The effects of higher education expansion in Italy

Veruska Oppedisano
University of Turin, DSE
(preliminary draft)

Abstract

Over the period 1995-1998, Italy experienced an expansion of its higher education supply motivated by the need of rebalancing universities premises across the territory with the aim of reducing regional differences in educational attainment. This paper evaluates the effects of this policy by combining differences across regions in the number of universities sites constructed with differences across cohort of secondary school graduates. A sequential model of educational choices with uncertainty is derived and estimated. Findings suggest that enrolment rose, particularly among middle ability individuals from less favorable background, together with the probability of being retained in the university system. Students' mobility across regions decreased. The decline in passed exams specially experienced by Southern regions raises doubts on the policy effectiveness in reducing regional disparities.

JEL Codes: I2

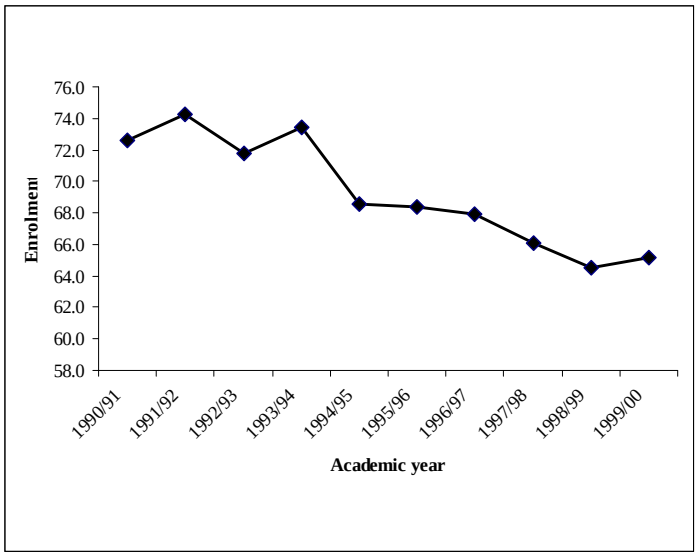
Keywords: Higher education, Italy, school choices, uncertainty

Email: veruska.oppedisano@unito.it. I am grateful to my thesis supervisor, Pietro Garibaldi, and to Giuseppe Bertola for invaluable help and assistance over the last few years. I also thank Claudio Michelacci for comments and discussions during my visit to Cemfi. I acknowledge helpful comments by Massimiliano Bratti, Christopher Flinn and Giovanni Mastrobuoni as well as partecipants in the seminar at Cemfi.

1 Introduction

Higher education is a central political and social issue in developed countries. At the beginning of the Nineties Italian higher education featured several failures, some of which still characterize the university system. Despite the fact that enrolment rate was in line with that of other Oecd countries (roughly 40 percent of young at the age of 19 was entering university), it was showing a sharp declining trend (see Figure 1[1]). Considering permanent failures, enrolment in Italy does not always lead to successful graduation: in fact, 25 percent of students drop out before entering the second year and 10 percent do not enter the third year[2]. Overall, only 30 percent of the students who enrolled actually attain a degree[3]. Among those, the majority complete their degree well beyond the envisaged time.

The range of enrolment rate, drop out and delayed graduation are especially very wide across macro area: for instance, in 1994 the enrolment rate in the North was as high as 67 percent, whilst in Southern regions it was nearly 10 percentage points lower, at 58 percent; the same range can be found by observing the drop out rate[4], which was close to 65 percent in Northern regions, and 76 percent in Southern ones. Differences were less sharp when delayed graduation is considered, as it reached 87 percent in the North and 89 percent in the South[5].



[1]The enrolment rate is measured as the percentage of students enrolled over the number of students graduated in that year from secondary school.

[2]Source: MIUR, (1998) "The evolution of higher education demand: students, graduates and equivalent students".

[3]This percentage rose to 42% at the end of the Ninenties (Education at a glance, 2000).

[4]Successful graduation is computed as the ratio between number of graduates in a given year over the number of students enrolled in the same year. Drop out is computed as the complement of successufull graduation.

[5]Source: ISTAT (1994) "Indagine sull'istruzione universitaria".

Enrolmet rate into Higher Education in Italy: 1990-1999

From the beginning of the Nineties, a supply side policy was implemented to tackle these issues, specially to reduce disparities between the North and South of Italy. Following the principles of decentralization and autonomy, this intervention resulted in a widespread increase in local institutions, homogeneously scattered across the territory, and in an expansion of established universities. Thus, the expansion encouraged the creation of local institutions, still under a regime of central approval, whilst old ones boosted the range of degrees offered.

This paper evaluates the effects of the Higher education expansion, which took place over the period 1995-1998, after which some regions increased their university sites, while others maintained the same sites provision.

The policy was explicitly justified by the need to balance higher education supply across the national territory and to decongest overcrowded universities, but instructions were not so clear about the allocation rule. After some worth discussions on the policy rule, a difference in differences estimation strategy can be implemented, where the treatment group is composed of regions directly affected by the higher education expansion, whilst the control group is represented by other Italian ones.

Higher education in Italy resisted stratification along standard dimensions, such as the public-private or the academic-vocational ones. This policy is interesting because, contrary to these standard cases, the expansion of the university system occurred by encouraging local institutions. Evaluating the effects of this program is helpful for two reasons. Firstly, it is useful to identify drawbacks and advantages of this specific higher education market structure, particularly as to how it affected some of the failures listed at the beginning of this section. Secondly, it allows the role of supply factors in shaping demand for higher education to be measured.

To better understand the channel through which the expansion affected schooling decisions, I have developed a sequential model of educational choices, within a framework of uncertainty over individual ability[6]. The main effect of having access to nearby higher education is to reduce the investment cost of entering university. There are three routes through which this cost reduction is likely to have had an impact in the short term period. Firstly, it should have loosened credit constraints and weakened the role of parental cultural capital on the decision to enter higher education. Above all, the investment is expected to have become more appealing for more groups in the society and educational opportunities more equalized across different social backgrounds. Secondly, widespread local access to university should have reduced students' mobility across regions: if university access is locally available and the title issued has the same legal value as that

[6]The first chapter of my Ph.D. thesis shows that educational choices in Italy (enrolment, drop out and delayed graduation) might be rationalized both within a theory of signalling and one of uncertainty. In both cases, heterogeneous academic performances correspond to different ability levels, but under the first framework students perfectly known their talent, whilst under the second one they learn it throughout university experimentation. Here I choose to develop the model under the uncertainty framework: the lack of information on graduate students' labour market outcomes would make cumbersome to test a general equilibrium theory of signalling.

issued by another institution, incentives to enter university in a region different from the one of residence should strikingly decline. Finally, local university might have exerted a negative impact on aggregate regional academic performance, measured in terms of drop out probability and number of exams passed. As the reduced cost of higher education increased enrolment, the average ability of the students' body should have declined in regions affected by the policy. Marginal students might have been more prone to drop out from university unless the lower monetary costs allowed them to continue it. If the overall drop out declined after the policy implementation performance, that should reflect students' talent, might have worsened.

The data originates from surveys of youth interviewed three years after graduating from secondary school, carried out by the National Institute of Statistics (ISTAT) in 1998 and 2001, merged with regional level data on higher education supply. Outcomes of interest are evaluated by means of a difference in differences estimation strategy.

Results show that the higher education supply expansion increased university enrolment, especially among marginal individuals, who are those with middle schooling ability and have less favorable family background, without increasing their probability to quit university. This first set of results shows that the expansion went in the direction of reducing the impact of family background on the decision to enrol university. Students mobility across regions significantly decreased, especially among more talented individuals from less affluent families. A last piece of evidence indicates that academic performance, measured by the number of exams passed three years after secondary school graduation, worsened, specially in Southern regions, without being explained by a more intense students' labour market participation. As a consequence, the policy spurred prolonged permanence into university and delayed the young' entry in the labour market.

The remainder of the paper is organized as follows: section II briefly reviews economic literature on the supply expansion of education. Section III describes the implementation of Italian policy data and presents the identification strategy. Section IV displays the conceptual framework using a sequential model of schooling choices. Section V is devoted to the estimation of the effects of the policy on a set of outcome indicators. Section VI performs some robustness checks and section VII concludes the paper.

2 Literature

Economic research on the effects of political changes, such as the expansion of education supply, is modest and quite recent. Indeed, many works have been focused on the impact of demand components on educational choices, while little attention has been placed on the role played by supply factors.

Duflo, E. (2001) evaluates the effects of a major primary school construction program launched in Indonesia between 1973 and 1978 using a difference in differences methodology. She exploits the fact that the exposure to school construction varied exogenously by date and place of birth to estimate the impact of the expansion on years of education and using the program as an instrument,

on earnings. Results point towards a significant increase in the proportion of the population with more years of primary education. The estimates of the return to schooling in this developing country corrected for endogeneity range from 6.8 to 10.6 percent.

Using a similar estimation strategy, Berlinski S. and Galiani S. (2007) assess the impact of a large construction pre-primary school facilities program in Argentina on pre-primary school enrolment and maternal labour supply. Their identification strategy relies as well on the heterogeneous intensity of program exposition across provinces and cohorts induced by the timing of the policy. They find that the construction program had a sizeable impact on pre-primary school enrolment among children aged 3-5 and increased maternal employment.

Few papers evaluate the effects of a tertiary education expansion program. Holzer, S. (2006) investigates the impacts of the expansion of higher education supply in Sweden, where at the end of the 1970s new regional university colleges had been established. At the beginning these colleges provided vocational and shorter programs, without conducting any research activities, meanwhile in the 1990s the status of these colleges was upgraded and they were permitted to carry out research more independently than before. The author focuses on two outcomes: the extent to which the expansion increased equality of opportunities in entering higher education and its effects on the allocation of students between the new shorter, vocational tertiary educational tracks and the old universities. Area of residence and cohort of birth jointly determine exposure to the program. Results provide some supports that social mobility increased and any evidence that students from lower socioeconomic backgrounds have been diverted from attending old universities to enrol at new university colleges.

For the Italian experience, Bratti, M. Checchi, D. and De Blasio, G. (2008) study the effects of the expansion of universities and faculties provision in Italy over the decade 1990-2000. Using data from the Bank of Italy Survey of Households Income and Wealth, they focus on the likelihood of holding a university degree and of being a university student. Exposure to the program depends on cohort of birth and actual region of residence, with the assumption that current region of residence coincides with the region of residence at age 19. Results show robust evidence of a positive effect of higher education expansion on student's enrolment and retention, but a non significant impact on the probability of graduation.

This paper complementary investigates the same issue, by using another database and highlighting different aspects related to the decision of university enrolment. Firstly, the time period of supply expansion under study is restricted to the 1995-1998 interval, over which some Italian regions increased the higher education provision, measured in terms of university sites, while in others existing supply remained unchanged. Since the time period over which I focus my attention is short, the regional variability on the intensity of university supplied is quite low. Thereby, differently from Bratti et al. who have analyzed the average effects across regions, I consider as treated regions those who expanded the number of sites in a significant amount. Most likely, this restriction has impact on the magnitude of the effects I estimate.

Secondly, information originates from the "Survey on the study and work experiences of sec-

ondary school graduates", a cross sectional sample of school leavers, interviewed by the national institute for statistics (ISTAT) three years after secondary school graduation, in 1998 and 2001. The vast wealth of information found concerning the region of residence at the age of 19 and the year of enrolment into university, allows me to make the identification assumption more precise. Lastly, my focus, complementary to Bratti et al.'s one, spreads some light on the channels through which this policy that encouraged local institutions, worked. The next section shows an overview of Italian's higher education policy since the beginning of the Nineties.

3 The program

3.1 Institutional framework and program information

During the Nineties the Italian higher educational system featured a striking expansion of its supply. This expansion has been driven by two broad rationales. Firstly, the necessity to spread the accessibility to university homogeneously across the territory in order to increase equality of opportunities in human capital investments. Secondly, the need to decongest overcrowded universities, which exceeded forty thousands students enrolled. This section is organized as follows: at first, a historical insight of Italian universities supply and the institutional framework within which it was organized; next, the legislation that regulated the higher education expansion; finally a focus on the specific expansion that took place over the period 1995-98.

