass the exam at their first attempt during the first year already before the reform. Another thing worth noting is that, although 10 credits have been attributed to the course, its workload is well below the measure of 250 hours.

When we consider the overall performance in the five first year exams all effects are highly statistically significant and go in the expected direction:

- first year workload reduced between 10 and 11 per cent. This roughly corresponds to 5 credits, that is to half of a first year exam;
- the average grade in first year exams increased between 2.2 and 2.8 per cent;
- the total number of first year exams failed reduced between 38 and 48 per cent;
- the total number of exams passed in the first year increased between 7.6 and 8.5 per cent.

Table 9: **Propensity score matching - Accounting**

methods/ variables	ATT		s.e.[(a)]	treated avg[(b)]	untr. avg[(c)]	n. obs.	Opt. bandw.[(d)]	R-bounds[(e)] Γ	p-value
				Workload					
radius/all	3.3		12.2	215	212	583	-	-	-
kernel/all	5.7		9.4	214	209	583	0.05	-	-
radius/sig	5.0		11.6	215	210	590	-	-	-
kernel/sig	6.4		8.8	214	208	590	0.05	-	-
				Grade					
radius/all	0.38		0.28	27.41	27.04	585	-	-	-
kernel/all	0.43	**	0.21	27.40	26.97	585	0.06	1.55	0.14
radius/sig	0.55	**	0.27	27.43	26.88	592	-	1.55	0.12
kernel/sig	0.48	**	0.21	27.41	26.93	592	0.05	1.60	0.11
				Probability of one or more failures					
radius/all	-0.02		0.03	0.06	0.08	669	-	-	-
kernel/all	-0.02		0.02	0.06	0.08	669	0.05	-	-
radius/sig	-0.01		0.03	0.06	0.07	680	-	-	-
kernel/sig	-0.02		0.02	0.06	0.08	680	0.04	-	-
				Probability of passing in the first year					
radius/all	0.05		0.04	0.92	0.87	592	-	-	-
kernel/all	0.05		0.03	0.93	0.88	592	0.05	-	-
radius/sig	0.05		0.03	0.93	0.88	599	-	-	-
kernel/sig	0.05	*	0.03	0.93	0.88	599	0.05	(§)	(§)

Note. See Table 6.

Table 10: **Propensity score matching - Economic History**

methods/ variables	ATT		s.e.[(a)]	treated avg[(b)]	untr. avg[(c)]	n. obs.	Opt. bandw.[(d)]	R-bounds[(e)] Γ	p-value
				Workload					
radius/all	5.4		9.1	175	170	635	-	-	-
kernel/all	-0.2		7.2	171	172	635	0.05	-	-
radius/sig	-10.5		8.6	173	183	644	-	-	-
kernel/sig	-3.0		7.0	171	174	644	0.05	-	-
				Grade					
radius/all	0.75	**	0.33	26.45	25.70	651	-	1.55	0.12
kernel/all	0.56	**	0.27	26.43	25.86	651	0.05	1.55	0.12
radius/sig	0.65	**	0.32	26.42	25.77	660	-	1.50	0.14
kernel/sig	0.48	*	0.26	26.44	25.95	660	0.05	1.70	0.13
				Probability of one or more failures					
radius/all	-0.00		0.02	0.02	0.03	718	-	-	-
kernel/all	-0.01		0.01	0.02	0.03	718	0.05	-	-
radius/sig	-0.01		0.02	0.02	0.03	730	-	-	-
kernel/sig	-0.01		0.01	0.02	0.03	730	0.04	-	-
				Probability of passing in the first year					
radius/all	-0.03		0.02	0.95	0.98	647	-	-	-
kernel/all	-0.02		0.02	0.94	0.96	647	0.05	-	-
radius/sig	-0.03		0.02	0.94	0.97	656	-	-	-
kernel/sig	-0.03		0.02	0.94	0.97	656	0.05	-	-

Note. See Table 6.

