LOMBARDY ADVANCED SCHOOL OF ECONOMIC RESEARCH

PhD in Economics

Cycle XXVIII

Topics in education and labor economics

PhD candidate: Daniele Biancardi

Thesis supervisor: Massimiliano Bratti

Secondary advisor: Alessandro Bucciol

Academic Year 2015/2016

Acknowledgments

In the last three years I discovered that the life of a PhD student can often be quite hard. However, I have been lucky enough to meet some special people who helped me along the way. I would have never been able to get to the end without them.

The first person I must thank is Massimiliano, my supervisor. He always encouraged and guided me with his advise and suggestions. He has been an excellent supervisor and I owe to him most of the things I learned.

Alessandro Bucciol has always been kind with me and gave me some very useful feedback and advice on my work. I thank Marco Leonardi for the opportunities and the encouragement he gave me during the past years, asking about my work and pushing me to improve. Ylenia Brilli helped me to get started with structural methods sharing here teaching material and some useful advice.

Tommaso Frattini made some helpful comments about my work in more than one occasion and Stefano Verzillo helped me with the data I used in chapters 1 and 2. I thank Paolo Bertoletti for allowing me to work in Pavia during the last three years.

I also need to thank my PhD colleagues Kinzica Soldano, Chiara Punzo and Fatma Usheva for a lot of interesting conversations about research (and many more topics) and for supporting me when I was downhearted. They have proved to be real friends in the office as well as in real life.

Last (but not least) I thank my parents and my brother because they have always been there for me when I needed them and my grandparents, that I lost along the way but I will always feel very close.

Contents

List of Figures

List of Tables

Co-authorship disclaimer

The second chapter of this thesis is a joint work with Massimiliano Bratti.

Preface

This thesis is composed by three distinct chapters. The first two contribute to the economics of higher education literature, while the third estimates a structural model of household behavior.

Chapter 1 presents a study assessing the impact of grading standards (GS) in Italian departments on the labor market outcomes of university graduates. The influence of heterogeneous GS on labor market performance can occur through two different channels: a productivity and a signaling effect.

The empirical papers trying to answer the same research question are quite rare due to limitations in data availability. This study provides first evidence on the dynamic effects of GS on university graduates in Italy, evaluating the impact on wages, employment and overeducation. The analysis is performed using unique data provided by Almalaurea on graduates in years 2008 and 2009 matched with department-level information on research quality and resources.

Italy is an interesting case study since university graduation rates are low but, at the same time, returns to higher education are below the average of other developed countries. The human capital accumulated is also quite low. The PIAAC data, measuring the level of skills in OECD countries by level of education, place Italian university graduates at the bottom of the ranking. For these reasons it is important to find policies that can increase the average productivity of highly educated workers. Furthermore, in the last decades the increased supply in the market for higher education and the 3+2 reform lead to a larger heterogeneity in quality and in GS between institutions.

The estimation strategy is divided in two steps. Firstly, we estimate a proxy for GS as the part of final grades which cannot be explained by differences in individual characteristics (student's quality) and other relevant inputs (quality of the institution attended). Then, the effect of GS on wage and other labor market outcomes is estimated. We show that differences in GS are large across departments. More generous grades are associated to a wage penalty on the labor market 5 years after graduation. In particular, graduates from 'generous' departments earn 3.4% less than people who studied in the 'strict' departments, they have a lower employment rate and a higher probability of being too educated for their jobs. The effects on wages are stronger for high ability workers while employment is more affected for low ability and female graduates.

Chapter 2 assesses the impact of the first Italian Research Evaluation Exercise (REE) on students' enrollment choices.

All Italian REEs have been followed by lively debates. Critics of REE maintain that they are very expensive and excessively based on quantitative (e.g., bibliometric) indicators. Advocates of REEs rebut that in a period of shrinking public funding of Higher Education it is more important than ever to allocate resources in an effective and efficient way. However, there is no evidence on the effect of the REE on students' choices.

Our paper is related to the literature which, especially in the US, has investigated the effects on student application and matriculation decisions of ratings and rankings of Higher Education Institutions (HEIs) produced by private 'intermediaries'.

We provide a first assessment of the impact of the 'Valutazione Triennale della Ricerca' (VTR) on student choices using a before-after estimator which exploits differential treatment intensities across HEIs. In particular, we investigate whether departments that received a better score also benefited of more student enrolments and enrolments of students with better entry qualifications after the VTR. This identification strategy enables us to control for both department-specific time invariant unobservable heterogeneity and

pre-existing department trends.
The analysis is performed using data on enrollment at the department level between 2002 and 2011 merged with data on research quality of departments from the first REE accomplished in Italy (the VTR).
Italy is an interesting case study since enrollment has been decreasing in Italy in the last decade, especially in the South, so assessing the effect of research quality on the quantity and quality of enrolled students is important.
Our analysis demonstrates that increasing the percentage of excellent products by one standard deviation at the department level increases student enrollments by 6.5 percent. Effects are larger for high quality students, namely those with better high school final marks (10 percent) or coming from the academic track (11.8 percent).
Departments in the top quartile of the quality distribution gained more from a good performance in the evaluation exercise. Effect magnitudes appear to be similar across all macro-regions (North, Centre and South and Islands), but are precisely estimated only for universities in Northern Italy.
Finally, Chapter 3 presents and estimates a model of household behavior with endogenous labor supply and fertility choices. The estimated model is then used to assess the effect of a childbirth transfer on household decisions.
We contribute to the recent literature (Adda et al., 2015) performing ex-ante structural evaluations of policies having the objective to modify the fertility and labor supply behavior of households.
The model is estimated using the Italian Survey on Household Income and Wealth (SHIW) for the period 1984-2014, a dataset collected by the Bank of Italy every two years. The model parameters are estimated through the Method of Simulated Moments. We obtain moments from household in the 1960 cohort, i.e. people born in years 1957-1963.
Structural estimation offer some important advantages with respect to reduced form approaches. First, it allows to model different sources of endogeneity (ex. self-selection into labor market participation). Second, it provides parameters from a theoretical model that can be used to simulate the effects of policy experiments.
The model is able to explain quite well the behavior of men and women in the cohort. Preliminary results show that the permanent childbirth transfer is successful in increasing the total fertility rate of married women, even if it has a negative effect on employment.

Chapter 1

The effect of grading standards on labor market outcomes: evidence from Italy

ABSTRACT - We evaluate the effect of different grading standards at the department level on labor market performance of university graduates in the few years after graduation. The estimation strategy is divided in two steps. Firstly, we estimate a proxy for grading standards as the part of final grades which cannot be explained by differences in individual characteristics (student's quality) and other relevant inputs (quality of the institution attended). Then, the effect of grading standards on wage and other labor market outcomes is estimated. We show that differences in grading standards are large across departments. More generous grades are associated to a wage penalty on the labor market 5 years after graduation. In particular, graduates from 'generous' departments earn 3.4% less than people who studied in the 'strict' departments, they have a lower employment rate and a higher probability of being too educated for their jobs. The effects on wages are stronger for high ability workers while employment is more affected for low ability and female graduates.

Keywords. Higher education · Grading standards · Italy
JEL Classification Numbers: I2,J2 ,J3.

1.1 Introduction

The effect of grade inflation and grading standards on learning and post-graduation performance has been a debated issue in the last decades. Economic research suggests that differences in grading standards, defined as different grades to students with the same ability, could have several effects both on students' and on graduates' outcomes. Some studies suggest that, especially at the higher education level, different grading policies can lead to distortions in courses and majors choice (Sabot and Wakeman-Linn, 1991; Butcher et al., 2014), especially for women (Goldin, 2013). There is a large and increasing amount of evidence suggesting that some fields of study tend to grade consistently higher than others. Recent evidence (Butcher et al., 2014) finds out that by forcing high grading departments to stick to stricter standards leads to a reallocation of students between majors and modifies course choices in favor of the low grading fields. Differences in grading standards are also likely to lead to changes in effort and learning. A study on a public college (Babcock, 2010) in California found that in generously graded courses the level of effort of students is much lower, while research at the elementary school level (Figlio and Lucas, 2004) discovered that high grading standards benefit students' learning as measured by test scores. However, high standards might have a negative effect on high school graduation rates for some disadvantaged minorities (Betts and Grogger, 2003). Large differences in grading standards across institution could bias selection into graduate schools. A recent study (Bailey et al., 2014) demonstrates that American law schools do not fully take into account differences in grading standards while selecting applicants, penalizing students from colleges with a rigid grading policy.

Finally, grading standards can affect the performance on the labor market. The influence can occur through two different channels: a productivity effect and a signaling effect. The effect on workers' productivity is due to a lower human capital accumulation in institutions with the lower standards (see (Babcock, 2010; Figlio and Lucas, 2004)). We expect the employer to be not fully informed (or not informed at all) about the grading standard and about the relationship of it with productivity. For this reason we could expect, in line with the (Altonji and Pierret, 1997) model of employer's learning, the positive relationship between standard and productivity, at first unobserved by the employer, to become evident over time. In this case, the coefficient for the grading standard in a wage regression will be not significantly different from zero at the time of entry in the labor market and will become increasingly important over time.

The pure signaling effect is just relevant in the short term and it is more complicated to foresee, depending on the amount of information the employer has about the grading standards. If employers observe the final grade but have no information about different standards, we expect that the students graduating in the easy grading departments will be advantaged (at least in the short run). On the other hand, if employers have perfect information about both the final grade and the standard, the high ability graduates in the generously grading institutions will be penalized, since they will be pooled with some lower ability workers and will end up receiving lower wages or accepting less satisfactory job offers leading to lower wages in the medium to long run, if they are not able to switch immediately to higher quality jobs. In any case, wages will just depend on workers' productivity in the long term.

The lack of signaling value of final grades in Italy has been a discussed topic especially after the introduction of the 3+2 system, which led to a compression towards the top of the distribution for 2nd cycle degrees. In figure 1 we plot [1] the final grade of graduates in unique-cycle degrees (usually 5 years) and master degrees (2 years after another 3-years university degree) against final mark in high school, that we use as a rough proxy of ability. We can see that the positive relationship between high school mark and final grade in university is much stronger in long degrees than in master degrees.

Despite the public debate on grade inflation and grading standards has been quite heated in many countries in the last decades, there is very little evidence linking differences in grading standards while in education and labor market outcomes. Up to our knowledge, the only empirical papers trying to study the same research question are (Betts and Grogger, 2003) and (Bagues et al., 2008). (Betts and Grogger, 2003) study the relationship between grading standards in high school, students achievement, educational attainment and entry level earnings. They find that higher standards positively affect test scores but the increase is greater for high-ability students and the effect on earnings turns out to be positive even if it is limited to high school graduates who did not enroll in college.

For the Italian case, (Bagues et al., 2008) find that large differences in grading standards exist between universities and fields of study and that more generous grading by

[1] Using the Almalaurea dataset on graduates in years 2008 and 2009.

departments tend to be associated with poorer performance after graduation.

Table 1.1: Literature review

Authors	Sample	Methodology	Results
Figlio and Lucas (2004)	Administrative data including every third, fourth, and fifth grader in a large school district (Alachua County, Florida) over four years	Regression with individual and school level FE	High teacher-level grading standards tend to have large, positive impacts on student test score gains in mathematics and reading. High standards also reduce student disciplinary problems in school
Babcock (2010)	12-quarter panel of course evaluations from the University of California, San Diego	Regression with FEs	Results indicate that average study time would be about 50% lower in a class in which the average expected grade was an A than in the same course taught by the same instructor in which students expected a C. Findings do not appear to be driven primarily by the individual student's expected grade, but by the average expected grade of others in the class.
Butcher et al. (2014)	Transcript-level data on student grades and courses from Fall 1998 to Spring 2008 at Wellesley College	Experimental design: some departments received treatment (pushed to decrease average grades) while others did not	Choice of treated majors declined
Betts and Grogger (2003)	High school graduates in the US (High school and beyond survey)	OLS and quantile regression	Positive effect of Math GS on log earnings
Bagues et al. (2008)	Italian survey on early career of university graduates	Regression	Large differences in GS, possible negative effects on the labor market of generous grading

In this paper we estimate how differences in the relative generosity of academic departments in awarding final marks affect the graduates' early performance on the labor market. In particular we want to evaluate if, after controlling for differences in a large set of individual-level and department-level variables, there is an effect of grading standards on employment rates and earnings. To do so, we use the data on labor market performance of university graduates collected by Almalaurea, a consortium of Italian universities. This is one of the two main sources of information available on the subject in Italy. We apply a three-step methodology similar to the one in (Brunello and Cappellari, 2008). First of all, we estimate the effect of each department on students' final marks after controlling for differences in observable individual-level variables. We use the data-set to run a linear regression of final grade in long university degrees (both unique-cycle and master degrees) on a large set of graduates' characteristics and department level dummies. In a second step, we use the department-level coefficients estimated in the first regression as a dependent variable and we control for differences in departments' characteristics. We exploited a large data-set collecting information on funding, research quality and human resources at the department-level. This step is necessary in order to exclude the differences in employment and wages that are driven by the quality of the institution attended. The residual of this second regression will be our proxy for the grading standard. Finally, we include our measure of grading standard in wage and employment regressions, together with many other control variables, in order to investigate the effect of interest.

This paper makes some important contributions with respect to the previous literature. First of all, we can rely on recent department-level data on research quality (Valutazione Qualita' della Ricerca) for the period 2004-2010 collected by the National Agency for the Evaluation of the University System (ANVUR) and we integrate them with data on financial and human resources provided by the Ministry of Higher Education and Research (MIUR). We make use of the best set of controls for department quality currently available and this constitutes a significant improvement with respect to (Bagues et al., 2008). Furthermore, most studies on Italian higher education make use of the data provided by the Italian Institute of Statistics (Istat), while we use a different source of data: Almalaurea. The dataset provided by Almalaurea allows us to track changes in earnings over time and provides larger sample sizes for each department, resulting in more precise estimates. Another important contribution of this paper is the ability to check the difference in the relationship studied between long degrees (single-cycle) and the 2nd cycle degrees taken under the 3+2 regime. Finally, we try to address a selection problem. In departments where grading standards are rigid we expect a higher dropout rate. To check for differences in dropout rates, we add a proxy for dropout in the wage and employment equations. The structure of the paper is as follows. The next section sets the empirical model and the conceptual framework for our analysis. Section 3 describes the data used for the estimation. The main results are outlined in Section 4. In particular, we estimate separately the effects of grading standards on earnings, overeducation and employment and we present some robustness checks. Section 5 summarizes our main findings and concludes.

1.2 The empirical model

We use an estimation strategy similar to the one in Brunello and Cappellari (2008). In the first step, we regress final grade on individual-level control variables and department-level dummies. Let us define the first equation of interest, the determinants of final grade (m_{ij}), as

$$m_{ij} = \beta_0 + \beta_1 X_i + \sum_{j=1}^{J} \beta_j D_j + \epsilon_{ij} \tag{1.1}$$

where i and j are individual and department subscripts respectively; X_i a set of individual characteristics influencing final grade; D_j dummy variables for each department; and u_{ij} an error term.

In a second equation we regress the estimated department-level coefficients in the first step, $\hat{\beta}_j$, on a set of controls for the quality of each department. We define the equation as:

$$\hat{\beta}_j = \gamma_0 + \gamma_1 Z_j + u_j \tag{1.2}$$

where $\hat{\beta}_j$ are the department coefficients from equation 1, Z_j is the set of quality variables for each department; and u_j an error term that will be our proxy measure for grading standards. In another version of this second step, we also add field of study dummies in order to take into account differences in grading practices across subjects:

$$\hat{\beta}_j = \gamma_0 + \gamma_1 Z_j + \gamma_2 field_j + u_j \tag{1.3}$$

This second version gives us the within-field variation in grading across departments. In the third and last step, we estimate the effect of our measure of grading standard $\hat{u}_j$ on wages, employment and overeducation:

$$Y_{ij} = \delta_0 + \delta_1 W_i + \delta_2 m_i + \delta_3 Reg_i + \delta_4 field_i + \delta_5 \hat{u}_j + \delta_6 \hat{\gamma}_1 Z_j + e_{ij} \tag{1.4}$$

controlling for a large number of individual characteristics W_i, local labor market effects Reg_i, a set of dummy variables for the field of graduation $field_i$, the final grade m_i and the predicted values of the department-level controls for quality and resources that we used in the second step $\hat{\gamma}_1 Z_j$. Local labor market effects are measured at the regional level. [2] We will interpret a statistically significant δ_5 as evidence for an effect of grading standards on labor market performance. We check for the robustness of our results by adding a measure of the dropout rate and the residual of the first step final grade equation. Our final specification will be:

$$Y_{ij} = \delta_0 + \delta_1 W_i + \delta_2 m_i + \delta_3 Reg_i + \delta_4 field_i + \delta_5 \hat{u}_j + \delta_6 \hat{\gamma}_1 Z_j + \delta_7 dropout_j + \delta_8 \hat{\epsilon_{ij}} + e_{ij} \tag{1.5}$$

Controlling for the dropout rate is necessary because in academic departments with relatively more generous grading policies the dropout rate is expected to be lower so that more people are expected to get a degree. For this reason the average quality of university graduates could be lower in institutions with less demanding standards since more students survive the selection process. The lack of controls for this phenomenon could lead to a bias in our coefficient of interest. The error term from the first step contains everything affecting the final grade that is not measured by the controls and by the department-level

[2] Results of regressions with province-level effects are not presented here but are very similar and available upon request.

fixed effects. In $\hat{\epsilon_{ij}}$ we have, for example, individual motivation or within-department differences in grading practices. If we do not include this term in the labor market outcome equation then we could have an omitted variable problem that could in theory bias the coefficient for the grading standard.

1.3 Data

Our analysis relies on a number of data sources, namely, Almalaurea microdata on university graduates, and department-level information collected by the Italian Ministry of Education and Research (MIUR) and the Italian Agency for the Evaluation of the University System (ANVUR). In this section we provide relevant details on these various sources.

1.3.1 Almalaurea

The most widely used dataset in empirical research on university graduates in Italy are the microdata provided by the National Institute of Statistics (ISTAT). However, the ISTAT dataset presents some important limitations for answering the research question of interest. First of all, the precise information we need about the final mark at university is missing. In order to avoid the problems we would have with the ISTAT dataset we perform the analysis using the microdata collected by Almalaurea. This second source has a lot of advantages. Since Almalaurea collects survey and administrative data on the whole population of graduates in the associated universities, we have larger samples available for each department, leading to more precise estimates. This is important when the object of interest are department-level effects, like in this case. Furthermore, while in the Istat dataset we just have information on labor market outcomes 4 years after graduation, Almalaurea collects information on employment and earnings 1, 3 and 5 years after graduation making possible to compare the effects of grading standards at different times. On the other hand, information is available just for the universities belonging to the Almalaurea consortium (48 for year 2008) and geographical location of unemployed is not updated, making more difficult to credibly estimate the effects of standards on employment. We exploit 2 waves of data on university graduates, focusing on students graduated in 2008 and 2009. These are the more recent waves for which we have complete histories on labor market outcomes until 5 years from graduation. Before starting the analysis we select the sample that we will use in the first step. The results of the procedure can be seen in Table 2. Since most students in Italy decide to continue their studies after their Bachelor degree, we decided to focus on graduates from long degrees, so we excluded all student graduated in first-cycle 3-years degrees. The remaining sample is split between graduates in single-cycle degrees (4 to 6 years) and second-cycle degrees (2 years master after the first cycle). We have a total of 115009 graduates in years 2008 and 2009 in the sample. We decided to drop graduates with missing information on the control variables used in step 1, namely the type of high school attended, final mark in high school, parents' education and distance from university. We then dropped graduates aged more than 35 and from defense and physical education. The last exclusion comes from the fact that we cannot control for quality of the department attended in the second step. Furthermore, these fields of study include a relatively small share of total graduates. The final sample for regression (1) consists of 93821 observations, summary statistics for the main variables used can be seen in table 4. More than 99% of graduates in the sample are Italian citizens. This number underlines three facts. The first one is the relative lack of attractiveness of Italian universities with respect to those in some other European countries. Secondly, the universities associated to Almalaurea in the years considered do not include some of the best institutions in the North of Italy, which are the more appealing ones for foreign born students. Finally, non-Italian students are much more likely to have missing information for the variables we used in the selection process.