Since the Italian constitution was established, the University system has traditionally been organized by central approval. The opening of a new university was subject to government authorization and the introduction of a new degree, selected among the admissible ones, involved the fulfilment of requirements of a specific law. This centrally organized procedure has always guaranteed that titles of higher education attainment were legally valid throughout Italy independently on the institution which issued them.

The list of state and free university sites, drew up under the Fascist regime within a unified framework law, remained substantially unchanged until the end of the Seventies[7]. At the end of the Seventies higher education admission was extended to all secondary school graduates; before that, students from technical or professional tracks were only allowed to sustain degrees which were relative to their course of earlier educational career. This new mass of potential entrants forced the legislator to approve the opening of new universities[8].

A substantial process of university reform started up at the beginning of the Eighties (law n. 382/1980). This law, besides new rules related to hiring and management procedures and teaching

[7]This included: Turin, Milan, Padoa, Pavia, Venice, Trieste, Genoa, Parma, Modena, Bologna, Ferrara, Pisa, Florence, Siena, Perugia, Urbino, Macerata, Camerino, Naples, Bari, Messina, Palermo, Catania, Cagliari and Sassari.

[8]In particular: Udine in 1978; 2nd university of Rome, Viterbo and Cassino in 1979; Potenza in 1981; l'Aquila, Chieti, Pescara, Brescia, Campobasso, Reggio Calabria, Verona and Trento in 1982.

loads, stated the necessity to include any variation in universities' provision in a development plan, but without explicitly defining standards and evaluation methods. The general concern was to balance funds allocation and university sites nation wide. According to the law, the scheme had to be put forward and approved by the Ministry of Education every four years. Within this framework the plan for the period 1986-90 established the institution of the Second University in Naples and the Polytechnic at Bari, both located in regions endowed with overcrowded universities (Campania and Puglia); also, the set up of other fourteen separate sites of existing universities: four in Emilia Romagna, three in Piemonte, two in Lombardia and Puglia and one in Veneto, Lazio and Campania. However, even though the plan had been drawn up, no resources were assigned for its realization.

Only with the Law 245/1990 new procedure and rules for the plan approval were established and financial resources have been made available. The plans' objectives were clearly stated. These were "..To ensure a balanced development and adjustment of higher education provision keeping into account local potential demand, big metropolitan areas, gaps between North and South and national instructive needs"[9]. The procedures for the plans design were slightly changed: universities gained autonomy in advancing proposals, to be presented one year before the current plan expiration. After long consultations with regional, parliamentary and university councils, plans had to be approved by central government. The law authorized the institution of new separate sites of existing universities and private institutions, suppressing the requirement of a specific parliamentary law; instructions for eligibility required financial resources, staff and available structures in order to ensure the correct functioning.

The amount of resources allocated to this program was around 1.5 thousand million of euros[10] in six years. To ensure a balanced site provision development, 40% of resources had to be assigned to Southern regions. The distribution of funds depended upon the type of structure (either a new site, a different faculty or a degree course) and the new subjects, in such a way that made the creation of scientific courses more worthwhile. The plans also specified the schedule for hiring new professors, programmed every two years. However, as competitions were never be complied within the fixed time, plans' actuation was consequently delayed.

The first plan under the new legislation was approved in 1991, covered the schedule for the period 1991-93 and provided the realization of the plan for the period 1986-90, for which resources had not previously been assigned. In addition, it established the institution of the Third University in Rome, a new university in Abruzzo (at Teramo) and four separate sites in Marche (at Ascoli Piceno), in Molise (at Isernia), in Basilicata (at Matera) and in Sicilia (at Trapani). To decongest overcrowded universities, the plan established the institution of a second university in Campania, university poles in Milan area and subsidiaries of Bologna's university at Rimini, Cesena, Ravenna, Forlì and Reggio Emilia. The plan for the period 1994-96 included the institution of a separate site in Piemonte (at Alessandria, Novara and Vercelli), in Lombardia (at Varese), in Campania (at

[9] Law 245/1990, art. 1 (a), my translation.

[10] Amount in euros, indexed at prices 2004.

Benevento) and in Calabria (at Catanzaro).

Over the period 1995-1998, which I focus on, in application to the forecasted provisions, the creation of the second university of Milan was authorized as well as new ones like Bicocca, Insubria at Varese and the separate site of the University of Bari, at Taranto. Similarly, University of Bologna opened subsidiary sites at Forlì, Ravenna and Reggio Emilia and University of Catania set up a new site for the faculty of Architecture at Siracusa. Universities in Molise (at Ascoli Piceno) and Marche (at Isernia) were effectively instituted over this time period.

Along with these initiatives private universities were also founded. In Trentino Alto Adige and Val D'Aosta, local authorities set up new universities, while in Lombardia and Lazio private enterprises were allowed to found respectively the San Raffaele and San Pio V. During the period from 1995 to 1998 the number of university sites in Italy increased from 69 to a total of 80.

Although the allocation rule was not very explicit, stated objectives were clear (art. 1, Law 245/1990). I checked whether the actual allocation rule decided upon by universities and Ministry of Education achieved planned objectives listed in that article. I used the Log of secondary school graduates in 1992 at regional level to get a proxy for potential demand at the beginning of the Nineties[11]. To control for metropolitan area, I built a dummy equal to one when the region was endowed with an overcrowded university and zero otherwise. Territorial disparities are controlled for by the Log of the professors-students ratio computed at regional level[12].

Table 1 presents results from a regression of the logarithm of the number of new university sites constructed in each region between 1991 and 1998 on the proxies. Since besides sites' expansion, all universities increased the variety of available degrees, by either creating new faculties or by offering new courses within existing faculties, I also controlled for the increase in the number of degrees supplied between 1991 and 1998. To avoid collinearity, the difference in degrees is computed by subtracting courses supplied in new sites. A substantial part of regional variation can be explained by these variables: the R-squared is 0.57. Coefficients have the expected sign: potential demand and overcrowded universities positively affect sites institution; whilst a proportionate teaching staff is negatively correlated with the expansion. The logarithm of the number of degrees created between 1991 and 1998 is negatively and significantly correlated with the sites' expansion. This may be explained by the substitutability of the two interventions. Indeed, over the period 1995-1998 all regions increased the supply of degrees within existing or new faculties. As the allocation of new sites could be correlated with the expansion of degrees supply, I will present specifications which also control for this parallel program.

The regression does not completely explain the rule as the allocation of funds was not tied to explicit and clear criteria. As it has been assessed by the *Ministero dell'Università e della Ricerca Scientifica*:".. With respect to the development and rebalancing of university premises prevailed – at least for the most part - an non selective "all over the place" approach, inspired by a barely

[11]Source: ISTAT. Data on secondary school leavers have not been collected at regional level before 1992.

[12]Source: MIUR (1997) "Allegato E - Verifica dei Piani di sviluppo dell'università 1986-1990 e 1991-1993".

incremental purpose..."[13].

One last point worth mentioning concerns the possibility that local labour market conditions, such as a high youth unemployment rate, might have influenced the policy rule, if it was a compensatory intervention, thus bringing more resources to more depressed economic areas. Since labour market conditions affect the opportunity cost of schooling, it could also have played a role in shaping demand for higher education. To assess the validity of the identification assumption, I also need to control for this factor. The second column of Table 1 presents the specification that includes the level of regional unemployment rate of young people in the age bracket 15-24 in 1990: it turns out that it negatively affected the program. The sign is not coherent with the argument presented, however the coefficient is not statistically significant.

Allocation of new sites		
	Log New sites	Log New sites
Log of secondary school leavers, 1992	0,344	0,397
	[1,48]	[1,67]
Overcrowded university	0,435	0,394
	[1,37]	[1,23]
Log of prof. for 100 students, 1990	-0,368	-0,302
	[1,47]	[1,23]
Log of new degrees, 1991-1998	-0,325	-0,346
	[2,52]**	[2,71]**
Youth unemployment rate, 1990		-0,756
		[1,09]
Observations	20	20
R-squared	0,57	0,61

Note: Robust t statistics in brackets. ** significant at 5 percent

Table 1: The allocation of sites

3.2 Data

I use data collected from the "Survey on school and work experiences of secondary school leavers", a cross-sectional sample of school graduates interviewed three years after graduation. The data contains a wide range of information on the school curriculum and on the post-school experiences, either in college and in the labour market. Moreover, information on personal characteristics, family background, region of residence during secondary school, region of university attendance and year of enrolment is available.

The Italian school system of secondary education is structured into tracks that can be divided between college oriented ("Licei") and labour market oriented ("Istituti tecnici"- technical schools- and "Istituti professionali" - professional schools-), which represent the largest share of supply of

[13]MIUR, (1997) "Verifica dei piani di sviluppo dell'università 1986-90 e 1991-93", doc.4/97, p. 10.

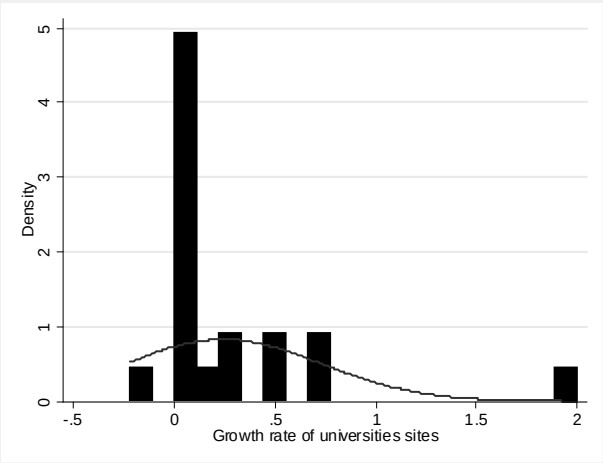


Figure 1: Distribution of growth rate of university sites between 1995-1998

secondary education. A minor share, between 7 and 8 percent, is composed of schools intended for individuals aiming at artistic professions. Given the specificity of this minor track, students coming from other secondary schools have been excluded from the estimation sample[14].

Excluding these observations yielded a final estimation sample of 36.629 observations, 16.535 of which belong to the 1995 cohort and the rest to the 1998 one.

The 1998 and 2001 repeated cross sections of individual data are pooled and information matched with regional-level data on the intensity of the higher education supply, both in 1995 and 1998 (years in which secondary school leavers were interviewed). Information over the regional supply of higher education, including sites, faculties and degrees courses, are taken from the annual ISTAT report "Statistics of Higher Education".

The measure for the intensity of the higher education supply at regional level is the percentage increase in the number of university locations between 1995 and 1998. Figure 1 shows the distribution of growth rate of sites supplied across regions. The growth rate, or intensity of the program P_j, has been compiled according to the formula:

$$P_j = \frac{2 * (s_{j,98} - s_{j,95})}{(s_{j,98} + s_{j,95})} \tag{1}$$

where $s_{j,year}$ identifies the number of sites in a given region and year. This formula has been used in order to account for the Val D'Aosta zero sites supply starting level.

There was an expansion in eight regions throughout the institution of new sites, the number of sites in the Veneto region was even reduced because of the closure of a separated site of the

[14] As a robustness check I have also estimated output equations including this sub sample of students. Results are very similar to those reported.

"Iulm" private university located in Milan, whilst eleven remaining regions maintained the same number of universities as forecasted. I defined as high intensity regions those where universities sites increased more than 30 percent, while low program regions those where the expansion was lower than 30 percent. As Figure 1 indicates, the majority of regions are located below this cut off point, whilst the others are well above.

Table 2 indicates the program intensity of treated and control regions. Regions in the treatment group are made up of Val D'Aosta, Trentino Alto Adige, Emilia Romagna, Molise and Puglia.