Table 11: **Propensity score matching- All 1^{st} year exams**

methods/ variables	ATT		s.e.$^{(a)}$	treated avg$^{(b)}$	untr. avg$^{(c)}$	n. obs.	Opt. bandw.$^{(d)}$	R-bounds$^{(e)}$ Γ	p-value
				Workload					
radius/all	-129.5	**	53.1	1133	1262	618	-	1.85	0.13
kernel/all	-144.8	***	44.5	1120	1266	618	0.05	2.30	0.11
radius/sig	-143.7	***	51.9	1122	1266	626	-	2.05	0.12
kernel/sig	-143.3	***	42.8	1118	1262	626	0.05	2.35	0.14
				Grade					
radius/all	0.56	**	0.24	26.10	25.54	621	-	1.50	0.13
kernel/all	0.58	***	0.18	26.06	25.49	621	0.05	1.70	0.14
radius/sig	0.70	***	0.22	26.07	25.36	629	-	1.70	0.10
kernel/sig	0.56	***	0.18	26.06	25.51	629	0.05	1.65	0.13
				Total number of failures					
radius/all	-0.44	***	0.16	0.73	1.17	718	-	2.35	0.13
kernel/all	-0.47	***	0.13	0.74	1.22	718	0.05	(§)	(§)
radius/sig	-0.65	***	0.16	0.70	1.35	730	-	(§)	(§)
kernel/sig	-0.48	***	0.12	0.73	1.22	730	0.04	(§)	(§)
				N. of exams passed in the first year					
radius/all	0.34	***	0.10	4.33	3.99	627	-	2.30	0.12
kernel/all	0.34	***	0.08	4.33	4.00	627	0.05	2.15	0.10
radius/sig	0.30	***	0.09	4.33	4.03	635	-	1.95	0.11
kernel/sig	0.31	***	0.07	4.33	4.02	635	0.05	1.60	0.11

Note. See Table 6.

The optimal bandwidths for kernel matching selected by cross-validation are quite small (ranging between 0.04 and 0.06). Therefore also ATT computed with this method is likely to suffer from a small bias. Rosembaum bounds usually show that ATT estimates turn out to be quite robust in the presence of one unobserved variable affecting both the probability of treatment and the outcomes.

In Appendix B, we report some diagnostic statistics for the matching procedure. In all cases the matching procedure produced quite satisfactory results, as shown by the reduction in the Pseudo R^2 of the probit model for treatment status before and after matching, the lack of joint significance of the covariates used for the estimation of the propensity scores and the sizeable reduction in the median bias after matching. All these diagnostics show that the covariates were well balanced between the two samples of treated and untreated individuals after matching.

3.5 An alternative definition of the treatment status

In the previous section, we considered individuals who enrolled after 2001 as 'treated'. This definition of treatment may pose some problems for the evaluation of the effect of the reform on workloads, grades and the probability of more than one failure since those students who enrolled before 2001 might have attempted and passed the exam in the new regime, i.e. after the reform. Therefore, some individuals who enrolled before 2001 may also have benefited from the reform. This suggests that the effects we estimated in the previous section might be lower bound estimates since some 'treated' individuals might have been matched with individuals who, although enrolling before 2001, passed the exams after the reform. However, we do not evidently have this problem when assessing the effect of the reform on the probability of passing the exams in the first year.

In Appendix C, we estimated the effect of the reform, i.e. having enrolled after 2001, only on individuals who passed the exams in the first year. When considering only this group we do not have the problem that some individuals enrolling before the reform passed the exam after the reform, however, the ATT we estimate may be different from that of the previous section because we consider relatively abler or more motivated individuals here. Moreover, given the small sample size, a caliper of 0.005 is probably too small and we will be unlikely to obtain significant estimates with radius matching. We report only the analysis on single exams and do not report the analysis on all first-year exams due to the very low number of individuals who passed all first-year exams in the first year.

For *Mathematics* we observe a reduction in workload generally significant at the 10% level when using kernel matching. The reduction is lower than that estimated in the previous section and ranges between 23 and 38 hours. There are no other statistically significant effects.

For *Private Law* we estimate a highly statistically significant reduction in workloads ranging between 22% and 25%, bigger than that estimated in the previous section.

For *Economics* we do not observe any statistically significant effect.

Finally both for *Accounting* and *Economic History* we estimate an increase in average grades ranging between 1.9% and 2.2% and 1.9% and 2.6%, respectively, generally statistically significant.