The sample consists for almost the 62% of females and the average high school mark is 86.6 out of 100. The share of students spending at least part of their studies abroad is close to 16%, while just 5.4% worked regularly while studying. Around two-thirds of the graduates come from academic high schools (classical studies and scientific tracks) and more than one third are from families where at least one parent got a university

Table 1.2: Sample selection

Sample step 1	Observations
1. Original sample	115009
2. Drop high school information missing	111799
3. Drop age above 35	105089
4. Drop defence	104095
5. Drop missing parents' education	94343
6. Drop close to university missing	93821
Sample step 3	
1. Final sample from step 1	93821
2. Drop unmatched with step 2	92200
3. Drop post-graduate education	73171
4. Final sample for wage (and overeduc) equation	21948
5. Final sample for employment equation	44259

degree (ISCED 5-6). For the third step of our analysis, i.e. the wage, overeducation and employment regressions, we carry out a further selection detailed in the second part of table 2. First of all, we drop students that was not possible to merge with the results of the regression in step 2 because of missing information on the specific department (universityXfield) of graduation. Then we drop students who decided to proceed with long postgraduate studies (e.g. PhDs) because their earnings in the first 5 years after graduation cannot be seen as a good proxy of their value on the market. This leaves us with 73171 observations. Then we keep observations with non-missing information on the outcome variables in each of the three post-graduation periods. We end up with 21948 observations for the wage and overeducation regressions and 44259 observations for the employment regressions.

Summary statistics for graduates who reported positive earnings 1, 3 and 5 years after graduation are presented in table 4. Monthly wages are left censored at 200 euros and right censored at 3250 euros. The average monthly wage increases from 1086.54 euros 1 year after graduation to 1415.93 euros 5 years after graduation. The employment rate (not reported) is slightly above 84% 3 years after graduation and increases to 88% 2 years later. Around 30% of the sample reports to feel overeducated for their job 5 years after graduation. The two measures of the standard reported in the table are calculated as the residuals of regressions 2 and 3 respectively.

1.3.2 Department-level information

In order to control for the quality of department in the second step we exploit a dataset produced by (Bratti et al., 2015). The data combine a large amount of information on research quality, financial and human resources available to each university department. In the first stage we defined as departments the university-field of study combinations. Research quality is measured by the Research Quality Assessment (Valutazione della Qualita' della Ricerca, (ANVUR, 2013)) for the period 2004-2010, in particular by a general indicator of research quality and by a variable measuring the quality, in terms of research output, of people hired and promoted in each department. The amount of competitive funding received by the government can be seen as a measure of both research quality and resources, while as a pure indicator of financial resources we make use of average yearly tuition fees paid by the enrolled students[3]. Total funding received by the department is used as a measure of size. Finally, we include a measure of human resources, the student-teacher ratio, and a measure of attractiveness, namely the share of students coming from the same province and the share from abroad in each department. Summary statistics are presented in Table 4.

We are able to match department-level effects on final grade with the data on quality for 354 departments. We can notice a very large variation in resources across departments. For example, the research quality indicator from the RQA ranges from 0.12 (lowest quality) to 0.98 (highest quality). The same is true for the average research performance of new hirings and promoted staff. Regarding resources, the average annual fees amount to 1069.6 euros per academic year, again with values ranging from 394 euros to 6046 euros. Student-teacher ratio, measured here as the number of people graduated over the number of professors in the department in a given year, ranges from 0.095 to more than 26.

1.4 Estimation results

In this section we summarize the main results of the empirical analysis.

[3]This is just available at the university level

1.4.1 Estimation of department-effects on grades

Results for the regression in equation (1) are presented in table 5. We run a simple linear regression of the final grade on a large number of controls for individual ability, family background, the region of origin and a set of dummy variables for each department. Heteroskedasticity robust standard errors are used. Departments are defined as university times field of study interactions. In the ISTAT classification 14 fields are present: Hard Sciences, Chemistry, Biology, Medicine, Engineering, Architecture, Agricultural Sciences, Economics and Statistics, Political Sciences, Law, Humanities, Languages, Teaching and Psychology. Since we need to match the subjects of study with indicators of research quality, we pool Humanities with Languages and Teaching with Psychology, reducing the number of fields from 14 to 12. We present the regressions with 12 fields in Table 5, the results do not change if the original 14 fields in the ISTAT classification are used.

In columns 1 and 2 we run regressions including only university and field of study effects on final grade respectively. In column 3 we add both university and field effects. Finally, in columns 4 and 5 we use department-level effects (universityXfield). A substantial amount of the variance in final grades can be explained by differences across departments. The coefficients in the table are quite similar across specifications and have the expected sign. In particular we observe a significant and positive coefficient for female graduates, high school final mark, regular attendance of courses and spending part of the studies abroad. On the other hand, work regularly during university studies and graduating late are associated with lower final grades. Second-cycle degrees (2-years long) tend to grade much more generously, more than 3 points in the complete specification in column 4, than single-cycle and pre-reform degrees, even after controlling for individual ability and background. An interesting finding is the negative association between final grade and parents' level of education. It could be partly explained by the fact that, in Italy, people from a low socio-economic background rarely complete university studies. Then, the few who reach university graduation are likely to be very smart and motivated. Controls for the region of origin associate lower grades to students coming from the south. This is probably due to a large heterogeneity in the quality of the school system over the national territory, as confirmed by international standardized tests (e.g. the PISA administered by OECD). The distribution of the fixed effects estimated in column 4 is presented in the first graph of figure 2. There are sizable differences in the departments' fixed effects on final grades, ranging from a negative coefficient of -7.97 (meaning that, other things equal, students in that department receive a final grade nearly 8 points below the reference value) to a positive coefficient of 5.29. We extract these effects to proceed with step 2.

1.4.2 Control for department characteristics

In the second step of our analysis we use as a dependent variable the estimated department-level coefficients and we run a regression on some controls for quality and resources. This is necessary since the effect of each department on final grade could be partly explained by differences in characteristics. A department with more resources and high quality professors could be better at teaching to students leading, other things equal, to higher grades. We use weights obtained calculating the inverse of the standard errors of the department-level coefficients estimated in equation (1).

The results are presented in table 6. In column 1 we present the baseline specification of equation (2), while in column 2 we add field of study effects as in equation (3). Departments with a better research environment, as measured by the quality of recently hired researchers, a general indicator of research quality and the amount of research funding received from competitive procedures, tend to award lower final grades to their students. The coefficient for the quality of hired and promoted staff decreases when we add field of study controls, however it remains significant at the 10% level. The other coefficients for research quality are also significant in the second specification and increase when moving from the first to the second column. Since direct measures of the quality of teaching are not available, the result might be due to the fact that better researchers spend less time teaching[4] so they are responsible for the relatively bad outcome of their students. This is unlikely to be generally true in the Italian case, we think a better explanation

[4]or are just worst teachers.

for the phenomenon is given by (Braga et al., 2014). They link students' evaluations of professors to teacher effectiveness, measured by grades in related future exams. The main finding of their paper is that lower grades are often not a proxy of poor teaching but of more demanding professors.

Variables indicating the amount of financial resources received, like average yearly fees per student and amount of funding from competitive procedures received, have a negative but not significant effect on final grades in column (1) while they get negative and significant adding field of study controls in column (2). The size of the departments, measured by the total amount of financial resources received, turns out to be irrelevant in explaining differences in grading practices. We proxy the quantity of human resources available by including the student-teacher ratio[5] as a control. When the ratio increases we observe a reduction in final grades. The effect is large and significant, even if decreasing, in the first specification but it gets much smaller and insignificant when we move to the second specification.

Finally, the share of foreign students, that is a measure of attractiveness of the department, is negatively related to grades, significantly in column (2). On the other hand, the larger the share of students from the same province in a given department, the more generous the department will be in awarding final grades. This finding reinforces the idea, suggested by the coefficients on research quality and average fees, that better departments might have, on average, more rigid grading standards. Adding field of study controls leads to a large increase in the regression R^2, meaning that a quarter of the variation in the department-level effects on final grade can be explained by heterogeneity in grading practices between fields. The results may simply be driven by a supply and demand explanation, departments with an excess of demand for students' enrollment will try to reach an equilibrium between enrolled students and resources available by making the grading standards more tough, while the opposite will happen where the demand for enrollment is weak.

We leave a more detailed explanation of the differences in grading practices across insti-

[5]It is given by the number of full-time students over the number of researchers and professors in each department.

tutions to future research since it is not the aim of this paper. We will use as a proxy for grading standards the residuals of the 2 regressions[6]. In the second and third graphs of Figure 2 we show the distribution of the two residuals obtained. The regional variation in grading standards is presented in Figure 3. The map shows that the standards are quite heterogeneous going from one region to another. Since high (tough) standards are associated to light shades of red, the regions with low (less demanding) grading standards in both specifications are Tuscany, Puglia, Sardinia, Molise and Basilicata. The result can be partly explained by (De Paola and Scoppa, 2007). They suggest that setting lower standards might be optimal in distorted labor markets, as is the case in the South of Italy. Large variations between fields of study, as documented by Figure 4, also exist, mostly confirming the results in the international literature. The fields keeping the highest standards are Chemistry-Pharmacy, Engineering, Economics-Statistics and Law, while the low standards fields are Social Sciences, Humanities, Medicine and Biology. However, the value found for Medicine may be at least partly explained by the fact that the admission to the degree is very selective, so student could be much more motivated or have a higher level of unobserved ability[7]. In Figure 5, we present the average standard for each university in the Almalaurea consortium. On the horizontal axis we have the ISTAT code for each university, that is increasing going from the North to the South of Italy. Except for an outlier, S. Raffaele University in Milan, with a very high positive value above 4, universities in the Centre-North of Italy tend to give lower grades than universities in the Centre-South. However, there are also sizable differences between universities within the same macro-area.

Finally, in Figure 6 we look at the distribution of grading standards between departments in the same aggregated field. We group the 12 fields of study in 6 areas (Science, Engineering, Medicine, Social Sciences, Law and Humanities) finding that there is a large amount of heterogeneity between departments within each area.

[6]Since the choice of the variables included as controls in the two regressions is somewhat arbitrary and the regressors are often highly correlated, we run the same regressions using the scores obtained from a Principal Component Analysis. The results are very similar.

[7]For this reason, as a robustness check, we will run the step 3 regressions excluding graduates in Medicine, as done in (Bagues et al., 2008). The results are almost identical to the ones that we show in the next section.

1.4.3 Labor market outcomes

The aim of this section is to present the main results on the relationship between standards and labor market outcomes. We start by showing the effect on earnings, then we analyze the impact on over-education and employment, finally we perform some robustness checks. We use bootstrapped standard errors at this stage[8]. The grading standard is re-defined as a 4 quartiles variable in order to check for non-linearities in the relationship between the variable of interest and the outcomes.

Earnings

Results of the wage equations are presented in Tables 7-9. In Table 7 the outcome variable is (log) monthly wage 5 years from graduation, while in tables 8 and 9 the dependent variables are (log) monthly wage 3 years and 1 year from graduation respectively. In column (1) and (4) we find the estimated coefficients from equation (4) for the grading standards found as the residuals in (2)-(3), while in columns (2) and (5) we estimate equation (5) for the two measures of standard. In columns (3) and (6) we run again the regression in (5) but just for graduates in 2-years master degrees. We will focus on the specification in column (5) since it is the more complete, incorporating both the more detailed measure of the standard and all the controls. However, the coefficients are quite similar across specifications. Having data for the same people over time enables us to study the variation in the coefficients of the variables. Final grades and parents' education have a positive effect on earnings that increases over time. In particular the effect of final grade on monthly wages ranges from a 0.2% increase per point 1 year after graduation to a 0.8% increase 5 years after graduation. While the growing relevance of parents' education in affecting wages over time is not surprising, since it is used in many studies [9] as a proxy for unobserved productivity, the increase in size of the coefficient on final grade is harder to explain. A possible explanation for the finding is that the employers do not consider the final grade as a reliable measure of quality in the selection process, so they

[8] 500 replications

[9] Altonji and Pierret (2001) being the most famous example

tend to evaluate the prospective employees based on other observables characteristics. Females have a 12.5% wage penalty 1 year after graduation that increases to 15.9% 4 years later. A similar pattern emerges for the unobserved heterogeneity in final grades [10] that is negative but insignificant 1 year from graduation but it gets increasingly negative and significant in the medium term. This fact makes clear that the error term contains more information about differences in grading practices within departments than about unobserved ability. The negative coefficient for year 2009 tells us that graduates in 2009 received lower wages, at least at the beginning of their career, with respect to the cohort of students graduated in 2008, due to the worsening of economic conditions. Finally, we observe a significantly positive coefficient, even though decreasing over time, for public sector employment, work on a regular basis during university studies and pre-reform degrees. This finding can be easily explained by the fact that these graduates entered the labor market before others, so they experienced higher wages right after graduation but this advantage starts disappearing when the other graduates gain labor market experience. Controls for field of study and region of employment give the expected results. We now turn to analyze the coefficient of interest. We divide the grading standard in 4 quartiles in order to check for non-linearities in the relationship with wages [11]. Using the first quartile as a reference, we see that there is no effect of the standard on wages 1 year after graduation. In fact, the coefficients on quartiles 2-4 are very close to 0 and insignificant. The standard gets increasingly important over time, 3 years after graduation the wage penalty is 2.2% for workers in the fourth quartile with respect to workers in the first quartile of the distribution. The penalty further grows to 3.4% after 2 years going from the first to the last quartile. The effect appears to be non-linear, with most of it taking place between the first and the second quartile. The results do not change much if we consider the alternative measure of standard or if restrict our analysis to second cycle degrees. It is unclear if the increasing importance of the grading standard as a determinant of wages is just due to a process of learning about productivity by the employer or if the workers tend to select themselves from the beginning in different occupations, with graduates

[10] the error term of regression (1).

[11] the transformation also makes our results easier to interpret.

from low standards departments clustering in lower quality jobs, and experience different wage growth afterward. In order to answer this question we turn, in the next section, to the analysis of how grading practices affect the probability of being over-educated for the job. Note that here we used a relatively restricted sample including graduates with non-missing and positive wages for all the 3 periods of the survey. If we run the same regressions with larger, less selected, samples we get the same dynamics for the coefficients of grading standards. They are still significantly negative and even stronger 5 years from graduation. This means that the effects found in table 7 are a lower bound and probably the grading standard affects wages also through a lower accumulation of labor market experience in the first years after getting the degree. We will check this hypothesis by looking at the effect of the grading standard on employment in section 4.3.3.

Overeducation

The results for the regression of overeducation on the grading standard and other controls are presented in tables 10-11[12]. In table 10 we use overeducation 5 years after graduation as a dependent variable, while in table 11 we study the effect of grading standards on overeducation 1 year after graduation. The results are very similar in the two tables, showing a very persistent effect of the grading standard on overeducation. While an increase in final grade has the intuitive effect of reducing the probability of being too educated for the job, we see that moving from tough to easy grading the probability of overeducation increases significantly. In fact, a 1 point increase in the final grade reduces the probability of being overeducated for the job by slightly more than 1 percentage point in both tables, while moving from the first to the last quartile of the grading standard distribution leads to a 6 percentage points increase in overeducation in our preferred specification (column 5) of table 10 and to a 7.3 p.p. increase in the same column of table 11. This result is interesting because the high persistence of overeducation over time could point to the fact that young workers, due to a lack of job opportunities in Italy caused by the financial crisis, remained often trapped in relatively low quality jobs.

[12] We run Linear probability models here, results from Probit regressions are very similar and available upon request.

Since we find a very strong effect of the grading standard already 1 year after graduation, we could argue that employers understand, at least partly, from the beginning that some workers coming from 'easy' degrees might be less productive and select them in lower quality occupations. In the following years graduates in higher quality jobs experience a faster wage growth and this contributes explain our finding on the increasing wage gap over time between different quartiles of the standard.

Employment

The main results for the regressions of employment on the grading standard are shown in tables 12 to 14. The effect of the control variables is very similar to the one seen in wage regressions. Focusing on the specification in column 3, one year after graduation the coefficients for all the quartiles below the first are negative, with the stronger effect, a 2.8 percentage points decrease in employment, taking place at the third quartile. As for wages, the grading standard coefficients grow over time, five years after graduation the negative effect on employment ranges from 2.6 p.p. in quartile 2 to 3.4 p.p. in quartile 4 compared to the first quartile. The results are almost unchanged if we use the other measure of standard or if we just consider graduates in second cycle degrees.

Robustness checks

We now study the robustness of the results obtained in different sub-samples. We consider the effect of the grading standard on (log) monthly wages 5 years from graduation in Tables 15-16 and on employment 5 years from graduation in Table 17. For the wage regressions we split the sample in employees and self-employed, workers in the public and private sector, males and females, high and low ability. The first result worth noting is that the effect of the grading standard on wages is negative and significant for both employees and self-employed and it seems to be larger for the second group. The finding could be explained by the positive effect of the standard on individual productivity, that is clearly more important (if not the only input) in determining the earnings of self-employed workers. While the grading standard has no effect on workers in the public sector, the coefficient for moving from the first to the fourth quartile of the distribution is negative but not significant, it has a strong negative effect on workers in the private sector, wages decrease by 3.5% - 5.2% (depending on the measure of standard considered) moving from the first to the last quartile. The effect appears to be stronger for high ability (wage penalty of 4.9% - 5.2% going from the first to the last quartile) than for low ability (penalty of 0.4% - 2.7% from quartile 1 to quartile 4) graduates. Considering the effects of standard on employment 5 years from graduation we find a stronger effect for females than for males and for low ability workers[13]. For females we also find that parents' education is also an important determinant of employment decisions while it is irrelevant for males. Putting the two results together, it is probably the case that while men decide to work anyway women stay in the labor market only if they are relatively more productive, with the grading standard and parents' education being both a proxy for productivity. The effect of the standard turns out to be quite large for females, since the probability of being employed decreases of 7.5 percentage points moving from the first to the fourth quartile. Finally, the coefficients tend to be more negative for low ability (employment decreases by 5.9 p.p. going from quartile 1 to 4) than for high ability graduates (employment decreases by 2.9 p.p. going from quartile 1 to 4). This last finding is opposite to what we found for earnings. One possible explanation is that while high ability university graduates tend to be employed with probability close to 1 five years after graduation, the low ability are more likely to remain unemployed. On the other hand, the positive effect of tough grading standard on wages is stronger for high ability graduates because they benefit the most from a challenging environment at university.