Program level	Mean	Min	Max	N
Low level regions	0,034	-0,222	0,286	15
High level regions	0,876	0,500	2,000	5

Note: high level regions are those were the expansion has been >30 percent

Table 2: Growth rate of universities sites between 1995-1998

Tables 3 and 4 show baseline summary statistics for treatment and control groups in terms of individual level data and regional level ones before the policy was implemented to assess whether the program created comparable groups. In Table 3, the first column shows statistics of individuals resident at the end of secondary schools in regions that kept locations' supply unchanged or increased it by less than 30 percent, while the second column statistics of individuals living in regions where sites' provision increased more than 30 percent. The third column presents differences of means between treatment and control groups and standard errors. There are no statistically significant differences between individual observable characteristics at treatment versus comparison groups, except for the mobility rate, higher in treated regions.

To provide a sense of the study population, its composition is described briefly. Slightly more than half of respondents are female; average respondent's age is about 22.7; roughly two thirds are composed of children whose parents have primary education, one third of children whose parents obtained a secondary school diploma, and less than 10 percent have college graduated parents. Marks at the end of lower secondary school are almost equally distributed among all respondents, with a prevalence of individuals who obtained a low D mark. Almost 27 percent of the youth attained the diploma in a high school, 50 percent in a technical secondary school and 22 percent in a vocational one. College enrolment and drop out were slightly lower among treated individuals and the number of exams passed were slightly higher.

Turning to regional level data (Table 4), it appears that there was a significant difference in the total number of university sites between control and treated regions, as the former were endowed with on average six sites per region and the latter with less than three. However, as I discussed in Section 3.1, among regions that boosted universities' provision between 1995 and 1998 it was the smallest ones, which needed less sites to satisfy local demand; indeed, by dividing the number of

regional sites by the population of secondary school graduates in that region, differences became less sharp, but still show that sites' provision increased more in regions less initially endowed, in accordance with legislator's objectives.

The fact that most of observable individual characteristics are similar across the treatment and comparison group provides evidence that the policy was somewhat randomized. To be more convincing, the next section discusses in detail the assumptions of the identification strategy.

3.3 Identification strategy

An individual's exposure to the program is jointly determined by his year of graduation from secondary school and his region of secondary school attendance. The young people who left secondary school in 1995 did not benefit from the program, since the higher education expansion only came into force between 1996 and 1998, whilst individuals who terminated secondary school in 1998 were fully exposed.

Delayed university enrolment could lead to mistakenly consider as non treated students that took advantage from the program; in that case, estimates would suffer a downward bias. Similarly, the intensity of the program exposure would be erroneously assigned to those who graduated in 1998 and delayed enrolment, since higher education expansion was kept-on after 1998. In that case, the effect of the policy would be overestimated. To avoid these problems, I drop from the pooled sample individuals who entered higher education in different years from 1995 and 1998.

A second source of variation arises from different intensity in higher education supply expansion across regions. Data contain information on both the region of secondary school and the region where university was attended. These two variables are highly correlated: 93.4 percent of students attended a university placed in the same region where they completed secondary school. Evaluating the enrolment decision and the academic performance according to the supply of higher education in the region of secondary schools could downward bias the coefficient estimate of the policy because migration introduces measurement errors. On the contrary, assessing enrolment and performance according to the supply of higher education in the region where university was attended could give positively biased estimates because of endogenous selection. Indeed, individuals living in areas with a low supply of universities are attracted to regions better endowed, leading to overestimate the impact of the program on enrolment (see Rosenzweig and Wolpin, 1988b). The problem is tackled by separately evaluating the two choices, that of enrolling and that of attending a university placed in a region different from the one of residence, on the basis of the exogenous supply of higher education in the region of secondary school. In this way, endogeneity issues are ruled out.

Identification of parameters of interest relies on the differential intensity of the program expansion across regions and differences in exposure across cohorts of graduates induced by the timing of the expansion.

The basic idea behind the identification strategy can be illustrated using a simple two-by-two table. Table 5 shows differences of outcomes' means, computed at regional level, between 1995 and 1998 by program level for each outcome of interest. It presents the main experiment providing an

	Comparison	Treatment	Comparison-Treatment
	Mean (s.d.)	Mean (s.d.)	Mean (s.e.)
Female	0,55	0,54	0,01
	(0,50)	(0,50)	(0,01)
Age	22,77	22,67	0,01
	(2,52)	(2,25)	(0,05)
Father college degree	0,09	0,08	0,01
	(0,28)	(0,27)	(0,01)
Mother college degree	0,07	0,07	0,00
	(0,26)	(0,27)	(0,00)
Father secondary degree	0,30	0,31	-0,01
	(0,46)	(0,46)	(0,01)
Mother secondary degree	0,27	0,28	-0,01
	(0,44)	(0,45)	(0,01)
Father compulsory degree or lower	0,61	0,61	0,00
	(0,49)	(0,49)	(0,01)
Mother compulsory degree or lower	0,66	0,65	0,01
	(0,47)	(0,48)	(0,01)
Junior school mark A	0,18	0,18	0,00
	(0,39)	(0,39)	(0,01)
Junior school mark B	0,19	0,21	-0,01
	(0,39)	(0,40)	(0,01)
Junior school mark C	0,28	0,29	-0,01
	(0,45)	(0,45)	(0,01)
Junior school mark D	0,35	0,32	0,03
	(0,48)	(0,47)	(0,01)
High school	0,27	0,26	0,01
	(0,45)	(0,44)	(0,01)
Technical secondary school	0,52	0,50	0,02
	(0,50)	(0,50)	0,01
Professional secondary school	0,21	0,24	-0,04
	(0,40)	(0,43)	(0,01)
College enrolment	0,50	0,48	0,02
	(0,50)	(0,50)	(0,01)
College drop out	0,18	0,16	0,02
	(0,38)	(0,37)	(0,01)
Mobility	0,13	0,22	-0,10
	(0,33)	(0,42)	(0,01)
Number of passed exams	8,59	9,00	-0,41
	(5,18)	(4,97)	(0,16)

Table 3: Baseline descriptive statistics (1995 survey) - Individual level data (N=16,535)

	Comparison	Treatment	Comparison-Treatment
	Mean (s.d.)	Mean (s.d.)	Mean (s.e.)
N. of sites	6,02	2,67	3,34
	(4,17)	(1,34)	(0,04)
N. of sites for 1000 sec. school leavers	0,17	0,13	0,04
	(0,07)	(0,08)	(0,00)

Table 4: Baseline descriptive statistics (1995 survey) - Regional level data (N=20)

illustration of the identification strategy. A list of outcomes of individuals who had little or no exposure to the program (those resident in low program regions at the end of secondary school in 1998) are compared to those of individuals who were exposed to the program (those resident in high program regions at the end of secondary school in 1998).

Outcomes of interest are the following: enrolment, a dummy equal to one when the individual entered higher education and zero otherwise; drop out, a dummy which takes value of one when the individual quitted university; mobility, which equals one when the student entered university in a region different from the one of residence at 19 years old and zero otherwise; finally, the average number of exams passed during the first three years of academic studies as an indicator of individual performance[15].

The first block in Table 5 presents the change in enrolment over the period for the two groups of regions. As already documented by Figure 1, in both groups the average enrolment dropped over the years. However, it decreased less in regions that set up more universities. Considering changes in withdrawal behavior for the population of students that entered higher education, it emerges that drop out diminished in all regions, but less in treated ones. Mobility decreased as well, but more in regions affected intensively by the policy. The number of exams passed fell in both groups of regions, but more where new universities have been opened[16].

The difference in the differences between treated and control groups can be interpreted as the causal effect of the policy, under the assumption that in the absence of the higher education expansion, the trend of the variables of interest would have not been systematically different between high and low program regions. These simple differences suggest that higher education expansion led to increase enrolment, sudden decrease mobility rates across Italian regions, increase drop out and, interestingly, caused a sharp reduction in the number of exams passed.

[15]Unfortunately, information on grades scored at university are not available. However, I argue that this is a minor problem as recent literature have found that, in Italy, faculties strategically adjust grading standards to affect enrolment (see Bagues, Sylos Labini and Zinovyeva, 2006).

[16]I computed the same differences using others cut off points, at 25% and 50%, obtaining, qualitatively, the same results.

	Program level		
	High	Low	Difference
Enrolment			
Graduated in 1998	0,402	0,353	0,049
	(0,069)	(0,044)	(0,033)
Graduated in 1995	0,465	0,49	-0,025
	(0,038)	(0,034)	(0,019)
Difference	-0,063	-0,137	**0,074**
	(0,011)	(0,039)	(0,041)
Drop out			
Graduated in 1998	0,117	0,135	-0,018
	(0,018)	(0,026)	(0,010)
Graduated in 1995	0,145	0,181	-0,036
	(0,042)	(0,031)	(0,020)
Difference	-0,028	-0,046	**0,018**
	(0,017)	(0,010)	(0,020)
Mobility			
Graduated in 1998	0,312	0,171	0,141
	(0,283)	(0,143)	(0,132)
Graduated in 1995	0,365	0,184	0,181
	(0,293)	(0,173)	(0,138)
Difference	-0,053	-0,013	**-0,040**
	(0,024)	(0,016)	(0,029)
Number of passed exams			
Graduated in 1998	9,087	8,388	0,699
	(1,178)	(0,907)	(0,577)
Graduated in 1995	9,297	8,455	0,842
	(1,012)	(1,075)	(0,531)
Difference	-0,210	-0,067	**-0,143**
	(0,117)	(0,149)	(0,189)

Note: Means and standard errors in brackets

Table 5: Means of outcomes of interest by cohort of graduation and level of program

To rely on this identification strategy and infer a causal effect of the program on university enrolment, drop out, performance and mobility, some comments are worth mentioning.

As Rosenzweig and Wolpin (1988a) discuss, one should not underplay the possibility of a compensatory intervention, meaning an universities allocation rule that geographically distributed expansion in such a way to provide less endowed regions with higher public investments. If this was the case, the effect of the program would result overestimated.

As discussed in Section 3.1, the allocation of funds to higher education investments has been driven by two broad economic rationales: the necessity of rebalancing universities premises across the national territory and that of splitting overcrowding universities. The first criterion resembled a compensatory policy intervention; the second one went in the opposite direction since overcrowded universities were principally located in regions already well endowed (Piemonte, Lombardia, Emilia Romagna, Lazio, Campania and Puglia). However, a standard way to circumvent the possibility that some pre treatment observable characteristics affected the policy rule is to condition the outcome equations to regional fixed effects dummies variables (as proposed by Rosenzweig and Wolpin, 1988a).

Another related but distinct issue regards the possibility that policy was endogenous (Besley and Case, 2000). Since the allocation rule was not explicit, it might be the case that investments in higher education infrastructures reflected some political, demographic and economic variables, that are time-variant regional specific. If these variables are also correlated with educational outcomes and are omitted from the outcome equation, the estimated effect of the policy would result biased. In this context, as shown in Section 3.1, the allocation rule was partly explained by variables related to objectives stated by the Law; however, roughly 40 percent of regional variation was not accounted for by these variables. A variable that might have driven the policy and could be correlated with outcomes is the proportion of young people unemployed within a region. A high local youth unemployment rate might have encouraged the allocation of more resources for an expansion of higher education; but also, if initial values are correlated with actual choices, it could have lowered the opportunity cost of attending university thus increasing the demand. The omission of this variable could mislead the causal interpretation of the policy of interest.

For these reasons, I add controls for the set of time varying variables that might have affected the policy to the baseline specification, as specified in Section 3.1. All these variables are interacted with the treatment status dummy. Moreover, I include the variation of the youth unemployment rate between 1995 and 1998 in order to control for possible changes in labour market opportunities that might have been correlated with educational choices.

Finally, estimates rely on the identification assumption that there is no omitted time-varying and region specific effect that might be correlated with the program. This assumption will be violated if the allocation of other programs were correlated with the establishment of new sites. As already discussed in Section 3.1, along with the new sites' set up, the Legislator spurred the expansion of existing universities by allowing the institution of new Faculties and/or new degrees courses; this policy was probably a substitute of the one under analysis. The identification

assumption might also, therefore, be satisfied, only after controlling for this second source of expansion.