Therefore, the analysis restricted to individuals who passed the exams in the first year qualitatively confirms the findings of the previous section, although, as expected, the significance of the estimated ATT tends to fall.

4 Concluding remarks

In this paper we used propensity score matching to evaluate the effects of the '3+2' university reform on student academic behaviour and performance introduced in Italy in 2001. We used data on graduates (in First Level degrees) from the Faculty of Economics of the Marche Polytechnic University in the period 2003-2005. Our findings suggest that the reform produced a differentiated effect on courses with different characteristics. In those courses in which before the reform the workload was much higher than that imposed by law (250 hours), such as Mathematics, Private Law and Economics, the reform produced a significant reduction in workload. In such courses the reform also produced a reduction in the likelihood of failing the exams and an increase in passing them in the first year. Since the average grade did not increase in such courses after the reform, we think that the last two effects were directly produced by the reduced workload and simplification of the courses. By contrast, in courses that were relatively simple also before the reform, in terms of required workload and likelihood of failing, the reform mainly produced grade inflation, i.e. teachers increased the average grade irrespective of any change in course workload.

Therefore, it appears that all these changes have gone in the direction of a reduction in drop-out rates, since exam failures may be an important motivation for interrupting university studies, and in graduation times, as first year exams are those that generally pose major problems to students. In other words, the reform might attain the objectives for which it was introduced. However, what will be the effects on graduates when they enter the labour market?

It is perhaps too early to evaluate the effect of the '3+2' reform on new graduates' employment outcomes, since only a few cohorts of students enrolled after the reform and have had the time to complete their studies and enter the labour market. However, based on the empirical evidence on this paper, as a suggestion for future research we put forward some possible labour market consequences of the reform, which could be tested when data on new graduates' employment become available. Up to now, from our study we have observed three kinds of effects of the reform: 1) reduction of course workload; 2) grade inflation; 3) increase in the probability of passing exams. In short, obtaining an undergraduate degree has become easier.

What does economic theory predict as possible consequences of this fact? According to the signalling theory (Spence, 1973) a reduction in the effort required to pass exams reduces the cost for low ability students (who have greater costs per unit of effort) to acquire higher education and raises their relative convenience of getting a university degree compared to high ability students. This implies an increase in the labour market of the quantity of low ability graduates with negative consequences on average returns to education, in the case of heterogeneity of returns to education according to ability. On the other hand, firms, which might use a degree qualification as a means to screen job applicants, will see an inflation of graduates in the labour market and will find more difficult to select high ability applicants. A further consequence is that students might decide to go on for postgraduate education, such as Second Level degrees, in order to signal their ability. And this seems to have been the case. Data collected by *Almalaurea*, a consortium of 44 Italian universities, show that about 76% of First Level degree graduates in 2004 planned to go for postgraduate studies (34% of whom planned to enrol in Second Level degrees).[37] However, only increasing the degree of selectivity of Second Level degrees (compared

[37] Source: `www.almalaurea.it`

to First Level degrees) high ability students will succeed in signaling their 'type' and firms will be able to select more able and competent individuals to fill high qualified positions.

It must be noted that also the other main competitive theory of education existing in the domain of economics, the human capital theory hypothesis (Becker, 1964), produces similar predictions. A reduction of workload due to simplification of course programmes might lead to a reduction of the amount of knowledge and skills (i.e. human capital) that graduates acquire during their studies, with negative consequences on employers who will dispose of younger but probably less skilled graduates than in the past. Also in this case, an excessive simplification of university courses might make students feel inadequate for the labour market and induce them to go for postgraduate education. Post-reform graduates might also suffer from a reduction in the economic returns to their degrees, as compared to pre-reform graduates, due to their reduced amount of skills.

We hope to see in the recent future an increase in studies analysing university students' outcomes, both academic and occupational, using data from several Italian universities or the whole university system. Such evidence might be extremely useful to fully evaluate the effects produced by the '3+2' reform and eventually make some adjustments in case they are needed to improve the current university system.

Appendix A: Probability to switch to the new regime

In this Appendix we show the results of a probit model of the probability that a pre-reform student chooses to opt for the new First Level degree programmes. During the transitory phase from the old system to the new one, universities had to complete degree courses that were still running and to grant students the opportunity to opt for the new system. The aim of the analysis is to find if there are any correlations between students' individual and academic characteristics and the probability of re-enrolling in the new, shorter, degree courses.