1.5 Concluding remarks

The aim of the paper is to estimate the effect of grading standards on labor market performance in the first years after university graduation. Using Almalaurea data for university graduates in 2008 and 2009, with data on department-level characteristics, we

[13] We define low (and high) ability workers as graduates with a final grade below (or above) the mean in their department.

found, in line with (Bagues et al., 2008), that there are indeed large differentials in grading standards across departments. We found that lower GS (i.e. more generous grades) are negatively associated with both employment and earnings. Furthermore, workers from departments with low grading standards are significantly more likely to feel over-educated for their job. Moving from a department in the first quartile to one in the fourth quartile of the grading standard distribution leads to a 3.4% decrease in monthly earnings, a 3.4 p.p. decrease in employment and a 6 p.p. increase in over-education 5 years from graduation. The results are robust considering different samples and specifications. We find that the effect of the grading standard on wages is stronger for females and high ability graduates, while the effect on employment is stronger for females and low ability graduates. Since the effect of the grading standard on wages and employment is negligible 1 year after graduation and gets increasingly negative over time, there are reasons to believe that the employers gradually discover the differences in productivity of their employees. However, the positive and significant coefficients for the effect of the GS on over-education 1 year from graduation point to the fact that employers have some rational expectations about differences in future productivity of graduates coming from departments with different grading practices.

Table 1.3: List of the variables used in the analysis

Group	Variables included
m: final grade	final mark in college
X: demographics, study career and family background	gender, Italian nationality, attendance, study abroad (erasmus or similar program), worked during studies, distance to college dummy, High School final mark, High School type, delay in university graduation, parents' education, degree type, year (2008 vs 2009), region of origin.
D: departments	dummies for graduation in a specific UniversityXfield department
Z: controls for department-level characteristics	research quality indicator, quality of hired/promoted staff, funding from competitive procedures, average fees, total financial resources, student-teacher ratio, share of foreign students, share of students from the same province.
field: field of study	dummies for the 12 fields of study considered
$\hat{\beta}$: dependent variable step 2	estimated coefficient of department fixed effect on final grade in step 1
$\hat{u}$: grading standard	Estimated residual from equations (2) and (3). It is the part of that remains unexplained afer controlling for department-level characteristics
Y: outcome variables 3rd step	(log) monthly wage, overeducation and full-time employment
W: controls 3rd step	For all Y: sex, year, parents' education, work during studies, degree type. In the wage and overeducation regressions also add controls for employees (vs self-employed) and working in the public sector.
Reg: region (NUTS-2)	Region of work for wage and overeducation regressions. For employment regressions both the region of study and the region of residence are included.
$\hat{\gamma_1} Z_j$	predicted measure of department quality
dropout: dropout rate	estimated dropout rate for departments (from MIUR data on enrollment).
$\hat{\epsilon_{ij}}$: error term 1st step	unobserved variation in final grades

Figure 1.1: Plot of final grade in high school and at university for unique and 2nd cycle degrees

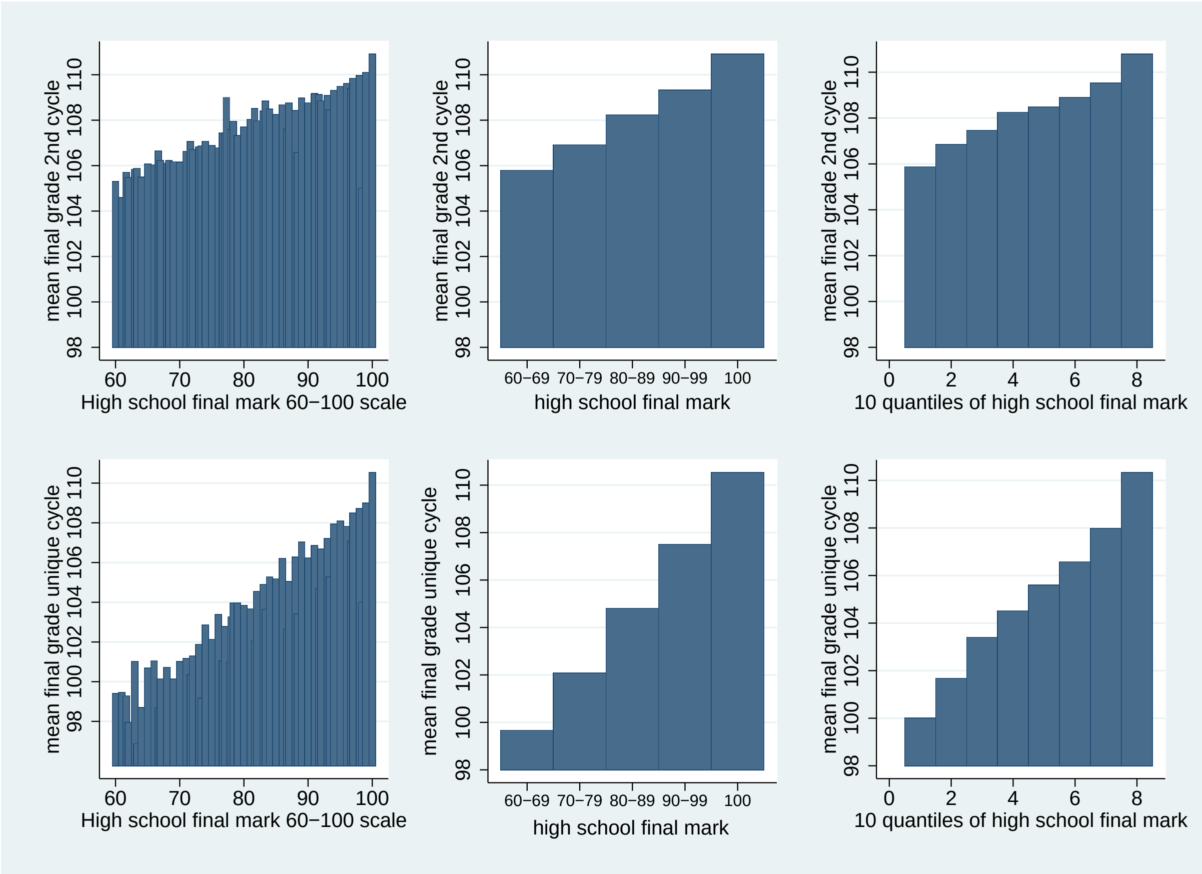
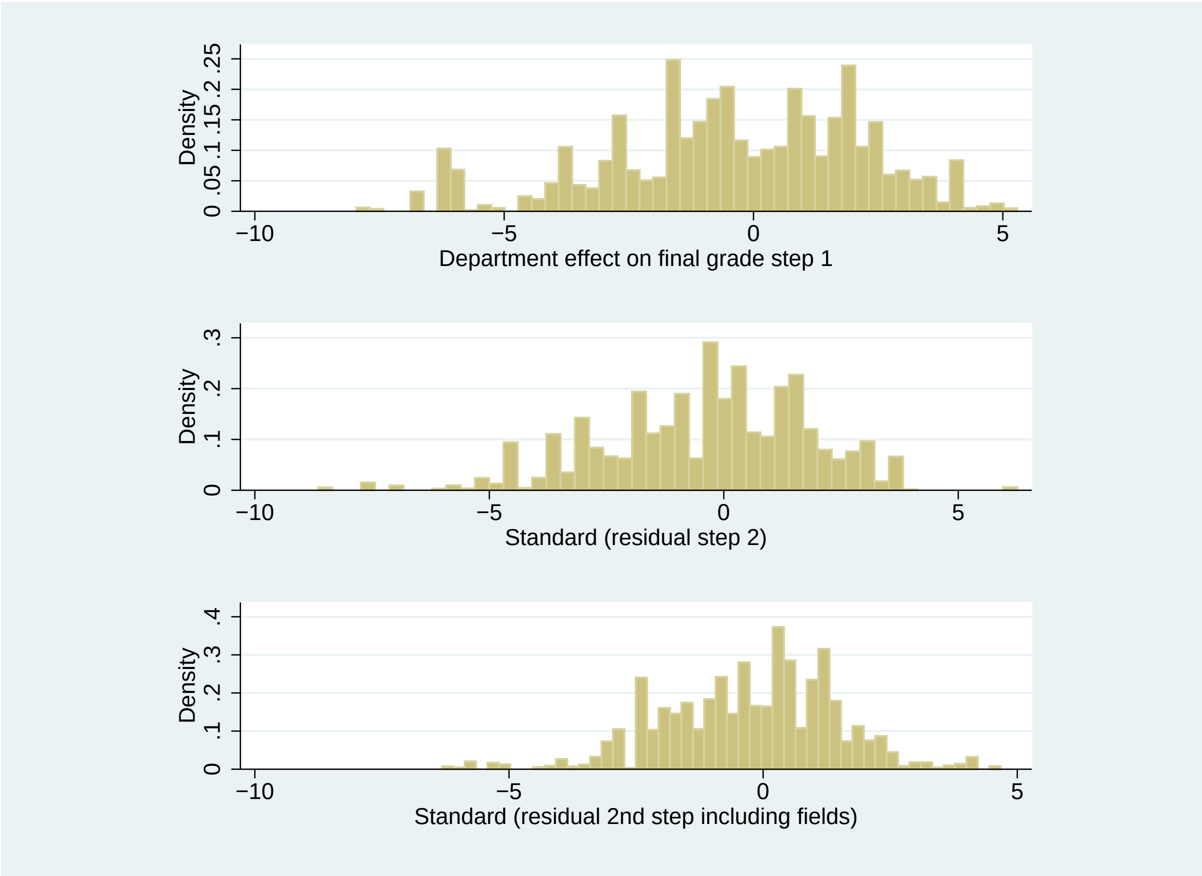
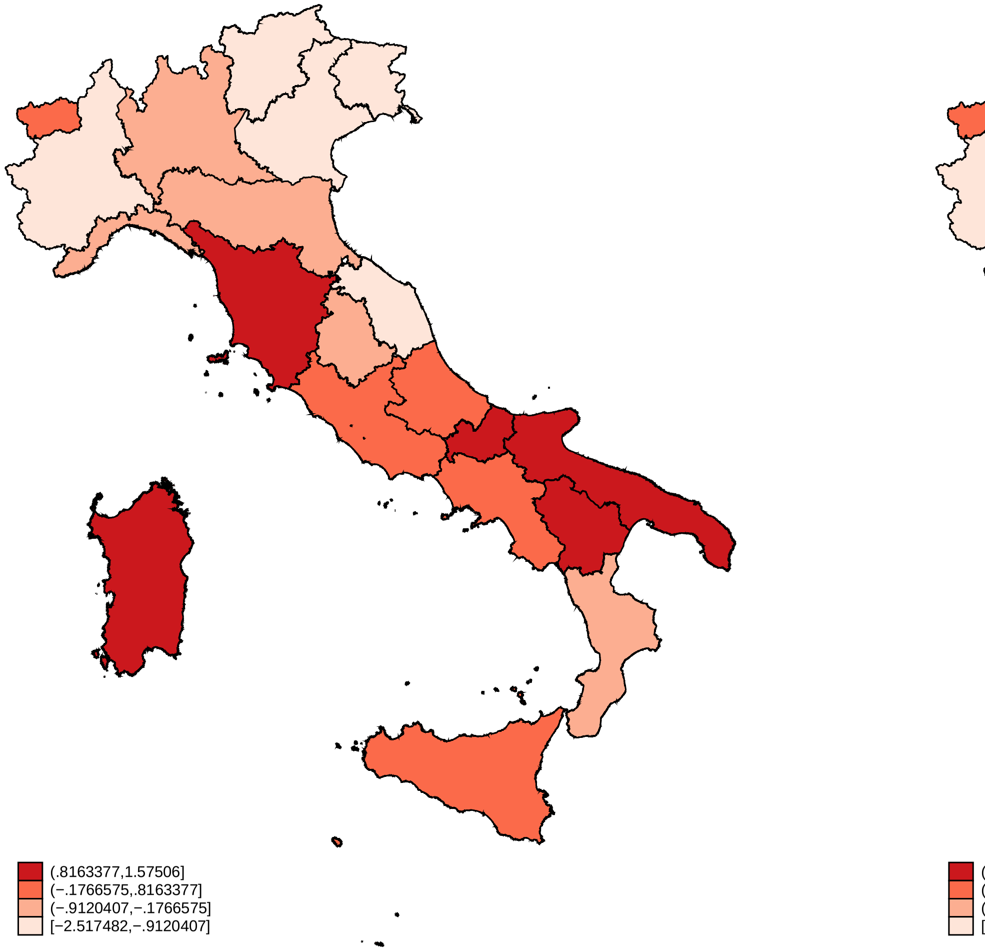


Figure 1.2: Histograms for department FE and the two measures of grading standard