4 Conceptual framework: a sequential model of schooling choices

In this section I use a simple sequential model to describe the young's decision of entering university, dropping out and their optimal human capital investment choice. I consider a risk neutral individual with uncertainty over his individual ability. The information set is described as follows. Let α_i be a measure of the unknown individual true ability. Students have prior beliefs over α_i; specifically, assume that this prior is normally distributed with mean μ_α and precision, which is the inverse of the variance, σ_α. Beliefs about α_i change as a function of an observed signal $s_{i\tau}$. The individual observes two signals, one before deciding whether to enter university or not and one after university enrolment. The first signal could be represented by the observable parents' education or the outcome from previous schooling career. The second might indicate the result of exams taken, either in terms of success or failure or of the grade scored.

More formally, following Jovanovic (1979), the signal takes the following expression:

$$s_{i\tau} = \alpha_i + \zeta_{i\tau} \tag{2}$$

where $\zeta_{i\tau}$ is noise, independent of α_i and normally distributed with mean 0 and precision σ_ζ[17]. Since both α_i and $\zeta_{i\tau}$ are normally distributed and independent, the conditional mean of α_i given $s_{i\tau}$, that is, the posterior about α_i, is a linear function of the observed signal. After the first signal has been received, beliefs are updated according to:

$$\begin{aligned} E(\alpha_i|s_{i0}) &= \mu_\alpha + \frac{cov(\alpha_i, s_{i0})}{var(s_{i0})}(s_{i0} - \mu_\alpha) \\ &= \frac{\sigma_\alpha}{\sigma_\alpha + \sigma_\zeta}\mu_\alpha + \frac{\sigma_\zeta}{\sigma_\alpha + \sigma_\zeta}s_{i0} \end{aligned} \tag{3}$$

Once entered university, students receive a second signal s_{i1} and update their beliefs as follows:

$$\begin{aligned} E(\alpha_i|s_{i0}, s_{i1}) &= E((\alpha_i|s_{i0})|(\alpha_i|s_{i0}) + \zeta_{i1}) \\ &= \frac{\sigma_\alpha}{\sigma_\alpha + 2\sigma_\zeta}\mu_\alpha + \frac{\sigma_\zeta}{\sigma_\alpha + 2\sigma_\zeta}(s_{i0} + s_{i1}) \end{aligned} \tag{4}$$

[17] To add realism to the model, σ_ζ could be let vary between the first and the second signal as they capture different dimensions. In particular, the first signal could be represented by the observable parents' education or the outcome from previous schooling career, whilst the second one could approximate the number of exams tried during university attendance to capture the idea that the more intensive is experimentation, the faster is the learning process. Here, as I am unable to separately identify the two sources of information, I consider a more general formulation.

More generally, if a new signal is received in each period τ, the posterior distribution of α will stay normal with mean and variance given by:

$$E(\alpha_i|s_{i\tau+1}) = \frac{\sigma_\alpha}{\sigma_\alpha + \tau\sigma_\zeta}\mu_\alpha + \frac{\sigma_\zeta}{\sigma_\alpha + \tau\sigma_\zeta}\sum_{\tau=0}^{\tau} s_{i\tau} \tag{5}$$

$$var(\alpha_i|s_{i\tau+1}) = \frac{1}{\sigma_\alpha + \tau\sigma_\zeta} \tag{6}$$

These expressions show that the conditional value of α follows a random walk with incremental variance that declines deterministically to zero as $\tau \to \infty$.

The timing of the model is the following. At time 0 the individual chooses whether to enrol university or enter the labour market according to the first prior about his ability. In the same period, the decision of eventually attending university in a different region is taken. At time 1 the student enrolled receives a further signal correlated with his true ability, updates his beliefs and chooses whether to continue university or to drop out. Finally, at time 2, if remaining in university, he decides the optimal amount of human capital to attain, which will be empirically proxied by the number of exams passed[18].

The utility function of individual i, born in region j, in cohort t, at time τ takes the following form:

$$U(Z,x)_{ij\tau t} = B(x)_{ij} - C(Z,x)_{ij\tau t} \tag{7}$$

where $B(x)$ is the income earned by an individual with higher education x and $C(x)$ is the cost of the human capital investment. I assume that returns to graduation for an individual i born in region j are linear and discounted by a a factor $\beta^{2-\tau} < 1$:

$$B(x)_{ij} = \beta^{2-\tau}\left(\mu_j + \rho x_i\right) \tag{8}$$

Returns to higher education are affected by regional economic conditions μ_j and the premium ρ for the observable number of exams taken x_i. The assumption of constant labour market reward to skills across regions is justified by information available in the data. Individuals are interviewed three years after secondary school completion and wages earned after university graduation are not observed. Thus, with these data, the effect of the policy can be estimated only in the short term[19].

[18] In reality students do not decide sequentially whether to drop out and then the optimal human capital investment; rather, the two decisions might occur simultaneously. This simplified assumption makes easier the analytical solution of the model, but will be relaxed in the empirical implementation.

[19] As noticed by Heckman, Lockner and Taber (1998) and Angrist (1995), policies that reduce the monetary cost of education might, as an indirect effect in the long run, change the equilibrium in the market of skills.

The cost of education $C(Z,x)$ is a function of two monetary components of schooling and a non monetary one. In particular:

$$C(Z,x)_{ij\tau t} = \delta_\tau Z_{jt} + c_j + \frac{x_i^2}{2E(\alpha_i|s_{i\tau})} \tag{9}$$

The first component is modeled as a linear function of the number of regional universities (Z_{jt}) for cohort t in region j and captures the most important modification induced by the program. An higher number of sites spread out in the region reduces travelling distances, thereby lowering the investment cost. The coefficient δ_τ is allowed to vary according to the outcome taken under consideration and is expected to have a negative sign.

The second term c_j represents other regional characteristics influencing the cost of schooling, but not affected by the program. The third one, $\frac{x_i^2}{2E(\alpha_i|s_{i\tau})}$, indicates the non monetary cost of human capital acquisition which is lower the higher the expected conditional ability $E(\alpha_i|s_{i\tau})$. The hypothesis of convex cost ensures that a given amount of investment is less costly when spread out over multiple periods than when it is concentrated in a single period.

Substituting equations (8) and (9) into equation (7) yields:

$$U_{ij\tau t} = \beta^{2-\tau}(\mu_j + \rho x_i) - \delta_\tau Z_{jt} - c_j - \frac{x_i^2}{2E(\alpha_i|s_{i\tau})} \tag{10}$$

Finally, I define the utility derived from entering the labour market as a linear function of μ_j, which summarizes labour market conditions for unskilled labour and represents the outside option of education.

$$B(L)_j = \mu_j \tag{11}$$

μ_j is taken constant over time. The assumption might not be fully realistic if labour market outcomes are characterized by some degree of uncertainty, which would call for time varying local labour market conditions. Again, since data at my disposal do not feature a time varying dimension, the assumption is coherent with information available.

The model has to be solved backward, starting with the optimal choice of exams in the last period.

4.1 Number of exams

Working backward, I look for the optimal number of exams at $\tau = 2$ for an individual who continues higher education:

$$\underset{x}{Max} U_{ijt\tau=2} = \mu_j + \rho x_i - \delta_2 Z_{jt} - c_j - \frac{x_i^2}{2E(\alpha_i|s_{i\tau})} \tag{12}$$

Simple algebra shows that the optimal number of exams is:

$$
\begin{aligned}
x^*_{ijt} &= \rho E(\alpha_i | s_{i\tau}) \\
x^*_{ijt} &= \rho \left[\frac{\sigma_\alpha}{\sigma_\alpha + 2\sigma_\zeta} \mu_\alpha + \frac{\sigma_\zeta}{\sigma_\alpha + 2\sigma_\zeta} (s_{i0} + s_{i1}) \right] \qquad (13) \\
&= p_0 + p_1 (s_{i0} + s_{i1}) \qquad (14)
\end{aligned}
$$

Where $p_0 = \rho \frac{\sigma_\alpha}{\sigma_\alpha + 2\sigma_\zeta} \mu_\alpha$ is a constant term representing the average number of exams taken by the population of enrolled students, and $p_1 = \rho \frac{\sigma_\zeta}{\sigma_\alpha + 2\sigma_\zeta}$ weights the individual deviation from the national average[20].

The optimal amount of exams is increasing in individual conditional ability and in the labour market premium to skills.

4.2 Drop out

I now proceed to characterize the optimal stopping criterion at time 1, after information on individual ability has been updated. First, I substitute the optimal number of exams x^* in equation (10) and replace $E(\alpha_i | s_{i\tau})$ with equation (4):

$$
\begin{aligned}
U_{ijt\tau=1} &= \beta \left(\mu_j + \rho^2 E(\alpha_i | s_{i\tau}) \right) - \delta_1 Z_{jt} - c_j - \frac{\rho^2 E(\alpha_i | s_{i\tau})}{2} \\
&= \beta \mu_j - \delta_1 Z_{jt} - c_j + \pi_0 + \pi_1 (s_{i0} + s_{i1}) \qquad (15)
\end{aligned}
$$

where $\pi_0 = (\beta - \frac{1}{2}) \rho^2 \frac{\sigma_\alpha}{\sigma_\alpha + 2\sigma_\zeta} \mu_\alpha$ is a constant term that incorporates the premium for skills and the characteristics of the ability distribution at time 1 and $\pi_1 = (\beta - \frac{1}{2}) \rho^2 \frac{\sigma_\zeta}{\sigma_\alpha + 2\sigma_\zeta}$ represents the coefficient of the sum of the two signals. At time $\tau = 1$, the student chooses the maximum between university continuation and the outside option in the labour market:

$$
Max \left[U_{ij\tau=1}; \mu_j \right] \qquad (16)
$$

Students are indifferent between drop out and university continuation if:

$$
U_{ij\tau=1}(S) = \mu_j
$$

[20] In a more detailed specification σ_ζ could be let vary across regions. This hypothesis would account for the fact that skills are distributed in a non homogeneous way across the territory (as documented by the stylized facts presented at the beginning of the paper). Also, heterogeneity along σ_ζ would capture regional differences in the quality of the information revealed in universities. Here, since in the most general case I am able to estimate only p_0 and p_1, these two additive sources of variation would not be identified. As a consequence, p_0 and p_1 are constrained to be constant across regions.

The individual propensity to drop out d^* is indeed:

$$d^*_{ijt} = \mu_j(1-\beta) + \delta_1 Z_{jt} + c_j - \pi_0 - \pi_1 (s_{i0} + s_{i1}) \tag{17}$$

Only actual drop out is observed, so the outcome of interest is a dummy variable $d^*_{ijt} = I(d^*_{ijt} > 0)$ that equals one when individual i quits university. Consistently with economic theory, the propensity is increasing in the monetary cost of education and in the value of the opportunity cost μ_j, whilst it is decreasing in the labour market premium for graduation ρ and in the number of universities' sites if $\delta_1 < 0$.

4.3 Entry choice

I now turn on defining the optimal entry choice at time 0. The expected utility at $\tau = 0$, after replacing x with x^* and $E(\alpha_i|s_{i0})$ with equation (3) in (10) is expressed as follows:

$$U_{ij\tau=0} = \left(\beta^2 \mu_j - \delta_0 Z_{jt} - c_j + \pi'_0 + \pi'_1 s_{i0}\right)(1 - E(d_i)) + \beta \mu_j E(d_i) \tag{18}$$

where $\pi'_0 = (\beta^2 - \frac{1}{2})\rho^2 \frac{\sigma_\alpha}{\sigma_\alpha + \sigma_\zeta} \mu_\alpha$ is the modified constant terms that embeds information known at time 0, $\pi'_1 = (\beta^2 - \frac{1}{2})\rho^2 \frac{\sigma_\zeta}{\sigma_\alpha + \sigma_\zeta}$ the coefficient of s_{i0} and $E(d_i)$ the expected probability of drop out next period. At time $\tau = 0$, the student chooses the maximum between the expected value of university entry and the outside option in the labour market:

$$Max \left[E(U_{ij\tau=0}(S)); \mu_j\right] \tag{19}$$

By imposing indifference between the two terms in squared brackets, exploiting the fact that μ_j is constant over time and for $\beta \to 1$, the propensity to entry university e^*_{ijt} reads as follows:

$$e^*_{ijt} = (\beta - 1)\beta \mu_j - \delta_0 Z_{jt} - c_j + \pi'_0 + \pi'_1 s_{i0} \tag{20}$$

which equals one when $e^*_{ijt} > 0$. The fraction of students that enroll at university is increasing in the number of sites.