We refer to all those students who got their degree (old or new) from the Faculty of Economics of the Marche Polytechnic University, in the 2003-2005 period, enrolled between 1999 and 2000. Our sample includes 516 students, 475 of them re-enrolled in the new system while only 41 remained in the 4-year courses. In this analysis, unlike for the PS matching analysis, we included individuals with missing values in specific variables in order not to reduce the sample size too much.

Table 12 reports the results from the probit model where the dependent variable is the probability of opting for one of the new degree courses. Results show that male students, students with lower *Maturità* grades and those who enrolled for the first time in the year before the introduction of the reform (2000) were more likely to choose a shorter degree programme.

Table 12: **Pre-reform students' probabilities to switch to (3-year) First Level degrees**

Variables	Marg. Eff		Std. Error
Gender (men)			
women	-0.034	*	0.023
Age at the enrolment (less than 19)			
19	0.041		0.026
20	0.009		0.049
21	(dropped)[a]		
22 o more	-0.145		0.097
School (technical)			
accounting	0.019		0.058
liceo classico	-0.022		0.083
liceo scientifico	-0.008		0.057
languages, arts, education	0.024		0.046
Maturità grade	-0.003	*	0.002
Father's education (no or elementary school)			
lower secondary school	-0.027		0.035
higher secondary school	-0.033		0.043
degree	-0.063		0.070
missing	-0.147		0.176
Mother's education (no or elementary school)			
lower secondary school	0.008		0.032
higher secondary school	-0.016		0.043
degree	-0.099		0.087
missing	(dropped)		
Father's occupation (entrepreneur, manager)			
independent worker	-0.037		0.053
white collar	-0.004		0.035
blue collar	-0.030		0.057
other	-0.038		0.040
missing	-0.004		0.052
Mother's occupation (entrepreneur, manager)			
independent worker	-0.023		0.076
white collar	0.009		0.046
blue collar	0.002		0.050
housekeeper	-0.006		0.047
other	0.057		0.041
missing	0.040		0.055
Year of enrolment (1999)			
2000	0.152	***	0.018
Resident in Ancona province (no)			
yes	0.043		0.039
N. obs		516	
Pseudo R^2		0.220	
$Prob > \chi^2$		0.000	

Source: *Marche Polytechnic University administrative data.*
Note. Reference categories for categorical variables are reported in brackets. [a] Categorical variables that corresponds to empty cells or are perfectly collinear to other variables are dropped from the analysis.

Appendix B: Diagnostics

Table 13: **PSM Diagnostics - Mathematics**

methods/ variables	Pseudo R^2 before$^{(a)}$	Pseudo R^2 after$^{(b)}$	$P > Chi^2$ after$^{(c)}$	% out of support$^{(d)}$	% bias reduction$^{(e)}$
			Workload		
radius/all	0.143	0.041	0.994	18.950	52.875
kernel/all	0.143	0.017	0.955	0.000	49.592
radius/sig	0.113	0.021	0.765	19.020	39.988
kernel/sig	0.113	0.012	0.984	0.000	50.949
			Grade		
radius/all	0.141	0.035	0.953	26.453	37.988
kernel/all	0.141	0.015	0.942	0.000	43.727
radius/sig	0.106	0.016	0.995	29.023	53.557
kernel/sig	0.106	0.011	0.975	0.000	60.280
			Probability of one or more fails		
radius/all	0.122	0.034	0.943	26.238	41.711
kernel/all	0.122	0.014	0.971	0.000	43.061
radius/sig	0.106	0.023	1.000	23.601	61.867
kernel/sig	0.106	0.010	0.970	0.000	60.106
			Probability of success in the first year		
radius/all	0.139	0.056	0.771	21.802	51.310
kernel/all	0.139	0.016	0.917	0.000	44.255
radius/sig	0.109	0.016	0.990	22.989	54.020
kernel/sig	0.109	0.010	0.976	0.000	52.582