(a) Standard 1

(b) Standard 2

Figure 1.3: plots of regional variation in the 2 standards

Figure 1.4: Variation in standards by field of study

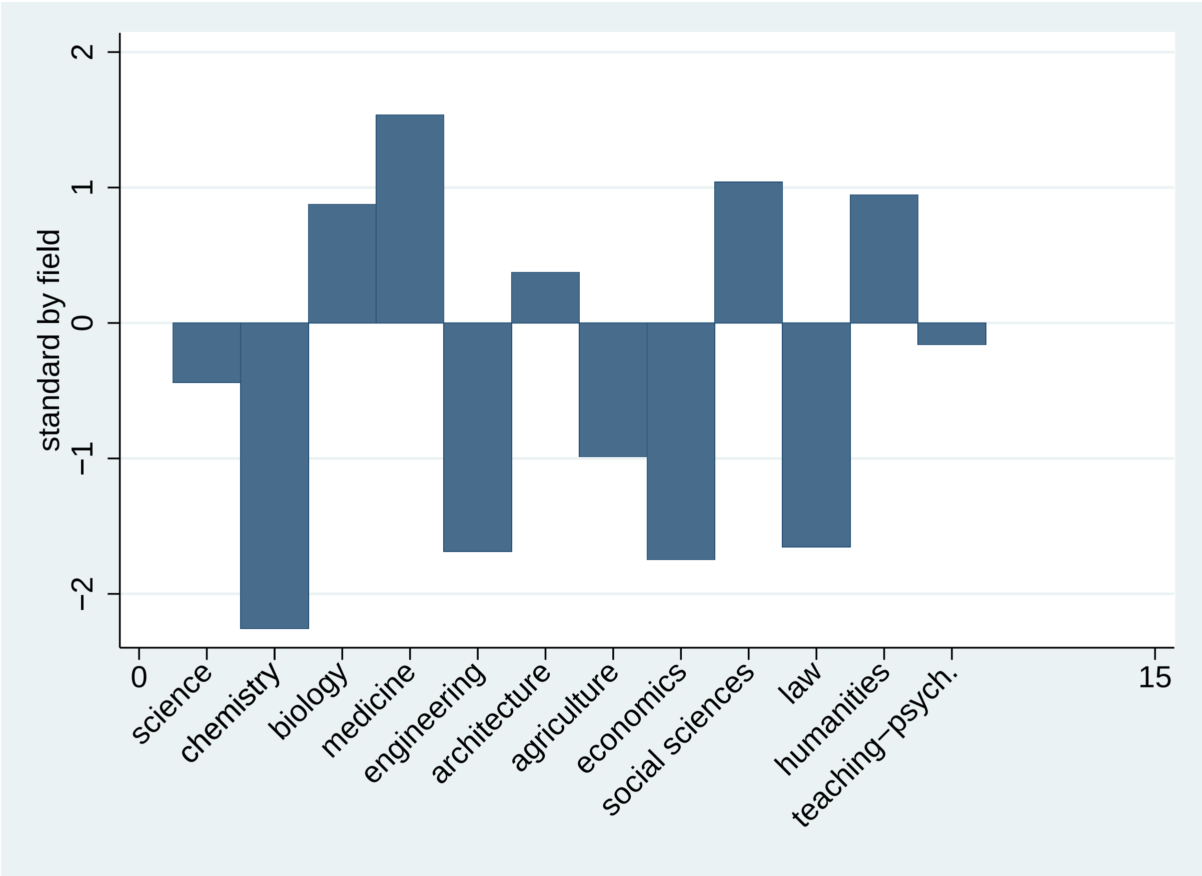


Figure 1.5: Variation in standards by university

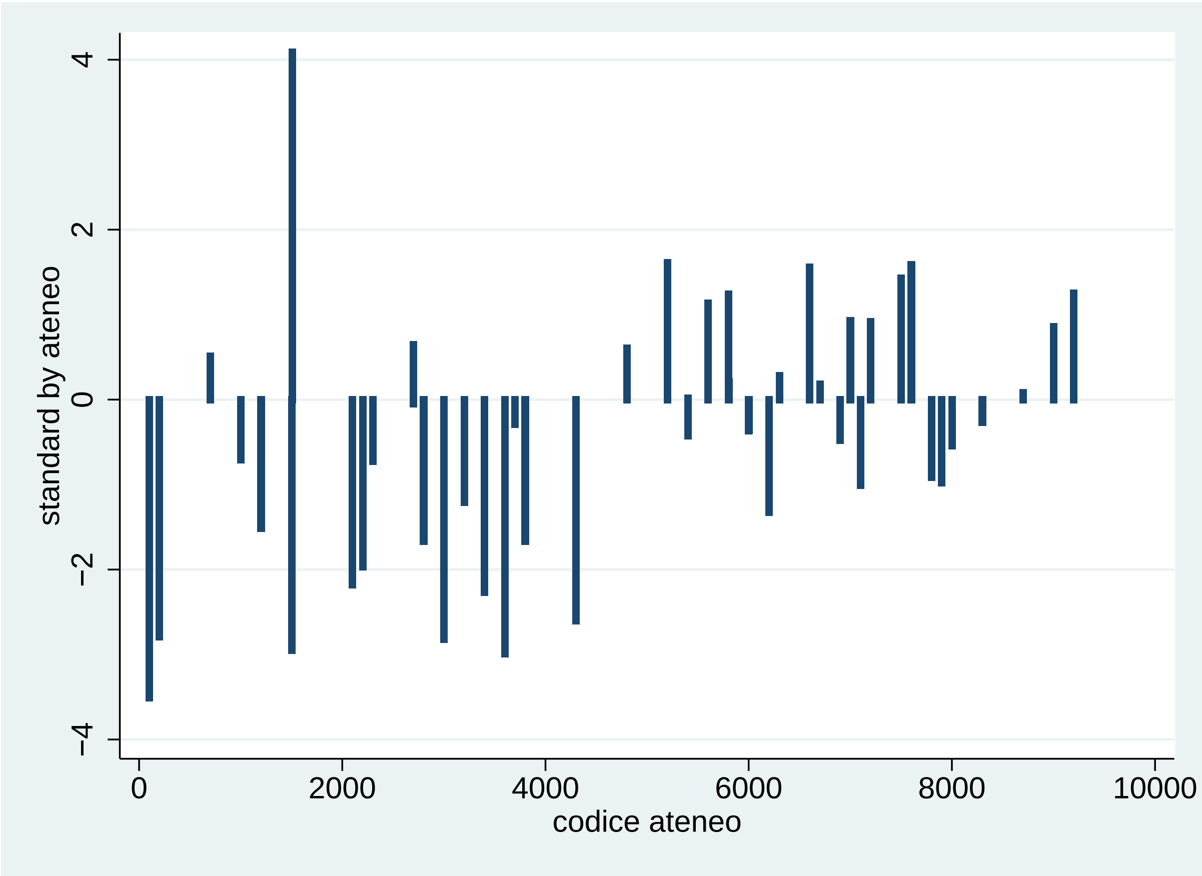


Figure 1.6: Distribution of standards by fields

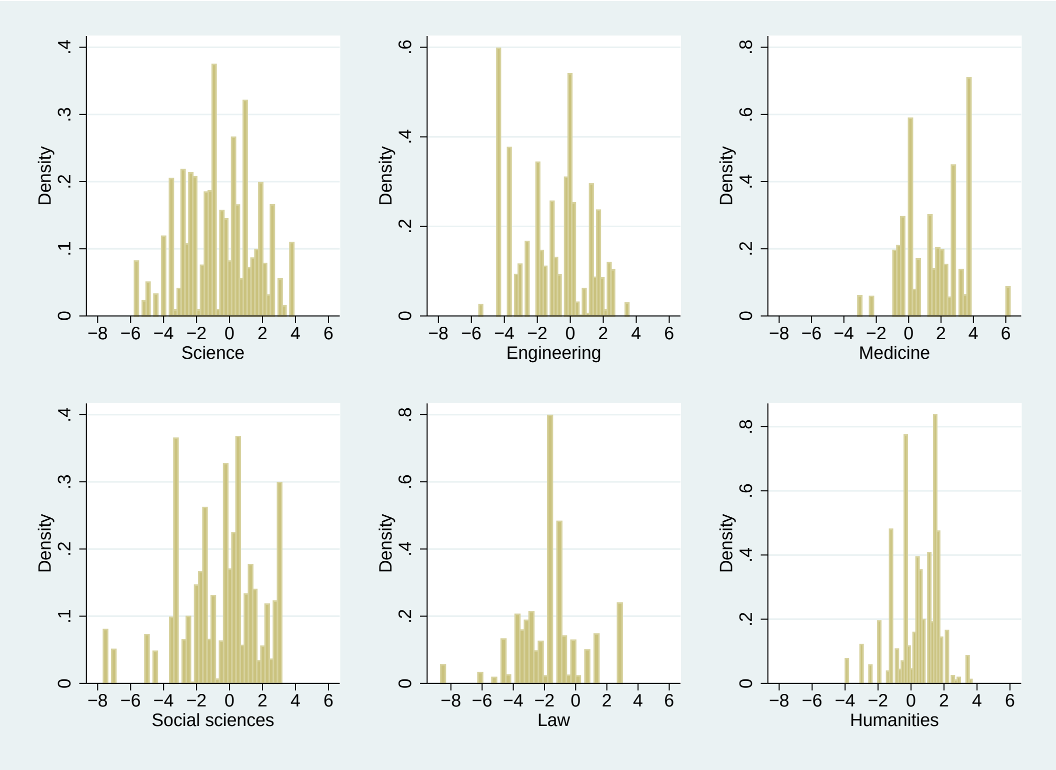


Table 1.4: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Variables Step 1					
courses attended	93821	0.742	0.437	0	1
italian citizen	93821	0.994	0.078	0	1
parents' ISCED 1-2	93821	0.218	0.413	0	1
parents' ISCED 3-4	93821	0.444	0.497	0	1
parents' ISCED 5-6	93821	0.337	0.473	0	1
study abroad	93821	0.157	0.364	0	1
work and study	93821	0.054	0.227	0	1
academic high school	93821	0.66	0.474	0	1
final grade	93821	108.112	5.941	72	113
Age at degree	93821	26.044	1.864	22.178	34.997
High school mark	93821	86.614	12	60	100
sex	93821	1.616	0.486	1	2
close to univ.	93821	0.693	0.461	0	1
Variables Step 2					
depart. effect on grade	354	0.193	2.513	-7.965	5.29
research quality indicator	354	0.565	0.174	0.12	0.98
funding from competitive procedures (000 euros)	354	5895	9182	0	72763
quality of hired/promoted staff	354	1.153	0.25	0.344	2.743
average fees	354	1057	874	394	6046
total financial resources (000 euros)	354	29536	28036	402	120061
student-teacher ratio	354	2.735	2.996	0.095	26.433
Share foreign students	354	0.023	0.044	0	0.582
Share same province	354	0.509	0.184	0.082	0.955
Variables Step 3					
standard	21948	-.802	2.192	-8.654	6.258
standard field	21948	-.493	1.728	-6.32	4.673
Xb	21948	-.194	1.131	-3.888	4.086
Xb field	21948	-.503	1.644	-5.078	5.039
year	21948	2008.52	.5	2008	2009
monthly wage 5 years	21948	1415.93	510.642	200	3250
employee 5 years	21948	.816	.388	0	1
overeducation 5 years	21948	.299	.458	0	1
public sector 5 years	21948	.154	.361	0	1
monthly wage 3 years	21948	1284.08	453.411	200	3250
employee 3 years	21948	.83	.376	0	1
public sector 3 years	21948	.151	.358	0	1
monthly wage 1 year	21948	1086.54	420.184	200	3250

Table 1.5: Dependent variables: final grade

	Full sample	Full sample	Full sample	Full sample	2nd cycle only
Ref: 2nd cycle					
Unique-cycle	-3.261***	-3.215***	-3.615***	-3.450***	
	(0.05)	(0.08)	(0.08)	(0.10)	
Pre-reform	-2.197***	-1.918***	-2.414***	-3.119***	
	(0.10)	(0.11)	(0.11)	(0.13)	
attendance	1.216***	1.029***	1.064***	1.115***	1.135***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
italian	0.829***	0.981***	0.900***	0.881***	0.777***
	(0.24)	(0.24)	(0.24)	(0.22)	(0.23)
study abroad	0.898***	0.582***	0.711***	0.738***	0.639***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
work and study	-0.466***	-0.592***	-0.704***	-0.785***	-0.949***
	(0.09)	(0.09)	(0.08)	(0.08)	(0.09)
Delay index	-3.482***	-3.183***	-3.469***	-3.252***	-2.500***
	(0.08)	(0.07)	(0.07)	(0.07)	(0.08)
HS mark	0.147***	0.152***	0.153***	0.154***	0.131***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
sex	0.751***	0.361***	0.307***	0.324***	0.339***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
close to uni	0.384***	0.350***	0.436***	0.412***	0.353***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Ref: academic HS					
other HS	-0.855***	-1.255***	-1.260***	-1.351***	-1.131***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
technical HS	-1.614***	-1.345***	-1.359***	-1.457***	-1.329***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
professional HS	-2.583***	-2.698***	-2.807***	-2.843***	-2.493***
	(0.18)	(0.18)	(0.17)	(0.17)	(0.18)
Ref: year 2008					
year 2009	-0.221***	-0.214***	-0.221***	-0.240***	-0.270***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Ref: parents ISCED 1-2					
parents ISCED 3-4	-0.193***	-0.248***	-0.198***	-0.187***	-0.189***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
parents ISCED 5-6	-0.213***	-0.295***	-0.235***	-0.253***	-0.245***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
constant	93.389***	93.770***	93.855***	94.588***	96.794***
	(0.56)	(0.56)	(0.55)	(0.55)	(0.59)
Area of origin controls	Yes	Yes	Yes	Yes	Yes
Department fixed effects	No	No	No	Yes	Yes
Field of study effects	No	Yes	Yes	No	No
University effects	Yes	No	Yes	No	No
R2	0.28	0.29	0.33	0.38	0.32
N	93820	93820	93820	93820	71681

Table 1.6: Dependent variable: Department FE on final grade

	No field controls	Field controls
funding from competitive procedures (000 euros)	-0.018	-0.037*
	(0.02)	(0.02)
total financial resources (000 euros)	-0.004	-0.002
	(0.00)	(0.00)
research quality indicator	-2.794***	-4.655***
	(0.90)	(1.76)
quality of hired/promoted staff	-2.032***	-1.252*
	(0.65)	(0.76)
student-teacher ratio	-0.482***	-0.033
	(0.10)	(0.11)
student-teacher ratio squared	0.015***	-0.002
	(0.01)	(0.00)
average fees	-0.381	-0.401**
	(0.24)	(0.20)
Share foreign students	-1.062	-6.866**
	(3.47)	(2.76)
Share same province	0.456	1.263**
	(0.73)	(0.62)
Ref: science		
chemistry		-1.057
		(0.65)
biology		0.801*
		(0.43)
medicine		1.849***
		(0.67)
engineering		-0.682
		(0.62)
architecture		0.120
		(0.50)
agricultural sciences		0.470
		(0.70)
economics		-2.843***
		(0.78)
political sciences		-0.545
		(0.59)
law		-2.893***
		(0.54)
humanities		1.106***
		(0.34)
teaching-psychology		0.121
		(0.47)
constant	5.842***	5.129***
	(0.98)	(0.96)
R2	0.22	0.47
N	354	354

Table 1.7: Dependent variables: log monthly wage 5 years

	Standard 1 Full sample	Standard 1 Full sample	Standard 1 2nd cycle	Standard 2 Full sample	Standard 2 Full sample	Standard 2 2nd cycle
Final grade	0.004***	0.007***	0.007***	0.004***	0.008***	0.008***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Ref: standard quartile 1						
standard quartile 2	-0.020***	-0.028***	-0.029***	-0.020***	-0.027***	-0.027***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 3	-0.012*	-0.021***	-0.020**	-0.017**	-0.027***	-0.030***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 4	-0.032***	-0.047***	-0.047***	-0.020***	-0.034***	-0.032***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Variables step 2	-0.008***	-0.010***	-0.010***	-0.011***	-0.013***	-0.015***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
sex	-0.154***	-0.158***	-0.156***	-0.154***	-0.159***	-0.157***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: parents ISCED 1-2						
parents ISCED 3-4	0.014***	0.014***	0.013**	0.014**	0.013***	0.013**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)
parents ISCED 5-6	0.035***	0.033***	0.028***	0.035***	0.033***	0.028***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
work and study	0.092***	0.100***	0.128***	0.093***	0.101***	0.130***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2nd cycle						
Unique cycle	0.079***	0.090***		0.081***	0.095***	
	(0.01)	(0.01)		(0.01)	(0.02)	
Pre-reform	0.183***	0.193***		0.186***	0.198***	
	(0.01)	(0.01)		(0.01)	(0.01)	
public sector	0.017**	0.018**	0.010	0.017***	0.018***	0.011
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
employee 5y	0.137***	0.137***	0.144***	0.138***	0.137***	0.143***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2008 year						
2009 year	-0.009*	-0.008*	-0.011**	-0.009**	-0.008*	-0.011**
	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
dropout rate		0.094*	0.075		0.057	0.036
		(0.05)	(0.06)		(0.05)	(0.05)
unobserved heterog.		-0.004***	-0.004***		-0.004***	-0.005***
		(0.00)	(0.00)		(0.00)	(0.00)
constant	6.717***	6.376***	6.352***	6.704***	6.316***	6.269***
	(0.04)	(0.09)	(0.10)	(0.05)	(0.10)	(0.11)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Field of study fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.32	0.32	0.34	0.32	0.32	0.34
N	21948	21889	17903	21948	21889	17903

Region FE are for region of employment. Bootstrap standard errors (200 reps) in parenthesis.

Table 1.8: Dependent variables: log monthly wage 3 years

	Standard 1 Full sample	Standard 1 Full sample	Standard 1 2nd cycle	Standard 2 Full sample	Standard 2 Full sample	Standard 2 2nd cycle
Final grade	0.004***	0.006***	0.006***	0.004***	0.006***	0.007***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Ref: standard quartile 1						
standard quartile 2	-0.017**	-0.024***	-0.029***	-0.007	-0.013*	-0.016**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 3	0.001	-0.005	-0.006	-0.008	-0.015*	-0.016*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 4	-0.023**	-0.033***	-0.034***	-0.012*	-0.022**	-0.019**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
sex	-0.134***	-0.137***	-0.136***	-0.134***	-0.138***	-0.136***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
Ref: 2008 year						
2009 year	-0.028***	-0.027***	-0.030***	-0.028***	-0.027***	-0.030***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Ref: parents ISCED 1-2						
parents ISCED 3-4	0.016***	0.016***	0.019***	0.016***	0.015***	0.018***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
parents ISCED 5-6	0.032***	0.031***	0.032***	0.032***	0.031***	0.032***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
work and study	0.119***	0.124***	0.150***	0.119***	0.125***	0.151***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2nd cycle						
Unique-cycle	0.042***	0.050***		0.045***	0.055***	
	(0.01)	(0.02)		(0.01)	(0.02)	
Pre-reform	0.203***	0.212***		0.207***	0.215***	
	(0.01)	(0.02)		(0.01)	(0.01)	
public sector	0.049***	0.049***	0.042***	0.049***	0.049***	0.042***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
employee	0.182***	0.181***	0.195***	0.182***	0.182***	0.195***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Variables step 2	-0.003	-0.004	-0.002	-0.005**	-0.007***	-0.007**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
dropout rate		0.109*	0.145**		0.057	0.089
		(0.06)	(0.06)		(0.05)	(0.06)
unobserved heterog.		-0.003**	-0.003***		-0.003***	-0.003***
		(0.00)	(0.00)		(0.00)	(0.00)
constant	6.603***	6.347***	6.258***	6.596***	6.313***	6.226***
	(0.05)	(0.10)	(0.11)	(0.05)	(0.10)	(0.10)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Field of study fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.29	0.30	0.32	0.29	0.30	0.32
N	21948	21889	17903	21948	21889	17903

Region FE are for region of employment. Bootstrap standard errors in parenthesis.

Table 1.9: Dependent variables: log monthly wage 1 year

	Standard 1 Full sample	Standard 1 Full sample	Standard 1 2nd cycle	Standard 2 Full sample	Standard 2 Full sample	Standard 2 2nd cycle
Final grade	0.002***	0.002**	0.002	0.002***	0.002**	0.002*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Ref: standard quartile 1						
standard quartile 2	-0.019**	-0.022**	-0.025***	0.005	0.004	-0.008
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 3	-0.002	-0.003	-0.004	-0.006	-0.007	-0.017*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 4	-0.015	-0.017	-0.017	0.004	0.001	0.004
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Variables step 2	-0.002	-0.002	0.000	-0.008***	-0.009***	-0.008***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
sex	-0.124***	-0.125***	-0.125***	-0.124***	-0.125***	-0.126***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2008 year						
2009 year	-0.018***	-0.018***	-0.021***	-0.018***	-0.018***	-0.020***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: parents ISCED 1-2						
parents ISCED 3-4	0.013**	0.013**	0.013*	0.013**	0.013**	0.013*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
parents ISCED 5-6	0.014*	0.014	0.016*	0.013*	0.014*	0.016*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
work and study	0.186***	0.188***	0.223***	0.187***	0.188***	0.224***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2nd cycle						
Unique-cycle	0.018	0.020		0.020	0.022	
	(0.02)	(0.02)		(0.02)	(0.02)	
Pre-reform	0.321***	0.326***		0.323***	0.328***	
	(0.02)	(0.02)		(0.02)	(0.02)	
public sector 1y	0.097***	0.096***	0.077***	0.097***	0.096***	0.078***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
employee 1y	0.394***	0.394***	0.408***	0.395***	0.395***	0.408***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
dropout rate		0.039	0.038		-0.027	-0.020
		(0.07)	(0.07)		(0.06)	(0.07)
unobserved heterog.		-0.001	0.000		-0.001	-0.000
		(0.00)	(0.00)		(0.00)	(0.00)
constant	6.463***	6.404***	6.429***	6.452***	6.397***	6.410***
	(0.06)	(0.12)	(0.13)	(0.06)	(0.12)	(0.13)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Field of study fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.32	0.32	0.33	0.32	0.32	0.33
N	21948	21889	17903	21948	21889	17903

Region FE are for region of employment. Bootstrap standard errors in parenthesis.

Table 1.10: Dependent variables: overeducation 5 years, LPM

	Standard 1 Full sample	Standard 1 Full sample	Standard 1 2nd cycle	Standard 2 Full sample	Standard 2 Full sample	Standard 2 2nd cycle
Final grade	-0.005***	-0.011***	-0.014***	-0.005***	-0.012***	-0.015***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Ref: standard quartile 1						
standard quartile 2	0.013	0.025***	0.035***	0.017**	0.031***	0.036***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 3	0.029***	0.052***	0.067***	0.044***	0.065***	0.076***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 4	0.045***	0.076***	0.091***	0.030***	0.060***	0.071***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Variables step 2	0.009**	0.015***	0.017***	0.014***	0.021***	0.025***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
sex	0.032***	0.040***	0.050***	0.031***	0.041***	0.051***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: parents ISCED 1-2						
parents ISCED 3-4	-0.007	-0.006	-0.009	-0.007	-0.006	-0.008
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
parents ISCED 5-6	-0.034***	-0.032***	-0.035***	-0.034***	-0.031***	-0.035***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
work and study	0.046***	0.033***	0.019	0.045***	0.030**	0.015
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2nd cycle						
Unique cycle	-0.265***	-0.286***		-0.266***	-0.290***	
	(0.01)	(0.01)		(0.01)	(0.01)	
Pre-reform	-0.471***	-0.491***		-0.472***	-0.495***	
	(0.02)	(0.02)		(0.02)	(0.02)	
public sector	-0.063***	-0.064***	-0.076***	-0.064***	-0.065***	-0.077***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
employee 5y	0.116***	0.117***	0.119***	0.116***	0.117***	0.119***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2008 year						
2009 year	0.024***	0.021***	0.024***	0.023***	0.021***	0.022***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
dropout rate		0.082	0.132*		0.081	0.135*
		(0.06)	(0.08)		(0.06)	(0.08)
unobserved heterog.		0.008***	0.010***		0.009***	0.011***
		(0.00)	(0.00)		(0.00)	(0.00)
constant	0.799***	1.432***	1.736***	0.823***	1.560***	1.909***
	(0.06)	(0.13)	(0.15)	(0.06)	(0.12)	(0.15)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Field of study fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.14	0.15	0.08	0.14	0.15	0.08
N	21948	21889	17903	21948	21889	17903

Region FE are for region of employment. Bootstrap standard errors in parenthesis.

Table 1.11: Dependent variables: overeducation 1 year, LPM

	Standard 1 Full sample	Standard 1 Full sample	Standard 1 2nd cycle	Standard 2 Full sample	Standard 2 Full sample	Standard 2 2nd cycle
Final grade	-0.004***	-0.011***	-0.013***	-0.004***	-0.012***	-0.014***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Ref: standard quartile 1						
standard quartile 2	0.030***	0.047***	0.054***	0.021***	0.037***	0.045***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 3	0.038***	0.059***	0.070***	0.037***	0.059***	0.066***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 4	0.064***	0.095***	0.103***	0.044***	0.073***	0.081***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Variables step 2	0.014***	0.017***	0.018***	0.016***	0.021***	0.023***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
sex	0.041***	0.050***	0.056***	0.040***	0.050***	0.057***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: parents ISCED 1-2						
parents ISCED 3-4	-0.031***	-0.029***	-0.033***	-0.030***	-0.029***	-0.033***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
parents ISCED 5-6	-0.045***	-0.042***	-0.045***	-0.045***	-0.041***	-0.045***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
work and study	0.068***	0.053***	0.039***	0.068***	0.051***	0.036**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2nd cycle						
Unique cycle	-0.284***	-0.308***		-0.286***	-0.313***	
	(0.01)	(0.01)		(0.01)	(0.01)	
Pre-reform	-0.557***	-0.580***		-0.558***	-0.584***	
	(0.02)	(0.02)		(0.02)	(0.02)	
public sector	-0.038***	-0.040***	-0.045***	-0.038***	-0.040***	-0.044***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
employee	0.018*	0.019**	0.016	0.018**	0.019**	0.017*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2008 year						
2009 year	0.020***	0.018***	0.020***	0.020***	0.018***	0.019**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
dropout rate		-0.139**	-0.147*		-0.110*	-0.122
		(0.07)	(0.09)		(0.06)	(0.08)
unobserved heterog.		0.008***	0.010***		0.009***	0.011***
		(0.00)	(0.00)		(0.00)	(0.00)
constant	0.856***	1.584***	1.840***	0.879***	1.672***	1.974***
	(0.06)	(0.13)	(0.17)	(0.06)	(0.12)	(0.17)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Field of study fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.15	0.16	0.09	0.15	0.16	0.09
N	21948	21889	17903	21948	21889	17903

Region FE are for region of employment. Bootstrap standard errors in parenthesis.