4.4 Moving choice

The choice of moving to attend university in another region depends on the comparison between the utility gained by studying in the region of residence j and the utility gained by attending university in any other region $z \neq j$, paying an additive individual cost of moving from region j, b_i, which equals zero if the student enrol in his region of residence j. Let b_i be normally distributed in the population of movers, with mean μ_b and variance σ_b. The decision is taken in period $\tau = 0$, together with university entry. Individual i chooses the region z that yields the maximum utility net of moving cost:

$$U_{iz}(Z, x) - b_i = \ Max \ ((U_{i1}(Z, x) - b_i), .., (U_{iN}(Z, x) - b_i) \tag{21}$$

The comparison between the utility gained in the region of residence with that derived by moving to the region z yields the equation for the propensity to move to another region m^*_{jt}:

$$m^*_{ijt} = \beta^2(\mu_z - \mu_j) - \delta_0(Z_{zt} - Z_{jt}) - (c_z - c_j) - b_i \tag{22}$$

where m_{ijt} is an indicator equals one when individual i actually moves from region j to region z. Propensity to attend university in a different region is increasing in the unskilled wage differentials between the two regions $(\mu_z - \mu_j)$ and in differences relative to the number of universities' sites (if $\delta_0 < 0$), whilst it is decreasing in the cost of moving b_{ij} and in the monetary cost differential $(c_z - c_j)$.

4.5 Aggregation

In this section I compute the average outcome at regional level for a cohort t of individuals. Next, I compute the differences between the average outcomes of a cohort exposed to the program (students leaving secondary school in 1995), denoted t, with the average ones of those not exposed to the program (students graduated in 1998), denoted 0.

Aggregating equation (5.1) is straightforward and yields:

$$e^*_{jt} = (\beta^2 - 1)\mu_j - \delta_0 Z_{jt} - c_j + \pi'_0 + \pi'_1 \mu_\alpha \tag{23}$$

where $\mu_\alpha = E(s_{i0})$. Equation (22) becomes:

$$m^*_{jt} = \beta^2(\mu_z - \mu_j) - \delta_0(Z_{zt} - Z_{jt}) - (c_z - c_j) - \mu_b \tag{24}$$

Average drop out is censored from below given self selection into enrolment. Aggregating expression (17), it results:

$$\begin{aligned} d^*_{jt} &= \mu_j(1-\beta) + \delta_1 Z_{jt} + c_j - \pi_0 - \pi_1 \left[E(s_{i0}|s_{i0} > \alpha_t^{e^*}) + E(s_{i1})\right] \\ &= \mu_j(1-\beta) + \delta_1 Z_{jt} + c_j - \pi_0 - \pi_1 \left[2\mu_\alpha + \left(\frac{1}{\sigma_\alpha} + \frac{1}{\sigma_\zeta}\right)^{\frac{1}{2}} \lambda(\alpha_t^{e^*})\right] \end{aligned}$$

where $\lambda(\alpha_t^{e^*}) = \frac{\phi(\alpha_t^{e^*})}{1-\Phi(\alpha_t^{e^*})}$ is the correction for selection and $\alpha_t^{e^*} = \frac{\sigma_\alpha + \sigma_\zeta}{\sigma_\zeta}\left(\frac{2\left[\mu_j(1-\beta^2)+\delta_0 Z_{jt}+c_j\right]}{(2\beta^2-1)\rho^2} - \frac{\sigma_\alpha}{\sigma_\alpha+\sigma_\zeta}\mu_\alpha\right)$ is the critical point above which realizations of s_{i0} lead to enter university.

To aggregate the number of exams I need to account for the change of the average ability of the group of attendees due to the effects of the policy on enrolment and drop out behavior. In particular, I have to account for self selection into university continuation:

$$x^*_{jt} = \rho E(\alpha_i | s_{i\tau}, (s_{i0} + s_{i1}) > \alpha_t^{d^*}) \tag{25}$$

Exploiting the fact that the sum of two normal variables is normal, the distribution of $(s_{i0}+s_{i1})$ is normal with mean $2\mu_\alpha$, and variance $4\left(\frac{1}{\sigma_\alpha}+\frac{1}{\sigma_\zeta}\right)$. Censoring is accounted for as follows:

$$x^*_{jt}=p_0+p_1\left[2\mu_\alpha+2\left(\frac{1}{\sigma_\alpha}+\frac{1}{\sigma_\zeta}\right)^{\frac{1}{2}}\lambda(\alpha^{d^*}_{jt})\right]$$

where $\lambda(\alpha^{d^*})=\frac{\phi(\alpha^{d^*})}{1-\Phi(\alpha^{d^*})}$ is the correction for self selection into university at the end of the first period, when the decision to drop out has been taken, whilst $\alpha^{d^*}_{jt}=\frac{\sigma_\alpha+2\sigma_\zeta}{\sigma_\zeta}\left(\frac{2\left[\mu_j(1-\beta)+\delta_1 Z_{jt}+c\right]}{(2\beta-1)\rho^2}-\frac{\sigma_\alpha}{\sigma_\alpha+2\sigma_\zeta}\mu_\alpha\right)$ is the threshold below which realization of $(s_{i0}+s_{i1})$ implies drop out.

The differences between the average propensities to enrol and university, move away and drop out and the average number of exams take the following form:

$$e^*_{jt}-e^*_{j0}=-\delta_0(Z_{jt}-Z_{j0}) \tag{26}$$

$$m^*_{jt}-m^*_{j0}=-\delta_0(Z_{zt}-Z_{z0})+\delta_0(Z_{jt}-Z_{j0}) \tag{27}$$

$$d^*_{jt}-d^*_{j0}=-\delta_1(Z_{j0}-Z_{jt})-\pi_1\left[\left(\frac{1}{\sigma_\alpha}+\frac{1}{\sigma_\zeta}\right)^{\frac{1}{2}}(\lambda(\alpha^{e^*}_t)-\lambda(\alpha^{e^*}_0))\right] \tag{28}$$

$$x^*_{jt}-x^*_{j0}=p_1\left[2\left(\frac{1}{\sigma_\alpha}+\frac{1}{\sigma_\zeta}\right)^{\frac{1}{2}}(\lambda(\alpha^{d^*}_{jt})-\lambda(\alpha^{d^*}_{j0})\right] \tag{29}$$

The policy directly affected enrolment, mobility and drop out, and indirectly influenced drop out and the human capital investment, due to changes in the ability composition of students enrolled. More specifically, variations in Z between cohort 0 and t should have increased university enrolment if the monetary cost of education decreases with the number of universities' sites ($\delta_0<0$); the effect on mobility should be negative if universities sites increased more in the region of residence, i.e. $(Z_{jt}-Z_{j0})>(Z_{zt}-Z_{z0})$. Changes in drop out reflect two opposite forces. On one hand, the increased supply should have directly reduced the probability to quit university as continuing university became less costly ($\delta_1<0$). On the other, the policy increased the participation into university of marginal types, which might be indirectly more inclined to abandon studies. In fact, $\lambda(\alpha^{e^*})$ is a monotonic decreasing function of Z_j as a higher university supply lowers the thresholds above which realizations of s_{i0} induce enrolment. Therefore, the difference $\left(\lambda(\alpha^{e^*}_t)-\lambda(\alpha^{e^*}_0)\right)$ is negative. Which of the two effects prevails in equilibrium depends on the magnitude of the two forces and affects the direction of the change in the number of taken exams. If drop out increases because the composition effect exceeds the cost reduction induced by the policy, an increase in number of exams passed should be observed. If on the contrary the cost reduction effect prevails, performance should decrease after the policy implementation.

Denote with $P_j = \frac{Z_{jt} - Z_{j0}}{Z_{j0}}$ the increase in the number of university sites normalized with respect to the initial sites' level. As for descriptive statistics, this variable takes the values of one for regions that experienced an higher education expansion greater than 30 percent, while it takes values of zero otherwise. The strategy implemented in the paper consists in estimating equations (26)-(29) using weighted least squares. In practice, I estimate:

$$Y_{ijt} = a_j + \beta_t + \gamma(P_j * T_i) + \delta X_i + \theta T_i * Z_j + \varepsilon_{ijt} \tag{30}$$

$$Y_{ijt} = a_j + \beta_t + \gamma(P_j * T_i) + \eta\lambda(\alpha^{d^*,e^*}) + \delta X_i + \theta T_i * Z_j + \varepsilon_{ijt} \tag{31}$$

where equation (30) is used to estimate enrolment and mobility, whilst equation (31), which includes a non linear term that accounts for self selection in the previous stage (enrolment or drop out), is implemented to estimate drop out and exams passed. More specifically, Y is a variable indicating the outcome of interest for the individual i, resident in region j at the end of secondary school in period t; a_j is a region of secondary school fixed effect, β_t a cohort of graduation fixed effect, P_j denotes the intensity of the program in the region of residence at 19 years old. T_i is a "treatment dummy" which takes values of one for the 1998 cohort of secondary school leavers and zero for the control 1995 cohort; the coefficient γ measures the effect of the treatment on the treated. X_i is a vector of individual variables related to family background and past schooling career. These variables are added to account for possible changes in the composition of population characteristics over time.

Lastly, Z_{jt} is a vector of regional specific time varying controls including youth unemployment rate in 1990, number of youth who attained secondary education in 1992, measured at regional level, professors-students ratio in 1990, changes in the youth unemployment rate between 1995 and 1998 and differences in the number of regional degrees supplied between 1995 and 1998, net of degrees provided by new sites. The first three controls enter the baseline specification to check the sensitivity of results to the omission of time varying regional specific variables that should have affected the program and could have influenced educational outcomes when there is correlation between educational choices and initial levels values of potential demand, supply and unemployment. The fourth accounts for variations in regional labour market conditions and is included in order to purge the coefficients of interest from effects that are driven by variations of the outside option of education. The fifth variable is included to control for possible correlation of the program with the alternative intervention, that is degrees' expansion. $\lambda(\alpha^{d^*,e^*})$ is the correction term for self selection into enrolment or drop out. Finally, ε_{ijt} is a zero-mean stochastic error term, clustered at the region and cohort of graduates level to account for correlation of errors within region and time.

The identification framework outlined above can be generalized to an interaction term analysis to assess the specific effect of higher education expansion by family background, individual characteristics and type of secondary school attended. The objective is to identify the channels through

which the program exerted its effects. Consider the following specification between the individual outcome and his characteristics, generically identified by b_{il}:

$$Y_{ijt} = (\alpha_j + \beta_t + \delta X_{i,-b_i} + \theta T_i * Z_j) * \sum_{l=1}^{L} b_{i,l} + \sum_{l=1}^{L} \gamma_l (P_j * T_j * b_{i,l}) + \varepsilon_{ijt} \tag{32}$$

where b_{il} is a dummy that indicates whether individual i is of type l. More specifically, b_{il} is a vector of individual and family dummies characteristics. This vector includes the following variables: the mark at the end of junior school, broadly defined in four classes: A, B, C, D. This variable has been chosen as a proxy for individual ability instead of the mark at the end of upper secondary school because standards of the latter suffer high variability among different types of secondary school (a given mark in a vocational school does not convey the same information as the same mark in a high school), among schools of the same type in a given area and among different regions. Additionally, I account for the type of secondary school attended, that may be academic, vocational or technical; and, finally, for father's educational level, which is defined as follows: college degree, secondary school degree and compulsory school level or lower. Results for mother's educational level are not reported because very similar to that of father's one. Parents' occupation is disregarded since in Italy, as documented by Checchi (2003), parents' educational level is the strongest predictor of children's human capital investment[21]. Each coefficient γ_l can be interpreted as an estimate of the specific effect of ability, track chosen in secondary school and father's education on outcomes affected by higher education expansion.