Note. $^{(a)}$ Pseudo R^2 of the probit model of the treatment status before matching suggesting how well the covariates (X) used for matching treated with control units predict the probability of enrolling at university after the 2001 reform.
$^{(b)}$ Pseudo R^2 of the probit model of the treatment status estimated on the matched sample, suggesting how well PSM balance the covariates between the two samples of treated and control individuals.
$^{(c)}$ Test for the joint significance of all covariates (except the constant) in the probit model. Lack of joint significance indicates that covariates are well balanced in the two samples of treated and control individuals.
$^{(d)}$ It is the percentage of treated units falling out of the common support, i.e. with a PS less than the minimum or greater than the maximum of the propensity scores of control units giving an indication of the problem of lack of common support.
$^{(e)}$ It is the percentage reduction in the median absolute standardized bias before and after matching, with median taken over all regressors used to compute the PS. Following Rosenbaum and Rubin (1985), the standardized bias before and after matching for the covariate X, labelled as $B_{bm}(X)$ and $B_{am}(X)$ respectively, are defined as:
$B_{bm}(X) = \frac{\overline{X}_1 - \overline{X}_0}{\sqrt{[V_1(X)+V_0(X)]/2}}$, $B_{am}(X) = \frac{\overline{X}_{1M} - \overline{X}_{0M}}{\sqrt{[V_1(X)+V_0(X)]/2}}$
where $\overline{X}_0$ and $\overline{X}_1$ are means of the X covariate in the full samples of treated and untreated individuals, while $\overline{X}_{0M}$ and $\overline{X}_{1M}$ are the means of the X covariate in the matched samples of treated and control individuals, respectively, and $V_1(X)$ and $V_0(X)$ are the variances of X in the full treated and control groups.

Table 14: **PSM Diagnostics - Private Law**

methods/ variables	Pseudo R^2 before$^{(a)}$	Pseudo R^2 after$^{(b)}$	$P > Chi^2$ after$^{(c)}$	% out of support$^{(d)}$	% bias reduction$^{(e)}$
		Workload			
radius/all	0.115	0.024	0.999	10.305	49.313
kernel/all	0.115	0.008	1.000	0.000	73.505
radius/sig	0.092	0.008	1.000	4.151	75.198
kernel/sig	0.092	0.005	1.000	0.000	71.927
		Grade			
radius/all	0.120	0.022	0.999	6.630	67.196
kernel/all	0.120	0.009	1.000	0.000	71.937
radius/sig	0.100	0.011	1.000	6.000	71.897
kernel/sig	0.100	0.005	1.000	0.000	81.689
		Probability of one or more failures			
radius/all	0.112	0.012	1.000	9.121	62.256
kernel/all	0.112	0.008	1.000	0.000	75.527
radius/sig	0.095	0.008	1.000	8.626	78.500
kernel/sig	0.095	0.005	1.000	0.000	72.986
		Probability of passing in the first year			
radius/all	0.116	0.023	0.999	9.125	45.809
kernel/all	0.116	0.008	1.000	0.000	75.696
radius/sig	0.094	0.008	1.000	7.143	80.020
kernel/sig	0.094	0.006	1.000	0.000	75.268

Note. See Table 13.

Table 15: **PSM Diagnostics - Economics**

methods/ variables	Pseudo R^2 before$^{(a)}$	Pseudo R^2 after$^{(b)}$	$P > Chi^2$ after$^{(c)}$	% out of support$^{(d)}$	% bias reduction$^{(e)}$
		Workload			
radius/all	0.157	0.030	0.999	19.283	34.807
kernel/all	0.157	0.013	1.000	0.000	69.923
radius/sig	0.133	0.014	1.000	10.132	56.037
kernel/sig	0.133	0.010	1.000	0.000	62.166
		Grade			
radius/all	0.162	0.033	0.997	14.350	53.488
kernel/all	0.162	0.013	1.000	0.000	69.223
radius/sig	0.132	0.027	0.982	17.181	34.494
kernel/sig	0.132	0.010	1.000	0.000	60.305
		Probability of one or more failures			
radius/all	0.154	0.027	0.999	12.595	57.087
kernel/all	0.154	0.015	1.000	0.000	69.273
radius/sig	0.136	0.026	0.967	15.672	52.116
kernel/sig	0.136	0.013	0.999	0.000	62.348
		Probability of passing in the first year			
radius/all	0.155	0.033	0.997	15.695	58.245
kernel/all	0.155	0.013	1.000	0.000	59.676
radius/sig	0.129	0.029	0.964	10.132	25.396
kernel/sig	0.129	0.010	1.000	0.000	54.591

Note. See Table 13.