Table 1.12: Dependent variables: full-time employment 5 years. LPM

	Standard 1 Full sample	Standard 1 2nd cycle only	Standard 2 Full sample	Standard 2 2nd cycle only
Final grade	0.011***	0.011***	0.011***	0.012***
	(0.00)	(0.00)	(0.00)	(0.00)
sex	-0.106***	-0.100***	-0.106***	-0.100***
	(0.00)	(0.00)	(0.00)	(0.00)
Ref: standard quartile 1				
standard quartile 2	-0.031***	-0.030***	-0.026***	-0.024***
	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 3	-0.024***	-0.014*	-0.028***	-0.027***
	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 4	-0.044***	-0.037***	-0.034***	-0.029***
	(0.01)	(0.01)	(0.01)	(0.01)
Variables step 2	-0.003	-0.007**	-0.006**	-0.009***
	(0.00)	(0.00)	(0.00)	(0.00)
Ref: parents ISCED 1-2				
parents ISCED 3-4	0.018***	0.018***	0.018***	0.018***
	(0.00)	(0.00)	(0.00)	(0.01)
parents ISCED 5-6	0.016***	0.014***	0.016***	0.014**
	(0.01)	(0.01)	(0.01)	(0.01)
work and study	0.116***	0.122***	0.117***	0.123***
	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2nd cycle				
Unique cycle	0.023**		0.028***	
	(0.01)		(0.01)	
Pre-reform	0.430***		0.433***	
	(0.01)		(0.01)	
Ref: 2008 year				
2009 year	-0.011***	-0.012***	-0.011***	-0.011***
	(0.00)	(0.00)	(0.00)	(0.00)
dropout rate	0.015	-0.015	-0.001	-0.037
	(0.04)	(0.05)	(0.05)	(0.06)
unobserved heterog.	-0.007***	-0.007***	-0.007***	-0.008***
	(0.00)	(0.00)	(0.00)	(0.00)
constant	-0.253***	-0.316***	-0.289***	-0.377***
	(0.09)	(0.10)	(0.09)	(0.11)
Region fixed effects	Yes	Yes	Yes	Yes
Field of study fixed effects	Yes	Yes	Yes	Yes
R2	0.16	0.18	0.16	0.18
N	44259	36493	44259	36493

Controls not included in the table: parents' educ. Region FE are for region of employment. Bootstrap standard errors in parenthesis.

Table 1.13: Dependent variables: full-time employment 3 years. LPM

	Standard 1 Full sample	Standard 1 2nd cycle only	Standard 2 Full sample	Standard 2 2nd cycle only
Final grade	0.009***	0.011***	0.010***	0.011***
	(0.00)	(0.00)	(0.00)	(0.00)
sex	-0.089***	-0.086***	-0.090***	-0.087***
	(0.00)	(0.00)	(0.00)	(0.00)
Ref: standard quartile 1				
standard quartile 2	-0.020***	-0.015**	-0.015**	-0.023***
	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 3	-0.016**	-0.002	-0.026***	-0.028***
	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 4	-0.031***	-0.020*	-0.025***	-0.019**
	(0.01)	(0.01)	(0.01)	(0.01)
Variables step 2	-0.001	-0.004	-0.006*	-0.007**
	(0.00)	(0.00)	(0.00)	(0.00)
Ref: parents ISCED 1-2				
parents ISCED 3-4	0.013**	0.016***	0.013**	0.016***
	(0.01)	(0.01)	(0.00)	(0.01)
parents ISCED 5-6	0.006	0.007	0.006	0.007
	(0.01)	(0.01)	(0.01)	(0.01)
work and study	0.172***	0.181***	0.173***	0.183***
	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2nd cycle				
Unique cycle	-0.023**		-0.020**	
	(0.01)		(0.01)	
Pre-reform	0.450***		0.453***	
	(0.01)		(0.01)	
Ref: 2008 year				
2009 year	-0.015***	-0.015***	-0.014***	-0.015***
	(0.00)	(0.00)	(0.00)	(0.00)
dropout rate	-0.073	-0.077	-0.085*	-0.082
	(0.05)	(0.05)	(0.05)	(0.05)
unobserved heterog.	-0.007***	-0.008***	-0.007***	-0.009***
	(0.00)	(0.00)	(0.00)	(0.00)
constant	-0.127	-0.271***	-0.179*	-0.355***
	(0.09)	(0.10)	(0.10)	(0.10)
Region fixed effects	Yes	Yes	Yes	Yes
Field of study fixed effects	Yes	Yes	Yes	Yes
R2	0.19	0.19	0.19	0.19
N	44259	36493	44259	36493

Region FE are for region of employment. Bootstrap standard errors in parenthesis.

Table 1.14: Dependent variables: employment 1 year. LPM

	Standard 1 Full sample	Standard 1 2nd cycle only	Standard 2 Full sample	Standard 2 2nd cycle only
Final grade	0.003***	0.003***	0.003***	0.003***
	(0.00)	(0.00)	(0.00)	(0.00)
sex	-0.060***	-0.060***	-0.061***	-0.061***
	(0.01)	(0.01)	(0.01)	(0.01)
Ref: standard quartile 1				
standard quartile 2	-0.009	-0.004	-0.015**	-0.015*
	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 3	-0.009	0.001	-0.028***	-0.019**
	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 4	-0.005	-0.002	-0.016**	-0.007
	(0.01)	(0.01)	(0.01)	(0.01)
Variables step 2	0.001	-0.001	0.001	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)
Ref: parents ISCED 1-2				
parents ISCED 3-4	-0.003	-0.001	-0.003	-0.001
	(0.01)	(0.01)	(0.00)	(0.01)
parents ISCED 5-6	-0.023***	-0.027***	-0.023***	-0.027***
	(0.01)	(0.01)	(0.01)	(0.01)
work and study	0.288***	0.302***	0.289***	0.303***
	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2nd cycle				
Unique cycle	-0.019**		-0.015*	
	(0.01)		(0.01)	
Pre-reform	0.502***		0.507***	
	(0.01)		(0.01)	
Ref: 2008 year				
2009 year	-0.039***	-0.040***	-0.039***	-0.040***
	(0.00)	(0.00)	(0.00)	(0.00)
dropout rate	-0.078	-0.096*	-0.062	-0.081
	(0.05)	(0.05)	(0.05)	(0.06)
unobserved heterog.	-0.002*	-0.002	-0.003***	-0.002*
	(0.00)	(0.00)	(0.00)	(0.00)
constant	0.433***	0.428***	0.363***	0.379***
	(0.09)	(0.11)	(0.10)	(0.11)
Region fixed effects	Yes	Yes	Yes	Yes
Field of study fixed effects	Yes	Yes	Yes	Yes
R2	0.20	0.19	0.20	0.19
N	44259	36493	44259	36493

Region FE are for region of employment. Bootstrap standard errors in parenthesis.

Table 1.15: Dependent variables: log monthly wage 5 years

	Standard 1 Employee	Standard 1 Self-employed	Standard 1 Public sector	Standard 1 Private sector	Standard 1 Males	Standard 1 Females	Standard 1 High ability	Standard 1 Low ability
VotoLau	0.006***	0.011***	0.006***	0.007***	0.005***	0.009***	0.008***	0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Ref: standard quartile 1								
standard quartile 2	-0.031***	-0.037	-0.008	-0.028***	-0.007	-0.049***	-0.032***	-0.018*
	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 3	-0.019***	-0.060**	-0.004	-0.019**	-0.015	-0.035***	-0.026***	-0.007
	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 4	-0.032***	-0.122***	-0.017	-0.052***	-0.034**	-0.066***	-0.052***	-0.027*
	(0.01)	(0.03)	(0.03)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)
Xb	-0.012***	0.001	-0.010	-0.010***	-0.012**	-0.010**	-0.010***	-0.007
	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
dropout rate	0.058	0.220	0.133	0.105	-0.092	0.173***	0.093	0.082
	(0.05)	(0.19)	(0.10)	(0.06)	(0.10)	(0.06)	(0.06)	(0.08)
unobs heterog.	-0.003***	-0.008**	-0.004*	-0.004***	-0.001	-0.006***	-0.003**	-0.005**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
sex	-0.132***	-0.245***	-0.120***	-0.161***			-0.160***	-0.156***
	(0.00)	(0.02)	(0.01)	(0.01)			(0.01)	(0.01)
Ref: parents ISCED 1-2								
parents ISCED 3-4	0.013***	0.020	0.006	0.016**	0.016**	0.012*	0.013**	0.015*
	(0.00)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
parents ISCED 5-6	0.023***	0.059**	-0.008	0.038***	0.036***	0.031***	0.036***	0.027**
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
work study	0.093***	0.181***	0.091***	0.111***	0.126***	0.089***	0.103***	0.096***
	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2nd cycle								
Unique cycle	0.083***	0.075**	0.215***	0.064***	0.048**	0.125***	0.098***	0.078***
	(0.02)	(0.03)	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Pre-reform	0.174***	0.319*	0.174***	0.125***	0.087**	0.190***	0.178***	0.205***
	(0.01)	(0.16)	(0.03)	(0.02)	(0.04)	(0.02)	(0.02)	(0.02)
public sector	0.023***	-0.002			-0.028**	0.033***	0.008	0.032***
	(0.01)	(0.04)			(0.01)	(0.01)	(0.01)	(0.01)
employee			0.155***	0.138***	0.092***	0.184***	0.128***	0.147***
	(0.00)	(0.00)	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2008 year								
2009 year	-0.003	-0.028*	0.020**	-0.013***	-0.012*	-0.006	-0.015***	-0.001
	(0.00)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
constant	6.616***	5.865***	6.329***	6.362***	6.663***	5.959***	6.304***	6.586***
	(0.07)	(0.38)	(0.23)	(0.10)	(0.14)	(0.15)	(0.18)	(0.17)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field of study fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.33	0.33	0.33	0.33	0.33	0.25	0.34	0.30
N	10388	10388	10388	10388	9666	12223	12522	9367

Table 1.16: Dependent variables: log monthly wage 5 years

	Standard 2 Employee	Standard 2 Self-employed	Standard 2 Public sector	Standard 2 Private sector	Standard 2 Males	Standard 2 Females	Standard 2 High ability	Standard 2 Low ability
VotoLau	0.007***	0.011***	0.006***	0.008***	0.006***	0.009***	0.009***	0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
standard quartile 1								
standard quartile 2	-0.022***	-0.060**	-0.009	-0.024***	-0.022**	-0.033***	-0.023***	-0.029***
	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 3	-0.022***	-0.051*	-0.020	-0.024**	-0.022*	-0.036***	-0.023**	-0.029**
	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
standard quartile 4	-0.026***	-0.080***	-0.019	-0.035***	-0.027**	-0.046***	-0.049***	-0.004
	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Xb field	-0.016***	-0.003	-0.002	-0.015***	-0.013***	-0.015***	-0.013***	-0.009**
	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
dropout rate	0.008	0.186	0.128	0.057	-0.085	0.112*	0.040	0.077
	(0.05)	(0.17)	(0.08)	(0.06)	(0.11)	(0.06)	(0.06)	(0.08)
unobs heterog	-0.004***	-0.007**	-0.004**	-0.005***	-0.002	-0.007***	-0.004***	-0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
sex	-0.132***	-0.244***	-0.120***	-0.161***			-0.160***	-0.157***
	(0.01)	(0.02)	(0.01)	(0.01)			(0.01)	(0.01)
Ref: parents ISCED 1-2								
parents ISCED 3-4	0.012***	0.021	0.006	0.015**	0.016**	0.011	0.012*	0.015
	(0.00)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
parents ISCED 5-6	0.022***	0.059***	-0.007	0.038***	0.036***	0.030***	0.036***	0.027**
	(0.00)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
work study	0.095***	0.179***	0.091***	0.112***	0.127***	0.090***	0.105***	0.095***
	(0.01)	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2nd cycle								
Unique-cycle	0.090***	0.081***	0.216***	0.068***	0.049**	0.134***	0.104***	0.081***
	(0.02)	(0.03)	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Pre-reform	0.180***	0.317**	0.174***	0.132***	0.090*	0.195***	0.183***	0.210***
	(0.01)	(0.16)	(0.02)	(0.02)	(0.05)	(0.01)	(0.02)	(0.03)
public sector	0.024***	-0.002			-0.028**	0.033***	0.008	0.031**
	(0.01)	(0.04)			(0.01)	(0.01)	(0.01)	(0.01)
employee 5y			0.153***	0.138***	0.091***	0.185***	0.129***	0.147***
			(0.05)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
2008 year	0.000	0.000	0.000	0.000				
2009 year	-0.003	-0.028*	0.019**	-0.013***	-0.012**	-0.006	-0.015***	-0.001
	(0.00)	(0.02)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
constant	6.539***	5.870***	6.327***	6.296***	6.598***	5.892***	6.158***	6.593***
	(0.09)	(0.36)	(0.21)	(0.11)	(0.15)	(0.15)	(0.15)	(0.15)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field of study fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.33	0.33	0.33	0.33	0.33	0.25	0.34	0.30
N	10388	10388	10388	10388	9666	12223	12522	9367

Table 1.17: Dependent variables: employment 5 years

	Standard 1 Males	Standard 1 Females	Standard 1 High ability	Standard 1 Low ability	Standard 2 Males	Standard 2 Females	Standard 2 High ability	Standard 2 Low ability
VotoLau	0.006***	0.014***	0.010***	0.009***	0.006***	0.014***	0.010***	0.010***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
sex			-0.096***	-0.122***			-0.096***	-0.122***
			(0.01)	(0.01)			(0.01)	(0.01)
work study	0.090***	0.129***	0.129***	0.105***	0.091***	0.130***	0.129***	0.108***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: parents ISCED 1-2								
parents ISCED 3-4	0.009	0.021***	0.023***	0.012	0.009	0.021***	0.023***	0.012
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
parents ISCED 5-6	0.003	0.022***	0.016**	0.016**	0.003	0.022***	0.016**	0.016**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ref: 2008 year								
2009 year	-0.011**	-0.011**	-0.009*	-0.015**	-0.011**	-0.011**	-0.009*	-0.015**
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
Ref: 2nd cycle								
Unique-cycle	0.026*	0.024*	0.025**	0.026**	0.026*	0.024*	0.025**	0.026**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Pre-reform	0.318***	0.445***	0.417***	0.437***	0.318***	0.445***	0.417***	0.437***
	(0.05)	(0.01)	(0.02)	(0.02)	(0.05)	(0.01)	(0.02)	(0.02)
Ref: stand quart 1								
stand quart 2	-0.015*	-0.055***	-0.021***	-0.038***	-0.015*	-0.055***	-0.021***	-0.038***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
stand quart 3	-0.003	-0.057***	-0.022**	-0.022**	-0.003	-0.057***	-0.022**	-0.022**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
stand quart 4	-0.029**	-0.075***	-0.029**	-0.059***	-0.029**	-0.075***	-0.029**	-0.059***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
Xb	0.003	-0.010***	-0.008**	0.004	0.003	-0.010***	-0.008**	0.004
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
dropout rate	-0.040	0.013	0.052	-0.027	-0.040	0.013	0.052	-0.027
	(0.07)	(0.05)	(0.06)	(0.08)	(0.07)	(0.05)	(0.06)	(0.08)
unobs heterog	-0.002**	-0.009***	-0.009***	-0.004***	-0.002**	-0.009***	-0.009***	-0.004***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
constant	0.305***	-0.735***	-0.160	0.009	0.305***	-0.735***	-0.160	0.009
	(0.09)	(0.12)	(0.14)	(0.13)	(0.09)	(0.12)	(0.14)	(0.13)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field of study fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.33	0.25	0.34	0.30	0.33	0.25	0.34	0.30
N	9666	12223	12522	9367	9666	12223	12522	9367

Chapter 2

Assessing the impact of the first Italian Research Evaluation Exercise on students' choices

ABSTRACT - This paper assesses the impact of the first Italian Research Evaluation Exercise (VTR 2001-2003) on university students' enrolment choices. A before-after estimator with different 'treatment intensities' is used to investigate whether departments that received a better score also enjoyed more enrolments and enrolments of students with better entry qualifications after the VTR. Our analysis demonstrates that increasing the percentage of excellent products by one standard deviation at the department level increases student enrolments by 6.5 percent. Effects are larger for high quality students, namely those with better high school final marks (10 percent) and coming from the academic track (11.8 percent). Departments in the top quartile of the quality distribution gained more from a good performance in the evaluation exercise. Effect sizes appear to be similar across all macro-regions (North, Centre and South and Islands), but are precisely estimated only for universities in Northern Italy.

JEL Numbers: I21 I23
Keywords: research evaluation exercise, student enrolment, student quality, Italy

2.1 Introduction

Research evaluation is relatively recent in Italy. The first Research Evaluation Exercise (REE, hereafter) concerning the period 2001-2003 (VTR 2001-2003)[1] was completed in 2006 and the results made public in the same year. The results of a second REE covering the scientific production of the period 2004-2010 (VQR 2004-2010)[2] were publicly released in 2013, while the third REE (VQR 2011-2014) is currently ongoing.

[1] *Valutazione Triennale della Ricerca* (three-year research evaluation).
[2] *Valutazione della Qualita' della Ricerca* (seven-year research evaluation).

All Italian REEs have been followed by lively debates. Critics of REE maintain that they are very expensive and excessively based on quantitative (e.g. bibliometric) indicators. Advocates of REEs rebut that in a period of shrinking public funding of Higher Education it is more important than ever to allocate resources in an effective and efficient way.

Representing the first adoption of a Performance-based Research Funding System (PRFS) in Italy, the VTR, bust also the following REEs, attracted a considerable attention by researchers (Rebora and Turri, 2013; Geuna and Piolatto, 2016). However, following a well established stream of literature (see, among others, Jiménez-Contreras et al., 2003; Auranen and Nieminen, 2010) only the effect of the VTR on the supply side of Higher Education, namely on universities' research productivity, has been assessed (Cattaneo et al., 2016). Surprisingly enough, there are no studies on the effect of VTR on the demand side, i.e. on students. In the current paper, we aim at filling this gap by investigating whether the score obtained in the VTR had any consequence for Italian Higher Education Institutions (HEIs, hereafter) in terms of the number and the quality of enrolled students.

Our paper is related to the literature which, especially in the US, investigates the effects on student application and matriculation decisions of ratings and rankings of HEIs produced by private 'intermediaries' (e.g., the US News and World Report College Rankings). In general, they do find a positive effect of improving institutional ranking on student applications (see the literature review in Tutterow and Evans, 2016), whose size however is not very large, and is generally lower in time-series studies controlling for prior rank (Sauder and Lancaster, 2006). Moreover, the effect of ranking on the number of applications and matriculations is larger for top institutions (Bowman and Bastedo, 2009). The way information is presented also matters. A better rank is more effective at raising applications when HEIs are listed in rank rather than in alphabetical order, although this effect is smaller for top institutions which already have a well established reputation (Luca and Smith, 2013). A higher rank is also associated with more selectivity in admissions and lower acceptance rates (Monks and Ehrenberg, 1999; Meredith, 2004), and a higher student quality (Monks and Ehrenberg, 1999; Griffith and Rask, 2007).[3] Evidence exists also for the UK, where researchers have assessed the responsiveness of applications to the rankings produced by popular newspapers, like The Guardian or The Times. Results align with the US literature. A better ranking is associated with more applications, and the effect is stronger for the institutions in the top quantiles of the quality distribution and for overseas students, who pay higher fees and are more sensitive to quality (Chevalier and Jia, 2015). Papers which pool all subjects and analyse the effect of ranking on applications at the university level rather than at the departmental level generally find smaller effects (Soo, 2013; Broecke, 2015). This is partly due to the high heterogeneity existing in the quality of departments within an institution (Chevalier and Jia, 2015; Gibbons et al., 2015). Interestingly, also UK studies confirm that the salience of information matters. Information on student satisfaction only affects applications when it is incorporated in league tables, and ranking scores are more relevant when there is high competition among departments and institutions (Gibbons et al., 2015).