5 Results

5.1 College enrolment

The first set of results is presented in Table 7 which shows linear probability estimates of the effect of the higher education expansion on college enrolment. The dependent variable Y_{ijt} takes value of one if the individual is enrolled in university and zero otherwise. The baseline specification in column 1 includes controls for region fixed effects, cohort of graduates dummy, program intensity interacted both with the treatment dummy and with the number of regional sites in 1995 every ten thousand squared kilometers. This last variable is added to account for sites' concentration. The distribution of this variable for treated regions is shown in Table 6. The inclusion of this interaction is justified by the fact that change in enrolment might depend on the initial level sites' concentration.

The effect of higher education expansion turns out to be positive and statistically significant (at 1 percent level). The coefficient slightly increases after the inclusion of individual specific controls, initial level variables that should have driven the policy (column 2 and 3) and variation in youth

[21] Indeed, parents' occupation is found significant only in few cases, in which it is included as a control.

	Sites	Freq.
Val D'Aosta	0.00	239
Trentino Alto Adige	0.74	228
Puglia	1.55	1,023
Emilia Romagna	1.81	1,074
Molise	2.25	628
Mean	1.27	

Table 6: Number of regional sites in 1995 for 10.000 km2, by treated region

unemployment rate between 1995 and 1998. The table reports the coefficient net of the interaction with the initial sites' concentration, evaluated at its mean value[22]. The estimated coefficient in the specification with the full set of controls indicates that on average, the probability of enrolment increased by 6.1 percent in regions affected by the policy.

Having established that the reform increased significantly enrolment rates, it is important to assess whether this expansion attracted students from less affluent backgrounds. Subsequent columns of the table provide an answer by looking at the interactions between the reform and personal characteristics. The specification chosen is the fourth one, which includes the whole set of regional time varying controls[23]. Results show that the expansion mainly benefited middle ability individuals from poorer household cultural capital: indeed there is a positive and significant effect of the program on enrolment of students awarded with a C mark at exit of lower secondary school and to a lower extent of students awarded with an A mark; students whose fathers had only secondary and primary school education increased their probability to enter higher education by on average 6 percent. Looking at the effect of the reform on enrolment of students from different secondary school types, it turns out that the expansion generally affected enrolment, regardless of secondary school type.

These results show that the program improved educational opportunities across the population. University entry increased for the marginal individuals, who are middle ability students coming from less educated parents. Two possible explanations might be given: either the expansion, lowering the cost of the investment, slackened liquidity constraints; or, having access to nearby universities weakened the role of household's information (that, in some sense, is correlated with his cultural capital) on the opportunity of acquiring higher education.

[22] I do not report separately the two coefficients, but the signs suggest that enrolment increased at a rate decreasing with the number of pre-existing universities.

[23] Coefficients reported include the interaction with initial sites concentration evaluated at its mean value.

	-1	-2	-3	-4	-5	-6	-7
	LPM	LPM	LPM	LPM	LPM	LPM	LPM
Treat+sites conc.*Treat	0,033	0,039	0,043	0,061			
	[4,58]***	[5,27]***	[5,17]***	[6,34]***			
Junior school mark A* (T+sites conc.*T)					0,232		
					[2,89]***		
Junior school mark B* (T+sites conc.*T)					0,112		
					[1,76]*		
Junior school mark C* (T+sites conc.*T)					0,502		
					[6,20]***		
Junior school mark D* (T+sites conc.*T)					-0,070		
					[1,40]		
Academic school*(T+sites conc.*T)						0,064	
						[3,76]***	
Technical school*(T+sites conc.*T)						0,058	
						[5,16]***	
Vocational school*(T+sites conc.*T)						0,069	
						[3,06]**	
Father college degree*(T+sites conc.*T)							0,013
							[0,74]
Father secondary degree*(T+sites conc.*T)							0,049
							[3,20]***
Father lower degree*(T+sites conc.*T)							0,075
							[5,90]***
Controls	no	yes	yes	yes	yes	yes	yes
Controls: Zj*Treatment	no	no	yes	yes	yes	yes	yes
Controls: Degree expansion*Treatment	no	no	no	yes	yes	yes	yes
Observations	36135	31619	31619	31619	31619	31619	31619
R-squared	0,01	0,43	0,43	0,43	0,44	0,45	0,44

Note: Robust t statistics in brackets; * significant at 10 percent; ** significant at 5 percent;*** significant at 1 percent.
Standard errors are clustered at region x cohort level. Observations are weighted to population weights.
Controls include: sites concentration, gender, parents education and occupation, siblings, past scores, type of sec school.
Regional controls include: youth unemployment rate in 1990, secondary school graduates (1992), number of prof. for 100 students (1990) and variation of youth unemployment between 1995-98; degree espansion is the variation in n.ro of degrees between 1995-98

Table 7: Probability of college enrolment

5.2 College drop out

Results discussed in Section 5.1 indicate that part of the effect of the expansion worked through a reduction of the impact of parental background on the choice of entering university. This result would imply an increase in intergenerational mobility only if this additional inflow of students is not more largely inclined to drop out university. Otherwise, changes at entry would not translate into equivalent changes at university exit, without reducing the role of family background on human capital investment.

I can directly assess the effect of the expansion on withdrawal because in the survey respondents were asked whether or not they began and then interrupted academic studies. Since the question is answered three years after enrolment and since 35 percent of students generally quit university within the first three years[24], this variable is a good measure for drop out changes.

The theoretical specification suggests that learning over individual ability endogenously leads students to drop out if the signal acquired during university attendance reveals low talent. Also, the model shows that individuals are non randomly sorted into drop out and that the distribution could be truncated from below given previous self selection into enrolment, thus calling for a Heckman selection model. To identify the parameters of interest without excessive reliance on functional forms, it is necessary to instrument selection in enrolment with a variable that affects the choice of entering university, but without directly influencing the individual decision to drop out. The model suggests that observable family characteristics should be embedded in the first signal, whilst information on individual ability disclosed during university attendance is captured by the second one. A suitable exclusion restriction within this framework is a variable related to the family social class as parental occupation, which can be seen as a proxy for family permanent income. These variables (dummies for paternal class of occupation), when included as controls in the enrolment equation, are not found to be jointly statistically significant in the drop out equation and can be indeed considered valid restrictions[25].

As the institution of a new university is suggested to influence both the selection equation and the outcome, it is included as a regressor in the probit and OLS stages of the model. The interpretation of the marginal effect for this variable must consequently be adjusted to correct for selectivity bias. Hoffmann and Kassouf (2005) derive the correction for a regressor that is a binary variable[26], which reads:

[24] Source: MIUR (1998) "The evolution of the demand for higher education: students, graduates and equivalent students".

[25] Cinzano and Cipollone (2007) explain withdrawal decision as a consequence of short term family credit constraints. This interpretation is not in contrast with my exclusionary restriction in so far parental occupation captures permanent income and not transitory income shocks that could induce drop out.

[26] For a derivation of the marginal effect of a continuous regressor in the Heckman model see Greene "Econometric Analysis".

$$\frac{\partial(d^*_{jt} - d^*_{j0}|s_{i0} > \alpha^{e^*}_t)}{\partial(Z_{jt} - Z_{j0})} = -\delta_1 + \beta_\lambda \Delta\lambda$$
$$= -\delta_1 + \beta_\lambda \left(\lambda(\alpha^{e^*}_{Z_{jt} \neq Z_{j0}}) - \lambda(\alpha^{e^*}_{Z_{jt} \neq Z_{j0}})\right) \quad (33)$$

where β_λ is $\left(\frac{1}{\sigma_\alpha} + \frac{1}{\sigma_\zeta}\right)^{\frac{1}{2}}$, $\lambda(\alpha^{e^*}_{Z_{jt} \neq Z_{j0}})$ is the selection equation evaluated at $P_j = \frac{Z_{jt} - Z_{j0}}{Z_{j0}} = 1$ and all the other regressors at their mean value and $\lambda(\alpha^{e^*}_{Z_{jt} \neq Z_{j0}})$ is the same vector evaluated at $P_j = 0$[27].

Table 8 presents results from a Heckman two stages equation of a drop out indicator for the sub sample of individuals self selected into higher education. Coefficients reported are calculated according to the above correction. In all specifications supply expansion turned out to both negatively and significantly affect drop out decision. The average effect was in order of -2.5 percent. Coefficient estimates halve when regional controls are included. The negative estimate shows that the prevailing effect of the Higher Education expansion was the cost reduction associated with university continuation. When looking at the coefficient of P_j alone in the outcome equation it turns out that it is only slightly different from the one estimated with the correction, suggesting that the composition effect is small with respect to the monetary cost reduction one.

The interaction term analysis shows that drop out reduction occurred especially among students marked A at junior school, which reduced by 3.4 percent their probability of quitting university and students marked C, which decreased it by 5.6 percent; also, students with lower educated parents had a closer gap in drop out rate compared to students with university educated parents, by 4.3 percent. Such evidence proves that the additional intakes into university due to the expansion did not experienced larger drop out and that individual previously liquidity constraint reduced their propensity to quit higher education, being consistent with the idea that policy increased equality of opportunities.

[27] Standard errors for the corrected marginal effects are computed using the Delta method.

	-1	-2	-3	-4	-5	-6	-7
	Heckman	Heckman	Heckman	Heckman	Heckman	Heckman	Heckman
Treatment*(1-eta)	-0,045	-0,049	-0,025	-0,025			
	[2,66]***	[2,88]***	[2,68]***	[2,76]***			
Junior school mark A* Treat*(1-eta)					-0,034		
					[3,53]***		
Junior school mark B* Treat*(1-eta)					-0,047		
					[1,52]		
Junior school mark C* Treat*(1-eta)					-0,056		
					[2,55]**		
Junior school mark D* Treat*(1-eta)					0,033		
					[1,08]		
Academic school*Treat*(1-eta)						0,008	
						[0,60]	
Technical school*Treat*(1-eta)						-0,024	
						[0,95]	
Vocational school*Treat*(1-eta)						-0,024	
						[0,90]	
Father college degree*Treat*(1-eta)							0,035
							[2,96]**
Father secondary degree*Treat*(1-eta)							-0,043
							[2,26]**
Father lower degree*Treat*(1-eta)							-0,006
							[0,34]
Controls	no	yes	yes	yes	yes	yes	yes
Controls: Zj*Treatment	no	no	yes	yes	yes	yes	yes
Controls: Degree expansion*Treatment	no	no	no	yes	yes	yes	yes
Observations	31619	31619	31619	31619	31619	31619	31619

Note: Robust t statistics in brackets; * significant at 10 percent; ** significant at 5 percent;*** significant at 1 percent.
Standard errors are clustered at region x cohort level. Observations are weighted to population weights.
Controls include: gender, parents education, past scores and type of sec school.
Regional controls include: youth unemployment rate in 1990, secondary school graduates (1992), number of prof. for 100 students (1990) and variation of youth unemployment between 1995-98; degree espansion is the variation in nro of degrees between 1995-98. Selection eq includes: treatment, treatment*sites concentration, gender, parents education, parents occupation, past scores, type of sec school and variation of youth unemployment between 1995-98

Table 8: Probability of college drop out

5.3 Students mobility

A second expected effect of such a local sites expansion is to reduce mobility of students across regions, which in Italy has always been impressively limited. Table 9 presents the results from estimating the baseline specification of the effects of university sites expansion on college mobility. The outcome of interest is a dichotomous variable which takes the value of one when the student attended university in a region different from the region of secondary school attendance, and zero otherwise. The database do not allow me to assess the concept of mobility more deeply, which may refer to the choice of a university located in the same region but also in a different city, since information on the specific university entered are not disclosed. Thus, although restricted, definition of mobility is limited to the regional one.