Table 16: **PSM Diagnostics - Accounting**

methods/ variables	Pseudo R^2 before$^{(a)}$	Pseudo R^2 after$^{(b)}$	$P > Chi^2$ after$^{(c)}$	% out of support$^{(d)}$	% bias reduction$^{(e)}$
		Workload			
radius/all	0.123	0.015	1.000	8.882	65.563
kernel/all	0.123	0.011	1.000	0.000	69.507
radius/sig	0.096	0.011	1.000	3.257	57.105
kernel/sig	0.096	0.006	1.000	0.000	73.121
		Grade			
radius/all	0.129	0.021	0.999	11.842	62.818
kernel/all	0.129	0.011	1.000	0.000	73.788
radius/sig	0.101	0.013	0.999	5.537	61.471
kernel/sig	0.101	0.006	1.000	0.000	76.299
		Probability of one or more failures			
radius/all	0.114	0.021	0.996	6.287	56.328
kernel/all	0.114	0.009	1.000	0.000	76.506
radius/sig	0.098	0.016	0.987	5.540	55.700
kernel/sig	0.098	0.006	1.000	0.000	67.462
		Probability of passing in the first year			
radius/all	0.123	0.017	1.000	8.553	60.002
kernel/all	0.123	0.011	1.000	0.000	64.905
radius/sig	0.096	0.008	1.000	6.515	65.647
kernel/sig	0.096	0.006	1.000	0.000	77.419

Note. See Table 13.

Table 17: **PSM Diagnostics - Economic History**

methods/ variables	Pseudo R^2 before$^{(a)}$	Pseudo R^2 after$^{(b)}$	$P > Chi^2$ after$^{(c)}$	% out of support$^{(d)}$	% bias reduction$^{(e)}$
		Workload			
radius/all	0.107	0.016	1.000	12.705	60.382
kernel/all	0.107	0.010	1.000	0.000	67.041
radius/sig	0.096	0.014	0.997	7.184	41.124
kernel/sig	0.096	0.007	1.000	0.000	72.230
		Grade			
radius/all	0.113	0.022	0.995	9.302	45.027
kernel/all	0.113	0.010	1.000	0.000	65.792
radius/sig	0.100	0.014	0.997	6.897	65.609
kernel/sig	0.100	0.007	1.000	0.000	77.619
		Probability of one or more failures			
radius/all	0.106	0.016	1.000	7.178	41.563
kernel/all	0.106	0.009	1.000	0.000	65.103
radius/sig	0.101	0.010	1.000	8.759	43.545
kernel/sig	0.101	0.007	1.000	0.000	73.357
		Probability of passing in the first year			
radius/all	0.108	0.017	1.000	9.538	51.748
kernel/all	0.108	0.011	1.000	0.000	63.209
radius/sig	0.096	0.014	0.996	6.571	58.710
kernel/sig	0.096	0.007	1.000	0.000	76.495

Note. See Table 13.

Table 18: **PSM Diagnostics - All 1^{st} year exams**

methods/ variables	Pseudo R^2 before$^{(a)}$	Pseudo R^2 after$^{(b)}$	$P > Chi^2$ after$^{(c)}$	% out of support$^{(d)}$	% bias reduction$^{(e)}$
			Workload		
radius/all	0.113	0.016	1.000	7.872	72.881
kernel/all	0.113	0.010	1.000	0.000	71.232
radius/sig	0.100	0.006	1.000	6.916	73.664
kernel/sig	0.100	0.007	1.000	0.000	70.501
			Grade		
radius/all	0.115	0.022	0.997	9.329	57.887
kernel/all	0.115	0.010	1.000	0.000	68.728
radius/sig	0.101	0.013	0.999	6.628	42.937
kernel/sig	0.101	0.007	1.000	0.000	68.105
			Total number of failures		
radius/all	0.106	0.016	1.000	7.733	41.563
kernel/all	0.106	0.009	1.000	0.000	65.103
radius/sig	0.101	0.010	1.000	8.759	43.545
kernel/sig	0.101	0.007	1.000	0.000	73.357
			N. of exams passed in the first year		
radius/all	0.112	0.013	1.000	9.302	64.651
kernel/all	0.112	0.011	1.000	0.000	66.770
radius/sig	0.099	0.009	1.000	12.356	50.880
kernel/sig	0.099	0.007	1.000	0.000	71.344

Note. See Table 13.