Despite the existence of abundant evidence on the effects of league tables, none of the studies just mentioned has looked into the effect of 'official' rankings, e.g., those produced by national REE, on student choices. On this issue the evidence is to the best of our knowledge almost non-existent. We are only aware of one study by Horstschräer (2012), which focuses on a single field of study, however, and demonstrates that being awarded excellence status in a government-run excellence competition significantly increased Medical schools' student applications in Germany.

In this paper, we seek to contribute to this still scant literature by focusing on the effects of the first Italian REE on university students' enrollment decisions. Italy is an interesting case study. Italy has been always characterized by the so-called *legal value* of university degrees. This granted a formal equality among all degrees irrespective of the awarding institutions, e.g., in the access to public sector jobs. However, the progressive reduction in the universities' public funding,[4] together with a decrease in student number[5]

[3] An effect on the SAT score is not found by Meredith (2004) instead.

[4] The 'Fondo di Finanziamento Ordinario' (FFO), that is the main source of public funding for Italian HEIs, decreased from almost 7.5 billion euros in 2009 to less than 6.4 billion euros in 2015.

[5] The total number of students enrolled decreased from a peak of 338 thousands in the academic

has spurred increasing competition among HEIs creating a quasi-market. In the absence of an official quality assessment of HEIs, students had little guidance when choosing which institution to enrol in. Popular newspapers such as 'La Repubblica' or 'Il Sole 24 Ore' have exploited this lack of information starting to produce specialized publications with HEIs' league tables. On the one hand, in this context, the setting of an official REE is likely to have provided a reliable source of information to students and to have had an impact on their choices. On the other hand, since the object of evaluation was only research, it is not at all obvious that such information was deemed relevant by students when choosing HEIs. The main goal of this paper is to assess whether this was the case or not. The focus on Italy is also important in the light of the heated debate on the fact that REEs may make it worse the brain drain in Southern regions (Fondazione RES, 2016). Indeed, Northern regions have been historically characterized by a net inflow of university students, also thanks to their labor markets which offer students better employment prospects. However, the geographical gap between Northern and Southern HEIs seems to have widened after 2006, that is, incidentally, the year of the VTR (see Figure 2.5). It is then important to assess whether a bad performance in the VTR might have been an important factor in accelerating the hemorrhage of students that Southern HEIs are suffering.

We provide a first assessment of the impact of the VTR on student choices using a before-after estimator which exploits differential treatment intensities across HEIs.[6] The quality score obtained in the VTR is the 'dose' of the treatment administered to HEIs. In our analysis we compare HEIs' outcomes (total enrolments and student quality) before and after the VTR, and look at whether in the post-VTR period there were changes significantly (positively or negatively) associated with the score obtained in the VTR. The main identification assumption is that there are no omitted variables which may be responsible for these changes. Such unobservable factors must have two features to threat our identification strategy: 1) they must have the same timing as the diffusion of

year 2003/2004, after the 3+2 reform of 2001, to 255 thousands in 2014/2015, the last academic year for which the Ministry of Education, University and Research (MIUR) provides data.

[6]See **?** in a difference-in-differences context.

the VTR results; 2) they must be correlated with the VTR scores. This makes it clear the importance of exploiting differences in VTR scores for identification. When making a simple before-after comparison, i.e. by simply comparing outcomes between the pre- and post-2006 period, the effect of the VTR may be confounded, for instance, with that of the Great Recession starting in 2008. By contrast, by exploiting for identification also differences in treatment intensities between HEIs and scientific areas, we will be able to control for year-specific or even province-year-specific fixed effects absorbing *inter-alia* the effect of the Great Recession, even if it was different across the regions where HEIs were located.

We contribute to the extant literature in at least three ways. First, as we mentioned, our study is the first one to systematically examine the effect of an 'official' REE on students' choices. Second, in line with the most recent literature (Chevalier and Jia, 2015; Gibbons et al., 2015), we frame the analysis at the department level. This is important because like for newspapers' league tables, also in REE HEIs' ranking scores are very likely to differ across disciplines. We provide evidence that this was indeed the case in the VTR. Third, unlike the previous literature on privately-produced league tables, we compare the period in which an official REE was not in place to the period where a REE was functioning. Thus, our paper is not concerned with the effects of increasing HEIs' ranking, but with how HEIs' enrolments changed over time time as a consequence of having performed well (or badly) in the first REE. In this sense, our estimates can be roughly interpreted as the effect on student choices of establishing a REE. This is of interest not only to stakeholders in Italy but also to readers in the many countries which are thinking of implementing similar Research Evaluation Exercises.

The main findings of our paper can be summarized as follows. First, we show that while the VTR score (an indicator of average HEIs' quality) does not affect student demand at the department level, the percentage of excellent products is positively associated with student enrolment. This is partly due to the higher ability of the second research quality indicator to discriminate across departments. Second, the VTR has a larger effect on enrolment of high-quality students, i.e. those with better entry qualifications. Those are indeed the students who are likely to care most about HEIs' quality. Third, the positive effect of VTR rating on enrolment is stronger in the top quartile of the quality distribution. This is consistent with REE and in general PRFS to increase competition among those HEI which have some chances to win in the 'race for quality.' Fourth, and last, the effect of VTR seems to be very similar across geographical areas, although it is precisely estimated only for Northern Italy.

The paper proceeds as follows. Section 3.2 describes the context in which the first Italian REE was introduced and its main characteristics. In Section 3.4 we explain our empirical strategy. Section 3.5 describes the data used in the empirical analysis, whose results are commented in Section 3.6. Finally, Section 3.7 summarizes the main findings and concludes.

2.2 The Italian system of Higher Education and the first Research Evaluation Exercise

The Italian higher education system has always been characterized by a high degree of centralization. Law n. 382 11/7/1980 provided that any variation in the existing university supply had to be included in a development plan, to be approved by the Minister of Education every three years. Moreover, openings of new universities required a specific law to be passed by Parliament. University degrees had to meet some criteria fixed centrally by the Ministry of Education, concerning, among other things, their curriculum content. The fact that the system was (and still is) almost entirely public and directly managed by the central government, together with the very little differentiation between the degrees supplied by the different HEIs, led to the legal recognition of degrees in the same field as identical ('valore legale,' i.e. legal value).

On the demand side, until a few decades ago, the student body used to come almost entirely from families with a relatively high socio-economic background. Indeed, educational mobility has historically been lower in Italy than in other developed countries. For example, Checchi et al. (1999) report that less than 2% of people whose father did not complete compulsory schooling end up having a college degree in Italy, while the corresponding figure for the United States is 12%. The evolution from an elite to a mass university system started in 1969, when access to university was liberalized and enrolment in any field became possible for students holding all types of upper secondary school degrees (Law 11 december 1969, n. 910).[7]

On the supply side, the increased demand for higher education led to the foundation of many new HEIs, new Faculties, and new local branches. Indeed, reforms between the late 80s and the early 90s granted a much higher level of autonomy to universities regarding the management of teaching and financial resources. The requirement of parliamentary approval was abandoned in 1990 (Law n. 341 19/12/1990), whereas the inclusion in a university development plan was still retained. However, universities gained autonomy to advance proposals for new initiatives to the Ministry. Many institutions used this new autonomy to open branches in smaller cities and to increase dramatically the number of degrees offered (Bratti et al., 2008; Oppedisano, 2011). The entry of new actors in the higher educational market and the increasing fragmentation of educational provision contributed to enlarge the gap, in terms of quality, between HEIs. However, the Italian university system remains characterized by a much larger variance of quality within departments than between departments in the same field of study (see Bonaccorsi and Cicero, 2015, for a within-between analysis of research quality).

A further step towards a mass tertiary education system was taken in Italy with the completion of the Bologna process and the so called '3+2' reform (Ministerial Decree n. 509/99).[8] The older long degrees were replaced with two levels of degrees, three-year first-level degrees and two-year second-level degrees.[9] The large increase in the supply of degrees offered made more difficult for high school graduates to choose the best possible option given their preferences and constraints. This stimulated a growing interest of prospective students to know the relative quality of institutions and degrees. For this

[7]Before this law only individuals graduating from a specific academic upper secondary school track (*liceo classico*, i.e. classical lyceum) could enrol in all types of tertiary education.

[8]For a brief description of the '3+2' university reform see Di Pietro and Cutillo (2008) and Cappellari and Lucifora (2009).

[9]Other courses were also introduced such as first-level Master Degrees, and second-level Master degrees, but most students enrolled in the first two types of degrees.

reason, two of the main Italian newspapers ('Il Sole 24 Ore' and 'La Repubblica') started about 15 years ago to publish rankings (updated every year) of Italian universities and Faculties.[10]

With a similar purpose, i.e. to evaluate the quality of universities and other research institutions receiving public funds and to diffuse this information among the stakeholders, the Committee for the Evaluation of Research (CIVR) initiated the first Research Evaluation Exercise (VTR) in December 2003. The REE assessed the research produced by 102 Italian institutions (77 universities and 25 research agencies) for the period 2001-2003. The products evaluated were divided in 20 disciplinary areas, the 14 CUN areas plus 6 interdisciplinary sectors.[11] Each university had to send one (autonomously selected) product every 4 researchers, while research agencies were required to submit one product every 2 researchers. The first REE was entirely based on peer review. A total of 17,329 products were evaluated by 6,661 experts (Franceschet and Costantini, 2011). Each product evaluation, by at least two referees, led to 4 possible outcomes: excellent, good, passable and limited. Furthermore, universities communicated data on human resources, international mobility and research funding in order to make a complete and informed assessment possible. The total cost of the REE was around 3.55 million euros.

In contrast to what happened in the United Kingdom with the Research Assessment Exercise (RAE), initially there has been no funding linked to the results of the REE. However, starting from 2009, the results of the REE have been used, together with other indicators of HEIs' performance, to allocate a small share (see Rebora and Turri, 2013, for some details) of public funds.

The final results of the evaluation were released in February 2006, potentially affecting university enrolments from the 2006-2007 academic year. The assessment of each single research product has not been published, it has been disclosed just to Rectors (i.e., Chancelors). The final VTR ranking score has been built as a weighted average, with the number of 'excellent' (E) products multiplied by 1, 'good' (G) products by 0.8, 'passable'

[10]Faculties are the equivalent of Schools in the international context.

[11]CUN stands for *Consiglio Universitario Nazionale* (National University Council). CUN's members are elected to advise the MIUR on matters related to HEIs.

(P) products by 0.6 and 'limited value' (L) products by 0.2. The formula is

$$\text{final VTR score} = \frac{1 * E + 0.8 * G + 0.6 * P + 0.2 * L}{\text{total products evaluated}}. \tag{2.1}$$

This indicator can vary between 0.2, if all products are judged as 'limited value', and 1, in case all products are 'excellent'. For the purpose of the current study we will be using two main indicators of quality. The first is the *final VTR score* computed as described above, and the second is the share of excellent products (i.e. those which obtained the evaluation of 'excellent'). In order to make the results of the estimated regressions easier to read both indicators are included in the econometric models as standardized variables with zero mean and unit standard deviation (s.d., hereafter), so as their coefficients correspond to the percentage increase (as the dependent variable is measured in logarithm) in the dependent variable produced by a one-s.d. increase in the indicator.[12]

2.3 Empirical strategy

Our primary interest lies in the impact of VTR on the number of university enrolments and the quality of students. We use two measures of student quality. The first is the number of students coming from the upper secondary school academic track (*liceo*) and the second is the the number of students with grades in the upper secondary school final exam above 90 (grades vary in the 60-100 interval).[13] We use data on enrolment from year 2002 to 2011.[14] We base our identification strategy on a before-after estimator with differential treatment intensities.[15] The main idea is to look at whether departments

[12]The final VTR score for research quality has been used ny the Ministry of Education to build official rankings of universities in each of the 20 areas. For the purpose of the current study, we focus on the VTR score and not on the official rankings, since the latter were produced by university size groups (large, medium, small). We do not think such classification to be particularly informative to students who are interested in enrolling in high-quality HEIs, although it may be for the Ministry of Education which has to allocate public resources.

[13]Italy has a tracked upper secondary school system. Schools can be divided into three main tracks. The first is represented by the academic track, and we will refer to these schools as the academic high schools. The second is the technical track and the third the vocational track. Students who choose the academic track generally go on in tertiary education.

[14]All our data falls into the post-'3+2' reform period.

[15] Since all higher institutions are subjected to the VTR exactly at the same time, it is not possible to use a difference-in-differences (DID) strategy.

which performed well in the VTR attracted after the VTR a higher number of students and better students compared to the past relative to those departments that did not perform satisfactorily in the research assessment. Our empirical specification is described by the following equation

$$Y_{it} = \alpha_0 + \alpha_{1i} D_i + \alpha_{2jt} D_{jt} + \alpha_3 (V_i * POST_{2006}) + \epsilon_{it} \tag{2.2}$$

where D_i is an indicator defined at the HE institution (a) $\times$ field of study (k) level (which is defined 'department' for the sake of brevity); D_{jt} are province-year fixed effects; V_i a (time-invariant) continuous variable reflecting the score obtained in the VTR and $POST_{2006}$ a post-VTR dummy. In particular, the first academic year affected by the reform was 2006/2007, and starting from this academic year the $POST_{2006}$ indicator takes on value one. In this baseline specification, α_3 captures a higher or lower *level* of the outcome variable (e.g., student enrolments or student quality) after the 2006 for departments which obtained a higher score in the VTR. Department-level time-invariant factors are captured by α_{1i} while local factors (e.g. cost of housing, local unemployment) by α_{2jt}.

Since the information released by VTR may take time to diffuse, we estimate a variant of equation (2.2) in which the VTR term is also interacted with time elapsed since the end of the VTR. The idea is to capture a post-VTR differential *trend* in student enrolments or student quality correlated with the VTR score. The corresponding estimated equation is

$$Y_{it} = \alpha_0 + \alpha_{1i} D_i + \alpha_{2jt} D_{jt} + \alpha_3 (V_i * POST_{2006}) + \alpha_4 V_i * (t - 2006) * POST_{2006} + \epsilon_{it}. \tag{2.3}$$

In this specification, enrolments and student quality are allowed to *grow* differently after 2006 according to the VTR results. There are two main reasons why one might expect an effect of VTR that is increasing overtime. First, as we mentioned, it may take some time before the VTR information is diffused through the media, and students become aware of it. Second, as we said, it was only in the .... that the VTR results were used for the first time to distribute public funding to HEIs. Although in the beginning VTR results accounted for a very small share of the university finances received by the Ministry, the share increased over time up to become .... in 2011.[16] We may expect a two-fold effect from the VTR. The first is mainly reputational, and except for a 'student learning' effect, it should be quite constant in the post-VTR period. The second is related to the desire of students to avoid enrolling in under-funded HEIs, from which we expect the effect of VTR to be increasing over-time following the rise in the share of performance-related funding.

The specifications in (2.2) and (2.3) control for department-level fixed effects, i.e. departments are allowed to start from different intercepts as far as enrolments and student quality are concerned. However, we also estimate a more demanding specification including both department specific intercepts and department specific trends, which allows department to follow different pre-VTR trends in the outcome variables. This may address the concern that departments who saw an increase of enrolments or student quality after the VTR may have been already on an increasing trend before the research assessment. The corresponding specifications are

$$Y_{it} = \alpha_0 + \alpha_{1i} D_i + \gamma_i (D_i * t) + \alpha_{2jt} D_{jt} + \alpha_3 (V_i * POST_{2006}) + \epsilon_{it} \tag{2.4}$$

$$\begin{aligned} Y_{it} = \alpha_0 + \alpha_{1i} D_i + \gamma_i (D_i * t) + \alpha_{2jt} D_{jt} + \alpha_3 (V_i * POST_{2006}) \\ + \alpha_4 V_i * (t - 2006) * POST_{2006} + \epsilon_{it}. \end{aligned} \tag{2.5}$$

where the γ_is are the department specific trends.

Some of the existing literature mentioned in the Introduction has demonstrated that league tables may be more important for the top institutions, while average- or low-

[16]This share will further increase in the future. Article 60 of Decree Law 69/2013 (Law 98/2013) has established that the amount based on a competitive basis shall be not less than 16% for the year 2014, 18% for 2015 and 20% for the year 2016, with subsequent annual increments of no less than 2% and up to a maximum of 30% of public funds (*Fondo di Finanziamento Ordinario*, FFO); of this amount, at least 3/5 are distributed among the universities on the basis of the results achieved in the quality of research rating (VQR) and 1/5 on the basis of the evaluation of recruitment policies.

quality institutions may be little sensitive to rankings. In order to test this hypothesis we dichotomize the VTR score in two quartile dummies, one for the fourth quartile of the quality indicator (Q4), meaning higher quality, and the other for the lowest quartiles (Q1-Q3). Then V_i is replaced with the fourth quartile dummy in all specifications above.[17] The coefficient on the $POST_{2006}$*Q4 has to be interpreted as the differential effect with respect to lower quartiles of quality.

The VTR produced several indicators. In this study we use the overall VTR score and the share of excellent products (see data description). The first is an indicator of the average research quality of a department, while the second is more suitable to capture research excellence.

2.4 Data

Our analysis is based on data from two main sources. Information about the number of students enrolled in each year and department comes from the website of the Ministry of Education and Research (MIUR). This dataset allows us to study the number of enrolled students also by upper secondary school final grade and track. We focus our analysis on students enrolled in first level degrees. The choice is dictated by two main reasons. First of all, while many second cycle degrees admit a fixed number of students per year the same is not true for first level degrees, where access was free almost everywhere in Italy in the period that we consider. Since we are interested in the effect of research quality on enrollment, restrictions on the number of accepted students would be a potential threat to our identification strategy. We expect in particular that HEIs with a better score in the REE will tend to rely more on selective admissions, leading to a bias in our estimates. Secondly, the two indicators for the quality of enrolled students that we use are likely to be better proxies or student quality before starting first-level degrees, while for second-level degrees the final grade in the first-university degree is likely to be a better proxy of student ability. Unfortunately, the latter is not available in the data.

[17] Only one of the two interactions between quartile dummies and the post-VTR period can be included in the regression, as when are both included the two dummies are collinear with the province-year fixed effects.

The second data source is the report released by the Committee for the Evaluation of Research (CIVR) in February 2006. The document contains information on research quality divided by scientific areas for 77 universities. We decided to focus our attention on two measures of research quality, the final VTR score described in equation (2.1) and the share of excellent products.

Before running the analysis the two sources of data had to be merged. Indeed, although data are recorded for different 'fields,' they are different in the enrolment data (MIUR), where a classification based on teaching is adopted, and in the VTR data, where research fields are instead used. The mapping was done as described in Table 2.1. Merging the two datasets, we managed to obtain complete information about enrolment and research quality for 518 Faculties in 72 universities.