As for the probability of entering university, the universities' concentration before the policy was implemented is included and interacted with the treatment dummy and coefficients presented are net of the interaction. Looking at the results, it turns out that the probability of moving to a different region decreased significantly (at 1 percent level in all specifications) by 10 percent when new sites were set up in the region of residence. The estimate remains stable when regional time varying variables and complementary policy are added. This result suggests that there is no bias induced by omitted variables and that therefore the identification strategy is reasonable.

Looking at the composition effect, it turns out that the reduction was concentrated among high ability individuals from less affluent backgrounds. In particular, students who were marked A at junior school reduced by 17 percent points (significant at 1 percent level) the likelihood to change region and sons of parents with secondary or lower education decreased the probability by 10 percent.

A possible reason why smarter students reduced their mobility rate is that these individuals, who were previously forced to move to attend university, appreciated the opportunity to study in their region of residence: likely the cost of moving to another region was not compensated by higher return, probably because the degree has the same legal value regardless of the university attended.

	-1	-2	-3	-4	-5	-6	-7
	LPM	LPM	LPM	LPM	LPM	LPM	LPM
Treat + sites conc.*Treat	-0,096	-0,094	-0,076	-0,097			
	[5,81]***	[5,92]***	[6,18]***	[5,59]***			
Junior school mark A* (T+sites conc.*T)					-0,173		
					[7,99]***		
Junior school mark B* (T+sites conc.*T)					-0,054		
					[1,69]		
Junior school mark C* (T+sites conc.*T)					-0,034		
					[0,99]		
Junior school mark D* (T+sites conc.*T)					-0,062		
					[1,31]		
Academic school*(T+sites conc.*T)						-0,095	
						[4,05]***	
Technical school*(T+sites conc.*T)						-0,097	
						[4,27]***	
Vocational school*(T+sites conc.*T)						-0,079	
						[2,11]**	
Father college degree*(T+sites conc.*T)							-0,007
							[0,24]
Father secondary degree*(T+sites conc.*T)							-0,098
							[4,74]***
Father lower degree*(T+sites conc.*T)							-0,124
							[4,14]***
Controls	no	yes	yes	yes	yes	yes	yes
Controls: Zj*Treatment	no	no	yes	yes	yes	yes	yes
Controls: Degree expansion*Treatment	no	no	no	yes	yes	yes	yes
Observations	14998	13579	13579	13579	13579	13579	13579
R-squared	0,12	0,13	0,14	0,15	0,15	0,15	0,15

Note: Robust t statistics in brackets; * significant at 10 percent; ** significant at 5 percent;*** significant at 1 percent.
Standard errors are clustered at region x cohort level. Observations are weighted to population weights.
Controls include: gender, parents education, father's occupation, past scores and type of sec school.
Regional controls include: youth unemployment rate in 1990, secondary school graduates (1992), number of prof. for 100 students (1990) and variation of youth unemployment between 1995-98; degree espansion is the variation in nro of degrees between 1995-98

Table 9: Probability of moving to a different region

5.4 Exams passed

An interesting effect of this higher education expansion concerns the impact on individual academic performance. As discussed at the beginning, a major failure of the Italian higher educational system is due to the extremely long period of time many students take to graduate from university. Oddly, this longer period of time is not explained by a parallel activity carried out in the labour market. Rather, the fraction of students employed in the age bracket 20-24 was, in 1999, roughly 3.2 percent in Italy, against an average 11.8 value displayed by other Oecd countries[28].

The sequential model suggests that a reduction of the drop out rate would imply a lower performance. Indeed, changes occurred in the composition of the stock of stayers due to the higher intakes and the lower withdrawal should have affected the overall composition of students' characteristics in regions where the expansion has been implemented. Likely, since intakes and new stayers were composed of individuals with middle ability, the new students' body should feature a lower average ability. This composition effect should have negatively affected aggregate regional performance.

The data enable identification of the number of exams passed thanks to a direct survey question. However, since the question was not answered by about 16 percent of enrolled students, the number of valid observations is censored. Table 10 presents the distribution of non reported exams passed by results at exit of lower secondary school. Evidence demonstrates that individuals in the lower tail of the ability distribution tend more to not report their academic performance. This should imply a positive bias in the coefficient of low ability individuals when turning to the interaction term analysis[29].

As for drop out, estimating number of exams requires to account for self selection in the group of students that continue university. However, as noticed in note 18, the choice of drop out and the amount of human capital investment are probably made simultaneously and not in a strict sequential order. In fact, the number of exams given are also observed in the data for a group of students who dropped out. To empirically account for these observations selection is applied to the enrolment choice using the same identification assumption implemented in section 5.2. Moreover, the theoretical model indicates that the effect of the policy on the human capital investments worked throughout the non linear composition term. However, in the empirical specification I add a linear term for the policy to capture possible complementary between the non monetary cost of education and the monetary one. Thus, the marginal effect of new universities' set up is estimated with the correction presented in equation (33).

The first set of results are presented in Table 11, which shows coefficient estimate from a Heckman procedure applied to the number of exams passed on a set of controls, which included the type of degree entered. The effect of the policy, stable across all specifications, indicates that

[28] See: Education at a glance, 1999

[29] Looking at the distribution of non reported passed exams by region of residence, no differences emerge between Northern and Southern ones.

the number of exams passed decreased by -0.46 as a consequence of sites' expansion. Looking at the interacted analysis, it emerges a significant reduction in exams passed for students marked C at junior school and an increase for students marked D. However, given the high percentage of non reporting data in the group of students who obtained a C and D mark at the end of lower secondary school, both coefficients are likely upward biased, indicating that students with middle ability reduced by more than 1.5 the number of exams taken. The decline in exams passed is significant for individuals coming from vocational schools (-1) and to a less extent for those exiting academic ones (-0.7). Children of parents with secondary degree passed nearly 1 fewer exam. This evidence seems to hold up the idea that a composition effect has played a role. Indeed, individuals that slowed their progress down were those that in the absence of the policy, would not have entered university or would have dropped it.

Another possible explanation could rely on the fact that the reduction of the direct monetary cost of higher education might have made less worthwhile the investment and implied a lower non monetary cost of education as well.

Overall it seems that this type of policy went in the direction of retaining students in the schooling system, but slowed their performance down, thereby prolonging time passed in education. A different reasonable explanation, which is not captured by the outlined model, may rely on possible organizational differences between well established universities and new ones, that could take some times before setting up an efficient organization. However, I argue that this should be a minor issue, since, as recent literature highlighted (see Bagues, M., Sylos Labini, M. and Zinovyeva, N., 2006), new universities should have had more incentives to lower grades standard and facilitate academic students' progress in order to attract local demand.

	J.S. Mark A	J.S. Mark B	J.S. Mark C	J.S. Mark D
Missing data	283	431	697	862
Total	4218	3465	3669	2959
Percentage	6,7%	12,4%	19,0%	29,1%

Table 10: Distribution of missing data relative to the number of passed exams, by Mark scored at junior school

	-1	-2	-3	-4	-5	-6	-7
	Heckman	Heckman	Heckman	Heckman	Heckman	Heckman	Heckman
Treatment	-0,333	-0,425	-0,488	-0,463			
	[1,38]	[1,95]*	[1,84]*	[1,87]*			
Junior school mark A* Treatment					0,157		
					[0,69]		
Junior school mark B* Treatment					-0,143		
					[0,63]		
Junior school mark C* Treatment					-1,610		
					[2,95]**		
Junior school mark D* Treatment					1,460		
					[3,26]***		
Academic school*Treatment						-0,725	
						[2,64]**	
Technical school*Treatment						0,237	
						[0,94]	
Vocational school*Treatment						-0,975	
						[2,58]**	
Father college degree*Treatment							0,041
							[0,11]
Father secondary degree*Treatment							-1,023
							[2,71]**
Father lower degree*Treatment							-0,063
							[0,23]
Controls	no	yes	yes	yes	yes	yes	yes
Controls: Zj*Treatment	no	no	yes	yes	yes	yes	yes
Controls: Degree expansion*Treatment	no	no	no	yes	yes	yes	yes
Observations	28634	28634	28634	28634	28634	28634	28634

Note: Robust t statistics in brackets; * significant at 10 percent; ** significant at 5 percent;*** significant at 1 percent.
Standard errors are clustered at region x cohort level. Observations are weighted to population weights.
Controls include: gender, age, parents education, past scores, type of sec school, employment status and degrees dummies.
Regional controls include: youth unemployment rate in 1990, secondary school graduates (1992), number of prof. for 100 students (1990); degree espansion is the variation in nro of degrees between 1995-98. Selection eq includes: treatment, treatment*sites concentration, gender, parents education, parents occupation, past scores, type of sec school and variation of youth unemployment between 1995-98

Table 11: Number of passed exams

Another possible reason why a reduction of the educational progress is observed could rely on a more intensive activity in the labour market. In fact, shorter travelling distances to university could have released time to be allocated on working activities rather then on studies. However, by running a linear probability model on the likelihood of being a working student, it emerges that new sites had no effects at all on the propensity to carry on academic activities along with working ones.

6 Robustness checks

In this section I perform some checks to assess the robustness of results to other specifications. The first is the use of a different cut off point to verify whether findings are sensitive to the 30 percent increase in sites' provision threshold. Setting the cut off at 25 percent[30] implies the inclusion of the region Sicilia among treated ones. A point worth noticing is that the new site instituted in Sicilia between 1995-1998, was constituted only by the Faculty of Architecture.

Table 12 presents results for the outcomes of interest, according to the specification which includes both individual controls, regional time varying variables and degree expansion, all interacted with the treatment status. Enrolment and mobility estimates include the interaction with the sites concentration and drop out and exams passed estimates are corrected for self selection. For drop out and exams passed, coefficients are qualitatively and quantitatively similar to those estimated with a higher cut off point. For mobility the coefficients has the expected sign but is not significant at 10 percent level. For enrolment the coefficient is zero and not significant. These two last results can be explained by the low impact that the new site in Sicilia had for existing demand of education. Indeed, when instituted, the new faculty enrolled 0.38 percent of the demand for higher education in Sicilia[31]. The inclusion of this region in the group of treated likely lowered the average impact of the policy.

Another useful test regards the sensitivity of results to the control and treatment group composition. Since policy has been implemented in regions located in the North and South of Italy, it could be useful to restrict the analysis to a group of regions sited in the same macro-area, likely satisfying the parallel trend hypothesis. Two groups of regions are considered: Northern ones, including Piemonte, Val D'Aosta, Lombardia, Trentino Alto Adige, Veneto, Friuli Venezia Giulia, Liguria and Emilia Romagna; and Southern ones, which take in Molise, Campania, Basilicata, Puglia and Calabria. Among Northern regions three have been treated (Trentino, Val D'Aosta and Emilia Romagna), whilst only two, Molise and Puglia, are among the Southern ones[32].

Table 13 shows results[33] from estimating outputs of interest separately for the two sub samples

[30] I do not perform the analysis with a cutoff higher than 50 percent because the sample of treated observations is too small (roughly 3 percent) to give reliable estimates.

[31] See ISTAT (1999) "Lo stato dell'università: i principali indicatori", Appendice Statistica.

[32] Since the purpose of the exercise is to consider outcomes whithin similar economies, regions located in the Center of Italy, among which none has been treated, were excluded.

[33] As in previous specification, enrolment and mobility estimates include the interaction with the sites concentra-

	-1	-2	-3	-4
	LPM	Heckman	LPM	Heckman
	Enrolment	Drop out	Mobility	Exams
Treatment	-0,008	-0,024	-0,029	-0,585
	[0,50]	[3,69]***	[1,53]	[2,93]**
Controls	yes	yes	yes	yes
Controls: Zj*Treatment	yes	yes	yes	yes
Controls: Degree exp*Treatment	yes	yes	yes	yes
Observations	31619	31619	13579	28634
R-squared	0,43		0,14	

Note: Robust t statistics in brackets; * significant at 10 percent; ** significant at 5 percent;*** significant at 1 percent. Standard errors are clustered at region x cohort level. Observations are weighted to population weights. Each outcome is estimated according to the 4th column specification in tables 7-8-9-11. The coefficient reported for enrolment and mobility accounts for the interaction with the intial level sites' concentration, and that for drop out and number of exams is corrected for self selection into enrolment.