Appendix C: Analysis restricted to students passing exams during the first year

Table 19: **PSM - Exam passed during the first year - Mathematics**

methods/ variables	ATT		std.err $^{(a)}$	treated avg$^{(b)}$	untr. avg$^{(c)}$	n. obs.
			Workload			
radius/all	-33.3		23.8	251	285	440
kernel/all	-37.4	*	19.3	240	277	440
radius/sig	-22.9		23.7	234	257	444
kernel/sig	-38.2	*	20.1	239	277	444
			Grade			
radius/all	0.93		0.58	25.25	24.31	439
kernel/all	0.64		0.41	25.06	24.41	439
radius/sig	0.81		0.56	25.07	24.26	443
kernel/sig	0.65	*	0.37	25.04	24.39	443
			Probability of one or more failures			
radius/all	-0.10		0.08	0.19	0.29	444
kernel/all	-0.06		0.07	0.20	0.26	444
radius/sig	-0.06		0.08	0.20	0.26	448
kernel/sig	-0.07		0.07	0.20	0.27	448

Note. See Table 6. Analysis limited to students who passed the exam during the first year. PSM diagnostics are available from the authors upon request.

Table 20: **PSM - Exam passed during the first year - Private Law**

methods/ variables	ATT		std.err $^{(a)}$	treated avg$^{(b)}$	untr. avg$^{(c)}$	n. obs.
			Workload			
radius/all	-79.8	***	29.6	240	320	377
kernel/all	-67.9	***	20.6	241	309	377
radius/sig	-79.1	***	25.2	247	326	381
kernel/sig	-73.1	***	17.3	241	315	381
			Grade			
radius/all	0.45		0.61	25.94	25.49	377
kernel/all	0.28		0.44	25.88	25.60	377
radius/sig	0.34		0.58	25.87	25.53	381
kernel/sig	0.57		0.44	25.88	25.31	381
			Probability of one or more failures			
radius/all	0.02		0.04	0.08	0.06	383
kernel/all	0.01		0.03	0.06	0.05	383
radius/sig	-0.05		0.05	0.06	0.11	387
kernel/sig	-0.02		0.03	0.06	0.08	387

Note. See Table 6. Analysis limited to students who passed the exam during the first year. PSM diagnostics are available from the authors upon request.

Table 21: **PSM - Exam passed during the first year - Economics**

methods/ variables	ATT		std.err $^{(a)}$	treated avg$^{(b)}$	untr. avg$^{(c)}$	n. obs.
			Workload			
radius/all	-23.5		27.6	251	275	328
kernel/all	-25.8		19.6	246	272	328
radius/sig	-37.6		23.8	247	284	336
kernel/sig	-31.5	*	16.7	246	278	336
			Grade			
radius/all	0.68		0.69	26.28	25.60	333
kernel/all	0.46		0.50	26.34	25.88	333
radius/sig	0.43		0.62	26.53	26.10	341
kernel/sig	0.32		0.44	26.37	26.06	341
			Probability of one or more failures			
radius/all	0.01		0.09	0.25	0.24	333
kernel/all	-0.05		0.07	0.24	0.29	333
radius/sig	-0.03		0.08	0.24	0.27	341
kernel/sig	-0.04		0.06	0.24	0.28	341

Note. See Table 6. Analysis limited to students who passed the exam during the first year. PSM diagnostics are available from the authors upon request.