In Figure 2.1 we plot the variation between and within institutions in the final VTR score. The graph presents the lowest, the average and the highest score obtained by each institution. A large majority of universities have quite similar average values of the score, while a much larger variation occurs between fields of study within the same institution. Just to take an example, the University of Catanzaro obtained a maximum score of 0.87 in Biology and a minimum score of 0.2 in Economics, with an average score of 0.656. This makes it clear the advantages of shaping the analysis at the field of study level, since averaging enrolments and REE scores at the HEIs would wash out most of the variation. The relatively low amount of variation in the final VTR score is partly due to the design of the REE, since the number of products to be sent for the evaluation was quite low, one every four researchers.[18] Figure 2.2 presents the same information as Figure 2.1 for the second indicator of research quality, the percentage of products that were evaluated as excellent in each department. For this indicator the variance is larger, with many departments presenting no excellent product and some others for which all the research output sent was judged as excellent. Figure 2.3 presents the variation in the two measures of research quality, the VTR score and the percentage of excellent products,

[18]This changed in the following REE. In the VQR 2004-2010, each university research staff had to submit three research products.

between provinces.[19] A clear geographical divide emerges, with most institutions in the top 20 located in the North of Italy and the majority of institutions with low scores located in the South.

Figure 2.4 visualizes by plotting the raw data the kind of empirical exercise that we do in this paper. The figure plots the average number of enrolled students per year in departments who got a low (first quartile) vs. a high (fourth quartile) score in the VTR. The number of enrolled students per department decreased significantly during the period in both groups. However, the reduction was larger for departments that received a bad evaluation, i.e. with a score in the first quartile, in the REE, and a large share of the divergence took place right after the publication of the results. Thus the effect of a better VTR rating on enrolments appears to be positive in the raw data. The falling trend for the whole period is also evident for students graduating from high school with a high mark[20] while for graduates of academic high schools the initial decrease in enrolment is compensated by a similar increase between 2007 and 2011 for both high and low research quality departments.

2.5 Results

Each table of results in this Section consists of two panels. Panel A reports the results using the VTR score and panel B those using the share of excellent VTR products. Models in columns (1) and (2) only allow for an intercept effect of the VTR, the first including only department FEs and the second also department-specific trends. Models in columns (3) and (4) also allow for a differential trend after the VTR according to a HEI's performance in the VTR.

The results in columns (1) and (2) of Table 2.2 show an interesting pattern. While the VTR score does not seem to be associated with total (log) student enrolments, irrespective

[19] In Italy, a province (provincia) is an administrative division of intermediate level between a municipality (comune) and a region (regione). They correspond to NUTS-3 in the Eurostat's Nomenclature of Territorial Units for Statistics (NUTS) classification.

[20] Since students in Southern Italy have on average higher marks in the high school final exam (see Montanaro (2008)), this trend may just reflect a more sustained negative trend for HEIs located in the South of the country.

of the controls included, student enrolments are sensitive instead to the share of excellent VTR products. A possible interpretation is that given the limited number of research products submitted, the VTR score was less able to discriminate quality than the share of excellent products, which indeed exhibits larger variation both between and within HEIs (see Figure 2.3). Raising the latter by one-s.d. increases enrolments by 6.2 percent when only controlling for department FEs and by 5.8 percent when department trends are included. Allowing the post-2006 trend in enrolments to differ according to the VTR outcomes leads to similar estimates, 0.058 and 0.063, in column (3) and (4) respectively. The VTR score-post 2016-time trend interactions are never statistically significant. hence, there is no evidence of an increasing reputational effect overtime or of the VTR effect to be increasing with the share of funding allocated on a competitive basis (see Section 3.4). In what follows, we consider the more general model estimated in column (4) as our preferred specification.

Tables 2.3 and 2.4 investigate potential differential effects by students' entry qualifications.[21] We expect especially better students to respond to the new information concerning HEIs' quality. Results in Table 2.3 relate to the enrolment of high-mark students (i.e. those who graduated from high school with a mark of 90 or more, out of maximum of 100) and confirm those of the previous table regarding the salience of the share of excellent products compared to the VTR score. The estimated effects are very robust across specifications also in this case. Our preferred specification in column (4) returns a coefficient of 0.103, statistically significant at the 5% level. It is worth noting that the point estimate is larger than that obtained in the previous table, demonstrating the higher responsiveness of high-quality students to rankings with respect to the average student.[22] Results in Table 2.4 mimic those in Table 2.3. Our preferred model indicates that increasing the share of excellent VTR products by one s.d. raises the number of students coming from the academic track by 11.8 percent. These results are consistent with the evidence coming from individual level data that in Italy are especially the most talented students, irrespective of their family backgrounds, who value quality when making

[21]In these regressions we lose 48 observations for which data on student quality is not available.
[22]However, the two coefficients are not statistically different.

their university enrolment decisions (Pigini and Staffolani, 2015).

A robust finding of the past literature is that rankings especially affect enrolments of HEIs in the top of the quality distribution, while the other institutions are less sensitive to the release of quality information. We investigate this hypothesis by allowing the VTR coefficients to change between the fourth quartile (Q4) and the other quartiles (Q1-Q3) of the VTR score's and share of excellent VTR products' distributions, using interaction terms as described in Section 3.4. Interestingly, Table 2.5 shows that the effect of ranking in the fourth quartile of the VTR score on total enrolments is now statistically significant, and points to a 11.9 percent increase in enrolments compared to HEIs in lower quartiles of quality (column 4). Similarly, being in the fourth quartile of the share of excellent products produces a 14.5 percent increase in total enrolments.

The effects of being in the top quartile of the VTR on enrolments of high quality students are even larger. Column (4) of Table 2.6 shows that performing in the top quartile of the VTR score (share of excellent products) raises the number of enrolments of high-mark students by 20.4 (20) percent. Effects of similar magnitude are found in column (4) of Table 2.7 on enrolments of students coming from the academic high school track. The magnitude of the effect of having the VTR score (share of excellent products) in the fourth quartile is 0.25 (0.26) log points. All in all, these last estimates show, in line with the past literature on league tables, the existence of substantial non-linearities in the effects of rankings.

Since the direct impact of VTR on public funding was initially very limited, we expect the effect of the REE results on student enrolment, if any, to be the result of a change in HEIs' reputation. Given that student mobility increased in Italy over the last decade (De Angelis et al., 2016) [23] it is important to study if the phenomenon is due to a better awareness of differences in quality between HEIs or, since the trend accelerated dramatically after 2008, it can simply be explained by the fact that student increasingly prefer to enrol in universities located in labor markets offering more opportunities. Indeed,

[23] Mobility of high school graduates to HEIs in other areas of the country increased everywhere but the North-West. The area experiencing the largest growth is the South, where the share of high school graduates enrolling in the Centre-North increased from 16.5% in 2008 to 22.3% in 2014.

the search for quality is not the only, and probably not even the most important, factor motivating students' geographical mobility, and the state of the labour market both at origin and at destination plays an important role (Dotti et al., 2013). However, there are concerns that a bad performance of Southern Italy's universities in REEs may exacerbate the brain drain and increase South-North migrations of university students. One could expect especially Northern universities to gain from a good result in the VTR. Indeed, they are likely to enjoy a 'double dividend' from a high ranking in REEs by attracting a higher number of local students but also more external students (i.e. students from other regions). Southern students whose have decided to move out of their regions may, for instance, change their enrolment choices with respect to the past in favour of high-rated faculties. Southern universities, in contrast, are more likely to compete in more local markets and to enjoy much lower gains from a good ranking in the REE. Table 2.8 explores this hypothesis, by reporting estimates split by geographical area (North, Centre, South and Islands). The first three columns show no significant association between total enrolments and the VTR score for all geographical areas. When considering the percentage of excellent VTR products, in panel A, only the coefficient for North Italy is statistically significant at the 10% level. The coefficients for the two remaining areas, especially that for South Italy, are not very dissimilar in magnitude, but are much less precisely estimated. The same happens considering the number of enrolments of high-mark or academic-track students. As for the former, the coefficient on North (0.134) is statistically significant at the 5% level, while the coefficient for South in spite of having a similar magnitude (0.138) has a standard error that is two times larger. The percentage of excellent VTR products increases the enrolments of academic-track students by 0.174 log-points (statistically significant at 5%). The coefficient on the South is a bit smaller (0.133) but statistically insignificant. All in all, we conclude that there is no striking evidence that the VTR was only effective for the North of Italy, as the estimates for the other regions generally are not close to zero but are imprecisely estimated.

2.6 Concluding remarks

The Italian Higher Education system has always been characterized by the so-called "legal value" of university degrees. The degree content being strongly regulated at the central level by the Ministry of Education, all university degrees in the same field were (and still are) considered as formally equivalent. However, the progressive transition to mass tertiary education has been accompanied by a very rapid increase in the number of HEIs and degrees supplied, often leaving students with little guidance on the real value of the educational programmes offered. This lack of information has been exploited by private intermediaries —in Italy by newspapers— which have published annually universities' league tables. Although the impact of "unofficial" university rankings on student choice has been already object of several studies in the US and the UK, and to a smaller extent also in Italy, the same cannot be said for official ranking exercises.

This paper focuses on the first Research Evaluation Exercise (VTR) that was completed in Italy in 2006, and features the first assessment of its impact on student choice, namely on the total number of university enrollments and on enrolments of high quality students, proxied by high school mark and provenance from the academic high school track. To the best of our knowledge, our paper also represents the first study assessing the effect on student enrolment choices of establishing a REE.

We relate the number of enrolments at the departmental level to VTR ratings using a "differential intensity" before-after estimator. In particular, we investigate whether departments with a better VTR rating also had better enrollment outcomes after 2006.

Our analysis shows that final VTR scores did not affect the number and quality of students enrolled, while our second indicator of research quality, the percentage of excellent products, had a positive and significant effect on enrollment. In our preferred specification, a s.d. increase in the share of excellent research products leads to a 6.3% increase in total enrolments, a 10.3% increase on enrolments of students graduating from upper secondary school with a high mark (at least 90 out of 100) and a 11.8% increase for students coming from the academic high school track. We find, in line with the previous literature, that the coefficients are highly non-linear, with most of the effect occurring in the fourth quartile of the research quality distribution. In fact, the total number of students enrolled is 11.9% and 14.5% higher for HEIs in the fourth quartile with respect to HEIs in lower quartiles of quality for the first and second indicator of quality, respectively. The size of the effect is larger when we focus on high quality students. For high school graduates with a high mark the HEIs in the fourth quartile experience an increase in enrolments of around 20% for both measures of research quality, while for students coming from an academic track the effect is even larger at 25% and 25.8% for the VTR score and the share of excellent products, respectively.

Finally, we find that the effect of the VTR on enrolments appears to be very similar in the North and in the South of the country, although it is precisely estimated only in the former.

Some cautionary notes are in order. First, unlike with the following REEs, the VTR did not link initially the rating performance to the amount of public funding received by institutions. In this respect, we interpret its effect as being mainly "reputational." The VTR rating may have been informative to students who where planning to search employment in very high-skilled or research-related jobs. Indeed, research quality may positively impact on the labour market outcomes of these university graduates (**??**). However, the effect of the following VQR could be stronger, since a more substantial share of universities' public budgets were distributed according to their results. Within a few years 30% of the total public budget will be allocated according to quality indicators, and research quality, with a weight of 65% in the determination of the total score, will be the main determinants of these funds (Decree Law 69/2013). Second, since a similar evaluation of university teaching was not in place during the same period, a possible reading of our results is that in the absence of reliable information on teaching quality, students were using research rankings and ratings from the REE as proxies for the quality of teaching. However, little is known about the complementary between teaching and research activity (**??**). The two may also be substitutes, and once students realize it, outcomes in REEs may even become negatively correlated with the enrolment of students who mostly care about teaching.

Table 2.1: Mapping of VTR to Area-Mixed

Area Mixed	Disciplinary Area (VTR)	Field name
1	1+2	Hard sciences (math and physics)
2	3	Chemistry
3	4+5+15e	Biology
4	6	Medicine
5	7+15b	Agriculture
6	8	Architecture
7	9+15c+15d	Engineering
8	10+15f	Humanities
9	11	Teaching and Psychology
10	12	Law
11	13	Economics
12	14+15a	Political sciences

In the first column (Area Mixed) we show the 12 areas that we use in the analysis. They result from merging the Disciplinary Areas in the VTR and the fields of study as classified by the Ministry of Education, University and Research (MIUR). The disciplinary areas in the VTR are the 14 CUN areas (1 - Mathematics and Computer Sciences, 2 - Physics, 3 - Chemistry, 4 - Earth Sciences, 5 - Biology, 6 - Medicine, 7 - Agriculture, 8 - Architecture, 9 - Industrial Engineering, 10 - Humanities, 11 - Teaching and Psychology, 12 - Law, 13 - Economics and 14 - Political Sciences) plus 6 inter-disciplinary areas (15a - Science of information and communication, 15b - Science for food quality and safety, 15c - Science for Nano-Microsystems, 15d - Aerospace sciences, 15e - Science for sustainable development and governance, 15f - Science for the evaluation and enhancement of cultural heritage).

Table 2.2: Effect of VTR on total (log) enrolled students

	(1)	(2)	(3)	(4)
Panel A. VTR score				
VTR*$Post_{2006}$	0.024	0.021	0.014	-0.004
	(0.024)	(0.022)	(0.024)	(0.030)
VTR*$Post_{2006}$*(Time-2006)		0.001		-0.013
		(0.005)		(0.011)
Number of obs	7302	7302	7302	7302
R^2	0.840	0.840	0.887	0.887
Panel B. % Excellent products				
VTR*$Post_{2006}$	0.062***	0.058***	0.058**	0.063**
	(0.024)	(0.022)	(0.023)	(0.026)
VTR*$Post_{2006}$*(Time-2006)		0.002		0.003
		(0.005)		(0.010)
Number of obs	7302	7302	7302	7302
R^2	0.841	0.841	0.887	0.887
control variables (both panels):				
Department FE	Yes	Yes	Yes	Yes
Province × year FE	Yes	Yes	Yes	Yes
Department specific time trends	No	No	Yes	Yes

*, **, *** statistically significant at the 10, 5 and 1% level. Standard errors are clustered at the Department level.

Table 2.3: Effect of VTR on (log) enrolments of high-mark students

	(1)	(2)	(3)	(4)
Panel A. VTR score				
VTR*$Post_{2006}$	0.065**	0.026	-0.005	-0.016
	(0.031)	(0.034)	(0.042)	(0.053)
VTR*$Post_{2006}$*(Time-2006)		0.016**		-0.009
		(0.007)		(0.018)
Number of obs	7254	7254	7254	7254
R^2	0.778	0.778	0.835	0.835
Panel B. % Excellent products				
VTR*$Post_{2006}$	0.102***	0.082**	0.083**	0.103**
	(0.031)	(0.032)	(0.036)	(0.042)
VTR*$Post_{2006}$*(Time-2006)		0.008		0.015
		(0.007)		(0.016)
Number of obs	7254	7254	7254	7254
R^2	0.779	0.779	0.835	0.835
control variables (both panels):				
Department FE	Yes	Yes	Yes	Yes
Province × year FE	Yes	Yes	Yes	Yes
Department specific time trends	No	No	Yes	Yes

*, **, *** statistically significant at the 10, 5 and 1% level. Standard errors are clustered at the Department level.

Table 2.4: Effect of VTR on (log) enrolments of academic-track students

	(1)	(2)	(3)	(4)
Panel A. VTR score				
VTR*$Post_{2006}$	0.012	0.028	0.030	-0.001
	(0.033)	(0.033)	(0.039)	(0.052)
VTR*$Post_{2006}$*(Time-2006)		-0.006		-0.023
		(0.007)		(0.019)
Number of obs	7254	7254	7254	7254
R^2	0.801	0.801	0.854	0.854
Panel B. % Excellent products				
VTR*$Post_{2006}$	0.074**	0.102***	0.122***	0.118***
	(0.034)	(0.034)	(0.038)	(0.044)
VTR*$Post_{2006}$*(Time-2006)		-0.011		-0.004
		(0.008)		(0.015)
Number of obs	7254	7254	7254	7254
R^2	0.801	0.801	0.855	0.855
control variables (both panels):				
Department FE	Yes	Yes	Yes	Yes
Province × year FE	Yes	Yes	Yes	Yes
Department specific time trends	No	No	Yes	Yes

*, **, *** statistically significant at the 10, 5 and 1% level. Standard errors are clustered at the Department level.

Table 2.5: Effect of VTR on total (log) enrolled students by quartile of HEIs' 'quality'

	(1)	(2)	(3)	(4)
Panel A. VTR score				
$Post_{2006}$*Q4	0.178***	0.154***	0.128**	0.119**
	(0.058)	(0.051)	(0.054)	(0.059)
$Post_{2006}$*(Time-2006)*Q4		0.010		-0.006
		(0.012)		(0.022)
Number of obs	7302	7302	7302	7302
R^2	0.841	0.841	0.887	0.887
Panel B. % Excellent products				
$Post_{2006}$*Q4	0.174***	0.150***	0.133***	0.145***
	(0.052)	(0.046)	(0.049)	(0.052)
$Post_{2006}$*(Time-2006)*Q4		0.010		0.009
		(0.012)		(0.022)
Number of obs	7302	7302	7302	7302
R^2	0.841	0.841	0.887	0.887
control variables (both panels):				
Department FE	Yes	Yes	Yes	Yes
Province × year FE	Yes	Yes	Yes	Yes
Department specific time trends	No	No	Yes	Yes

*, **, *** statistically significant at the 10, 5 and 1% level. Standard errors are clustered at the Department level. Q4 stands for the fourth quartile of the VTR score or percentage of excellent products distribution.

Table 2.6: Effect of VTR on (log) enrollments of high-mark students by quartile of HEIs' 'quality'

	(1)	(2)	(3)	(4)
Panel A. VTR score				
$Post_{2006}$*Q4	0.239***	0.222***	0.210***	0.204**
	(0.077)	(0.074)	(0.077)	(0.082)
$Post_{2006}$*(Time-2006)*Q4		0.007		-0.005
		(0.016)		(0.033)
Number of obs	7254	7254	7254	7254
R^2	0.779	0.779	0.835	0.836
Panel B. % Excellent products				
$Post_{2006}$*Q4	0.209***	0.191***	0.188***	0.200***
	(0.068)	(0.066)	(0.069)	(0.073)
$Post_{2006}$*(Time-2006)*Q4		0.008		0.009
		(0.016)		(0.032)
Number of obs	7254	7254	7254	7254
R^2	0.779	0.779	0.835	0.835
control variables (both panels):				
Department FE	Yes	Yes	Yes	Yes
Province × year FE	Yes	Yes	Yes	Yes
Department specific time trends	No	No	Yes	Yes

*, **, *** statistically significant at the 10, 5 and 1% level. Standard errors are clustered at the Department level. Q4 stands for the fourth quartile of the VTR score or percentage of excellent products distribution.

Table 2.7: Effect of VTR on (log) enrollments of high-mark students by quartile of HEIs' 'quality'

	(1)	(2)	(3)	(4)
Panel A. VTR score				
$Post_{2006}$*Q4	0.196**	0.271***	0.301***	0.250***
	(0.080)	(0.074)	(0.081)	(0.091)
$Post_{2006}$*(Time-2006)*Q4		-0.031*		-0.037
		(0.017)		(0.034)
Number of obs	7254	7254	7254	7254
R^2	0.801	0.802	0.855	0.855
Panel B. % Excellent products				
$Post_{2006}$*Q4	0.186***	0.253***	0.287***	0.258***
	(0.071)	(0.066)	(0.072)	(0.084)
$Post_{2006}$*(Time-2006)*Q4		-0.028*		-0.022
		(0.016)		(0.034)
Number of obs	7254	7254	7254	7254
R^2	0.801	0.801	0.855	0.855
control variables (both panels):				
Department FE	Yes	Yes	Yes	Yes
Province × year FE	Yes	Yes	Yes	Yes
Department specific time trends	No	No	Yes	Yes

*, **, *** statistically significant at the 10, 5 and 1% level. Standard errors are clustered at the Department level. Q4 stands for the fourth quartile of the VTR score or percentage of excellent products distribution.