Table 12: Outputs estimated at sites expansion higher than 25 percent

of population with the full set of controls, individual and time varying regional specific. It emerges that results are consistent in both groups, without showing strong differences according to the area, except for the number of exams passed. In particular, enrolment increased by a similar amount in Northern and Southern regions. The probability of being retained in university was slightly higher in Northern regions. The propensity to move to a region different from the one of residence decreased more in Southern regions; the number of exams passed significantly fell in the North by nearly -0.2, whilst in Southern ones by -1.5.

A reasonable way to explain differences in magnitude of performance between North and South of Italy is to analyze the change in the composition of students enrolled, due to the supply expansion. To this, I look both at the composition of the flow of new intakes into university and of new stayers as a consequence of the reduced probability of quitting university. Mobility surely had an impact, but since all outcomes are evaluated considering the region of residence at the end of lower secondary school, the effects of endogenous migration didn't translate into performance. To see whether the flow of new stayers varied across areas, I perform a more detailed analysis of the previous regressions, by adding the interaction term analysis for the type of secondary school

tion but only in the specification restricted to Northern regions (the coefficient of the initial sites concentration has been evaluated at its mean value for Northern region, which is 1.62). Sites concentration has been dropped from the specification restricted to Southern regions because it was found not significant. Drop out and passed exams estimates are corrected for self selection.

North

	-1	-2	-3	-4
	LPM	Heckman	LPM	Heckman
	Enrolment	Drop out	Mobility	Exams
Treatment	0,008	-0,048	-0,062	-0,193
	[7,94]***	[2,93]***	[7,51]***	[2,65]**
Controls	yes	yes	yes	yes
Controls: Zj*Treatment	yes	yes	yes	yes
Controls: Degree exp.*Treat.	yes	yes	yes	yes
Observations	12364	12364	5176	11305
R-squared	0,04		0,05	

South

	-1	-2	-3	-4
	Enrolment	Drop out	Mobility	Exams
Treatment	0,011	-0,030	-0,136	-1,423
	[3,90]***	[24,29]***	[16,80]***	[7,29]***
Controls	yes	yes	yes	yes
Controls: Zj*Treat.	yes	yes	yes	yes
Controls: Degree exp.*Treat.	yes	yes	yes	yes
Observations	7525	7525	13579	7035
R-squared	0,41		0,10	

Note: Robust t statistics in brackets; * significant at 10 percent; ** significant at 5 percent;*** significant at 1 percent. Standard errors are clustered at region x cohort level. Observations are weighted to population weights. Each outcome is estimated according to the 4th column specification in tables 7-8-9-11. The coefficient reported for enrolment and mobility accounts for the interaction with the intial level sites' concentration for the group of Northern regions, and that for drop out and number of exams is corrected for self selection into enrolment.

Table 13: Outputs estimated for different subgroups of the population

attended[34]. Results are reported in Table 14[35]. By looking at the composition of the flow of new entrants, it turns out that the principal difference between the North and the South of Italy concerns the group of students coming from a professional secondary school: their probability of entering university strikingly fell in the North, whilst it sharply increased in the South. Even if this group of students rose the probability of drop out university in Southern regions, the majority of them were retained in university. Moreover, drop out decreased significantly in the North among students who attended high school, but who on the contrary increased their abandon rate in the South. Thereby, the change in the composition of the stock of stayers, due to a higher propensity to enter university and a lower propensity to abandon it, increased the overall participation of students who were less prepared to afford academic studies, specially in Southern regions. Indeed, looking at the composition of individuals who diminished the number of taken exams, some correspondences can be found with the group of students who increased the probability of entering university and decreased the probability of quitting it after the policy was implemented. Especially students coming from professional schools slowed down their performance, reducing the number of exams passed in Southern regions by nearly -5. This evidence holds up with the idea that a compositional effect has played an important role in shaping different outcomes of magnitude.

The reason why students with the same secondary school curriculum made different educational choices between the two macro areas could be due to heterogenous changes in labour market conditions. Indeed, in the theoretical framework optimal choices have been derived under the assumption that the outside option of education stayed constant over time, whilst in the empirical implementation I relaxed this hypothesis, by controlling for possible changes in youth unemployment over the period 1995-1998. Likely, this control has not captured all possible variations occurred in the labour market, since, besides unemployment rate, employment opportunities might have evolved differently among types of jobs and across regions. Future research is needed to highlight the role of uncertainty in the labour market on youth educational choices.

7 Conclusion

In this paper I used pooled data on two cohorts of secondary school graduates to assess the impact of university sites expansion on a series of indicator related to human capital investments, exploiting the quasi-natural experiment of this policy.

I find that new sites increased university enrolment and that the effect was largely concentrated among middle ability individuals with less favorable family background. This new flow of intakes significantly increased the probability of being retained in the university system. This evidence can be interpreted as a proof of existence of liquidity constraints that prevented a subset of population from investing in higher education.

[34]I do not report coefficients for the interaction by mark at exit of lower secondary school as differences between the North and South are less wide.

[35]Drop out is estimated with a linear probability model as the Heckman model does not converge with many interacted terms.

	LPM		LPM		Heckman	
	Enrolment		Drop out		N. exams	
	North	South	North	South	North	South
Highschool	0,004	-0,004	-0,033	0,014	0,296	-1,073
	[1,46]	[1,54]	[21,76]***	[4,88]***	[2,35]*	[8,67]***
Technical secondary school	0,032	-0,039	0,001	-0,01	-1,046	-0,977
	[8,52]***	[1,19]	[0,08]	[3,73]***	[4,42]***	[4,02]***
Professional secondary school	-0,041	0,068	-0,014	0,024	-3,064	-4,924
	[10,18]***	[12,81]***	[1,78]*	[13,42]***	[3,30]***	[2,69]**
Observations	12364	7525	12847	7785	11767	7035
R-squared	0,45	0,44	0,03	0,33		

Note: Robust t statistics in brackets; * significant at 10 percent; ** significant at 5 percent;*** significant at 1 percent. Standard errors are clustered at region x cohort level. Observations are weighted to population weights. Each outcome is estimated according to the 4th column specification in tables 7-8-9-11. Drop out is estimated with a LPM. The coefficients reported for enrolment and mobility account for the interaction with the initial level sites' concentration for the group pf Northern regions, and those for number of exams are corrected for self selection into enrolment

Table 14: Output estimated by type of sec. school for different subgroups of students' population

However, local universities did not boost successful academic performance: rather, incentives for doing well at university lowered, particularly in Southern regions affected by the policy, thereby causing a decline in the number of exams passed. This lower academic performance cannot be explained by a more intense activity in the labour market, since the probability of being employed and carrying on academic activities was not affected at all by the program. The explanation relies on the change in students' composition due to the new flow of intakes into university and the reduced outflow of drop out. As a consequence, the average preparation of the pool of students who entered university likely lowered. Another effect worth mentioning is the significant reduction in mobility of students across regions, especially among individuals with good schooling ability and less affluent family background. All results are robust to the inclusion of a wide set of controls and to different compositions of the control and treatment groups.

This university market structure, which encouraged local institutions, on one side increased equality of opportunities across the national territory by opening access to more groups in the society and reducing the impact of family background on the decision to enter higher education; but on the other hand, it went in the direction of strengthening across regional disparities, given wide negative effects exerted on individual performance in Southern regions.

In conclusion, the planned objectives, namely the necessity of rebalancing higher educational supply in order to reduce across regional disparities, have not been achieved. My results indicate that the institution of a new university in a less advantaged region has not been a good instrument to attenuate differentials with more developed ones. One last point worth mentioning concerns the type of expansion, which encouraged the set up of local institutions. Given the structure of the

secondary school system, which can be divided into academic and labour market oriented tracks, an expansion of the supply of higher education stratified along the academic-vocational dimension would have been preferable, in that this type of structure would have lead students to choose the type of tertiary education better suited to their curricula and aspirations.

References

[1] Angrist J. (1995) "Economic return to schooling in the West Bank and Gaza strip", American Economic Review, 85:5, 1065-1087

[2] Bargues, M., Sylos Labini, M. and Zynovyena, N (2006) "Endogenous grading standards and labour market mismatch", paper presented at the Brucchi Luchino 2006 conference (Padua)

[3] Becker, G. (1964) "Human Capital: theoretical and empirical analysis with special reference to education". Columbia University Press

[4] Berlinski, S., Galiani, S. (2007) "The effects of a large expansion of pre-primary school facilities on pre school attendance and maternal employment", Labour Economics, 41: 665-680

[5] Besley, T.,Case, A. (2000) "Unnatural experiment? Estimating the incidence of endogenous policy", Economic Journal, 110: 672-694

[6] Biggeri, L. (1998) "Programmazione e valutazione dello sviluppo del sistema universitario", MIUR

[7] Bratti, M., Checchi, D., De Blasio, G. (2008) "Does the expansion of higher education increase equality of opportunities? Evidence from Italy", Labour, forthcoming

[8] Cappellari, L., Lucifora, C. (2007) "University reforms and human capital investment", Iza Discussion Paper

[9] Checchi, D. (2003) "The Italian educational system: family background and social stratification", in ISAE 2003, Annual Report in Monitoring Italy. Roma, pp. 131-176

[10] Cingano, F., Cipollone P. (2007) "University drop out: the case of Italy", Temi di Discussione, Banca d'Italia n. 626

[11] Dixit, A.K. Pindyck, R.S. (1994) "Investment under uncertainty", Princeton University Press, Princeton, N.J.

[12] Duflo, E. (2001) "Schooling and labour market consequences of school construction in Indonesia: evidence from an unusual policy experiment", American Economic Review, 91: 795-813

[13] Heckman, J., (1979) "Sample selection bias as a specification error" Econometrica, 47:1, 153-161

[14] Heckman, J., Lochner, L., Taber, C. (1998) "General-equilibrium treatments effects: a study of tuition policy" American Economic Review, 88:2, 381-386

[15] Hoffman, R., Kassouf, A.L. (2005) "Deriving conditional and unconditional marginal effects in log earnings equations estimated by Heckman's procedure", Applied Economics 37:11, 1303-1311

[16] Holmstrom, B. (1999) "Managerial incentives problem: a dynamic perspective", The Review

of Economic Studies, vol. 66:1, 169-182
[17] Holzer, S. (2006) "The expansion of higher education in Sweden and the issue of equality of opportunities", mimeo
[18] Istat (1994) "Statistiche delle scuole secondarie superiori", Roma
[19] Istat (1997) "Statistiche delle scuole secondarie superiori", Roma
[20] Istat (1995) "Statistiche dell'istruzione universitaria", Roma
[21] Istat (1998) "Statistiche dell'istruzione universitaria", Roma
[22] Istat (1999) "Lo stato dell' università: i principali indicatori", Roma
[23] Johnson, W. (1978) "A theory of job shopping" Quarterly Journal of Economics, 92:2, 261-278
[24] Jovanovic, B. (1979) "Job match and the theory of turnover", Journal of Political Economy, 87: 972-990
[25] Meyer, B. (1995) "Natural and quasi-experiments in Economics", Journal of Business and Economics and Statistics, vol. 13 n.2
[26] MIUR (1998) "L'evoluzione della domanda di formazione universitaria: studenti, laureati e studenti equivalenti"
[27] OECD (2000), Education at a glance, Paris
[28] Rosenzweig, M. R., Wolpin, K. I. (1988) "Evaluating the effects of optimally distributed programs: child health and family planning interventions", American Economic Review, 76: 470-482
[29] Rosenzweig, M. R., Wolpin K. I. (1998) "Migration selectivity and the effects of public programs", Journal of Public Economics, 37: 265-289
[30] Vance, C., (2007) "Marginal effects and significance testing with Heckman's sample selection model: a methodological note", Economic Letters, 1-5