Table 22: **PSM - Exam passed during the first year - Accounting**

methods/ variables	ATT		std.err $^{(a)}$	treated avg$^{(b)}$	untr. avg$^{(c)}$	n. obs.
			Workload			
radius/all	3.8		13.3	212	209	526
kernel/all	8.4		9.9	213	205	526
radius/sig	8.2		12.0	212	204	532
kernel/sig	10.1		9.6	214	204	532
			Grade			
radius/all	0.59	*	0.31	27.50	26.91	527
kernel/all	0.49	**	0.23	27.47	26.98	527
radius/sig	0.52	*	0.29	27.45	26.93	533
kernel/sig	0.51	**	0.23	27.47	26.96	533
			Probability of one or more failures			
radius/all	-0.03		0.03	0.05	0.08	532
kernel/all	-0.02		0.03	0.05	0.07	532
radius/sig	-0.03		0.03	0.05	0.08	538
kernel/sig	-0.01		0.02	0.05	0.06	538

Note. See Table 6. Analysis limited to students who passed the exam during the first year. PSM diagnostics are available from the authors upon request.

Table 23: **PSM - Exam passed during the first year - Economic History**

methods/ variables	ATT		std.err $^{(a)}$	treated avg$^{(b)}$	untr. avg$^{(c)}$	n. obs.
			Workload			
radius/all	4.3		9.2	175	170	600
kernel/all	1.6		7.7	172	171	600
radius/sig	0.2		9.3	170	170	609
kernel/sig	-1.2		7.4	172	173	609
			Grade			
radius/all	0.65	*	0.37	26.48	25.83	604
kernel/all	0.66	**	0.31	26.49	25.84	604
radius/sig	0.48		0.33	26.49	26.01	613
kernel/sig	0.57	**	0.27	26.50	25.94	613
			Probability of one or more failures			
radius/all	-0.00		0.02	0.02	0.02	612
kernel/all	0.00		0.01	0.02	0.02	612
radius/sig	-0.01		0.02	0.02	0.03	621
kernel/sig	0.00		0.01	0.02	0.02	621

Note. See Table 6. Analysis limited to students who passed the exam during the first year. PSM diagnostics are available from the authors upon request.

References

Becker, G. (1964), *Human Capital.* New York: Columbia University Press.

Black, D. A., and Smith, J. A. (2004), 'How robust is the evidence on the effects of college quality? Evidence from matching', *Journal of Econometrics*, 121: 99-124.

Blundell, R., Dearden, L., and Sianesi, B. (2003), 'Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey', *Journal of The Royal Statistical Society, Series A*, 168: 473-512.

Boero, G., Laureti, T. and Naylor, R. (2005), 'An econometric analysis of student withdrawal and progression in post-reform Italian universities', CRENOS Working Paper no. 2005-04.

Broccolini, C. (2005), 'Una prima valutazione degli effetti della riforma universitaria: il caso dell'Università Politecnica delle Marche', Università Politecnica delle Marche, Dipartimento di Economia Working Paper no. 244.

Broccolini, C., and Staffolani, S. (2005), 'Riforma universitaria e performance accademica: un'analisi empirica presso la Facoltà di Economia dell'Università Politecnica delle Marche', Università Politecnica delle Marche, Dipartimento di Economia Working Paper no. 242.

Caliendo, M., and Kopeinig, S. (2005), 'Some practical guidance for the implementation of propensity score matching', IZA DP no. 1588. Bonn: IZA.

DiNardo, J., and Tobias, J. (2001), 'Nonparametric density and regression estimation', *Journal of Economic Perspectives*, 15: 11-28.

DiPrete, T. A., and Gangl, M. (2004), 'Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments', Wissenschaftszentrum Berlin für Sozialforschung (WZB), SP I 2004-101. Berlin: WZB.

Härdle, W. K. (1991), *Smoothing techniques with implementation.* New York: Springer Verlag.

Miur Cnvsu (2005), *Sesto rapporto sullo stato del sistema universitario.* Roma: Miur.

OECD (2002), *Education at a glance.* Paris: OECD.

Rosenbaum, P. R. (2002), *Observational studies*, Second Edition. New York: Springer.

Rosenbaum, P. R., and Rubin, D. B. (1985), 'Constructing a control group using multivariate matched sampling methods that incorporate the propensity score'. *The American Statistician*, 39: 33-38.

Spence, M. (1973), 'Job market signaling', *Quarterly Journal of Economics*, 87: 355-374.