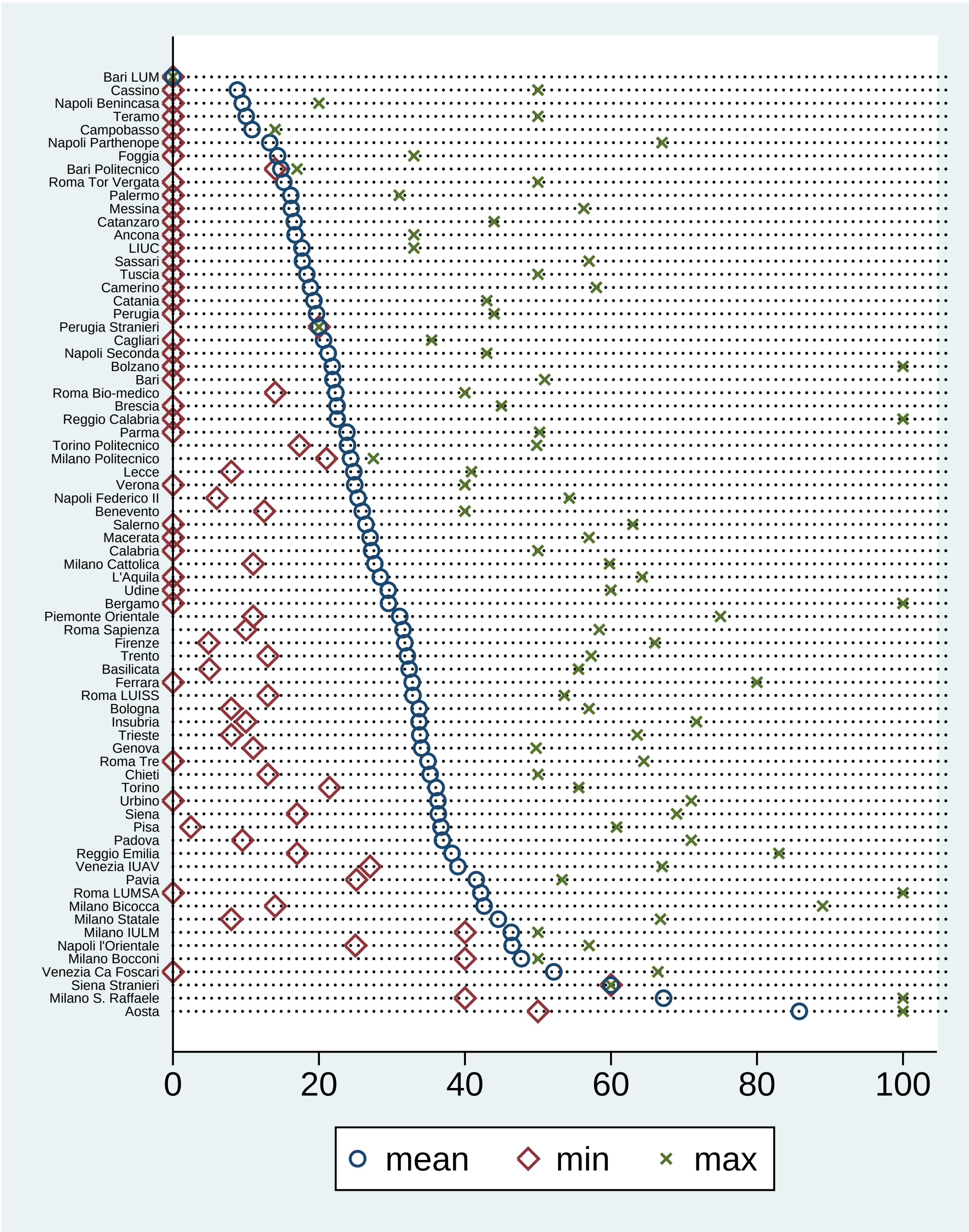
Table 2.8: Effect of VTR on (log) enrolment of students by geographic area

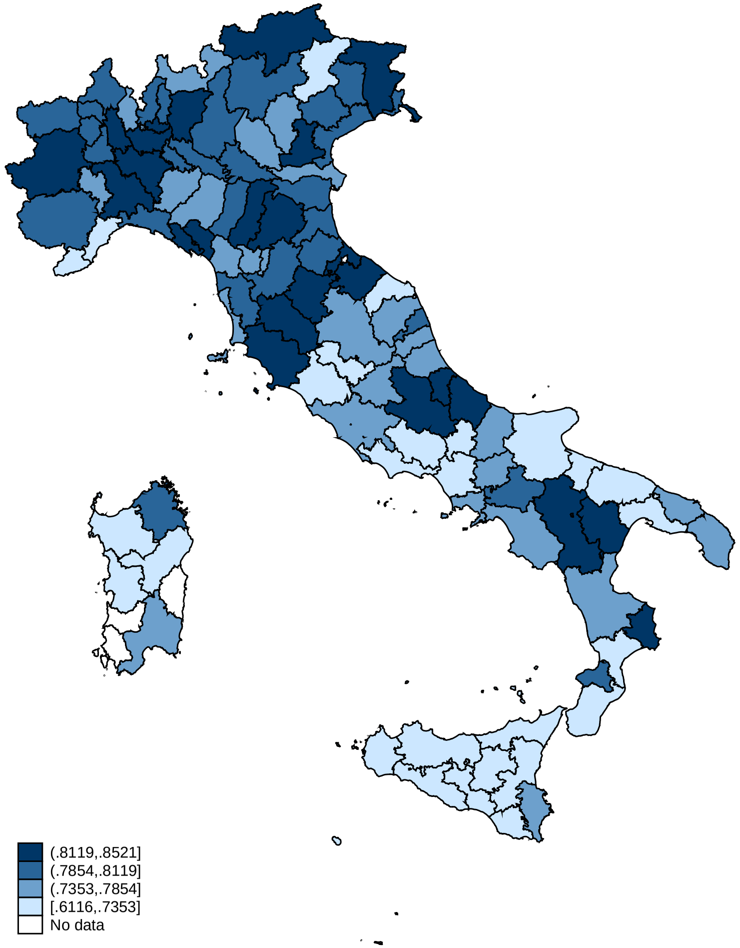
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. All students						
	VTR score			% Excellent		
	North	Centre	South	North	Centre	South
VTR*$Post_{2006}$	0.037	-0.010	0.032	0.064*	0.045	0.104
	(0.046)	(0.035)	(0.053)	(0.037)	(0.038)	(0.066)
VTR*$Post_{2006}$*(Time-2006)	0.001	0.002	-0.017	0.002	0.008	0.008
	(0.017)	(0.016)	(0.018)	(0.013)	(0.015)	(0.022)
Number of obs	2803	2114	2254	2803	2114	2254
R^2	0.876	0.883	0.909	0.876	0.883	0.909
Panel B. High-mark students						
	VTR score			% Excellent		
	North	Centre	South	North	Centre	South
VTR*$Post_{2006}$	0.097	-0.026	0.007	0.134**	0.056	0.138
	(0.064)	(0.057)	(0.088)	(0.054)	(0.060)	(0.102)
VTR*$Post_{2006}$*(Time-2006)	0.015	0.011	-0.003	0.014	0.006	0.038
	(0.022)	(0.023)	(0.026)	(0.019)	(0.024)	(0.038)
Number of obs	2793	2094	2236	2793	2094	2236
R^2	0.836	0.822	0.868	0.837	0.822	0.868
Panel C. Academic-track students						
	VTR score			% Excellent		
	North	Centre	South	North	Centre	South
VTR*$Post_{2006}$	0.145	-0.007	0.000	0.174**	0.066	0.133
	(0.089)	(0.046)	(0.079)	(0.071)	(0.060)	(0.083)
VTR*$Post_{2006}$*(Time-2006)	0.021	-0.009	-0.034	0.007	-0.010	0.002
	(0.026)	(0.024)	(0.025)	(0.020)	(0.027)	(0.030)
Number of obs	2793	2094	2236	2793	2094	2236
R^2	0.853	0.837	0.882	0.854	0.838	0.882
control variables (all panels):						
Department FE	Yes	Yes	Yes	Yes	Yes	Yes
Province × year FE	Yes	Yes	Yes	Yes	Yes	Yes
Department specific time trends	Yes	Yes	Yes	Yes	Yes	Yes

*, **, *** statistically significant at the 10, 5 and 1% level. Standard errors are clustered at the Department level.

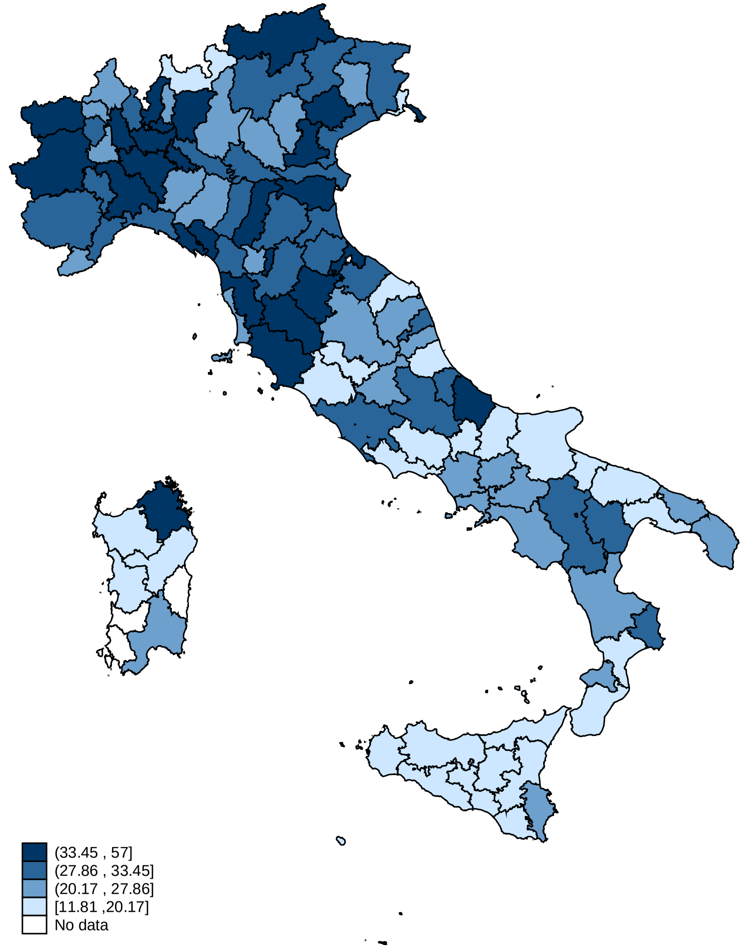
Figure 2.1: VTR final score by university

Figure 2.2: Percentage of excellent VTR products by university





(a) VTR score



(b) Percentage of VTR excellent products

Figure 2.3: Province-level variation in VTR research quality

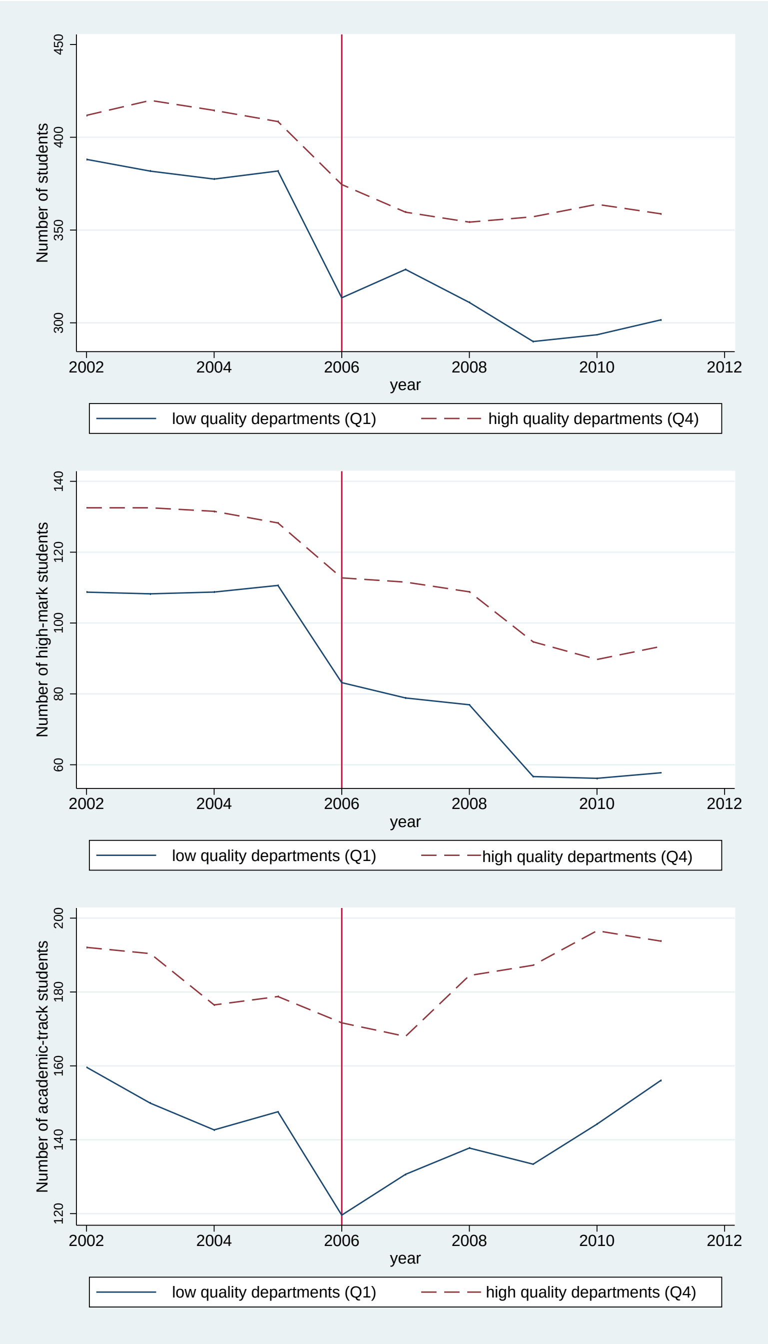


Figure 2.4: Average number of students enrolled by year for 1st (Q1) and 4th (Q4) quartiles of VTR score

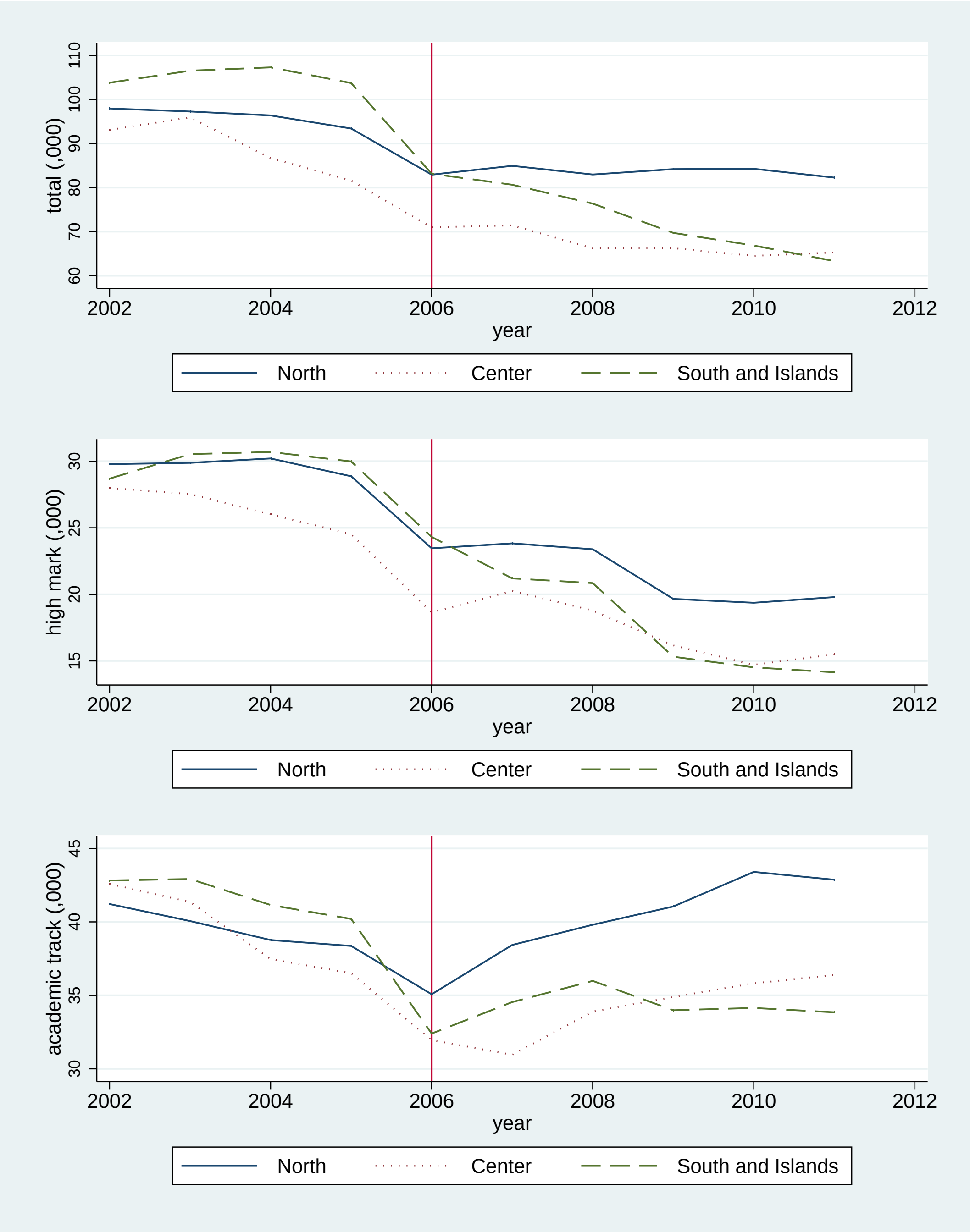


Figure 2.5: Number of enrolled students by year and geographic area

Chapter 3

Household labor supply: a structural estimation for Italian cohorts

ABSTRACT - This chapter analyzes the effect of a childbirth subsidy on fertility, employment and wages of both husband and wife. We build a model of labor supply and fertility choices within household following the Eckstein et al. (2016) framework. Then, we estimate the parameters of the model using Italian data for the 1960 cohort from the Survey of Household Income and Wealth and the Simulated Method of Moments. The model is able to explain quite well the behavior of men and women in the cohort. Preliminary results show that the permanent childbirth transfer is successful in increasing the total fertility rate of married women, even if it has a negative effect on employment.

JEL Numbers:

Keywords:

3.1 Introduction

Over the last decades the gender gap in terms of education disappeared in many developed countries and now women are on average more educated than men. Despite this improvement, married women are still working much less and earning lower wages than their husbands, even if there are large differences between countries (Arulampalam et al., 2007). At the same time fertility decreased sharply and in some European countries it reached such low levels (Kohler et al., 2002) that governments are starting to study subsidies to boost the number of childbirths without pushing women out of the labor force. In this paper we build and estimate a discrete-choice stochastic dynamic programming model in order to study the life-cycle labor supply and fertility decisions of Italian married men and women. For the estimation we use the Survey of Household Income and Wealth (SHIW) data from 1984 to 2014. We focus on the behavior of the 1960 cohort, composed of people born in 1957-63.

We start from a simplified version of the Eckstein et al. (2016) model, which extends the Keane and Wolpin (1997) model for men and Keane and Wolpin (2010) for women. Men and women make their decisions jointly as a married couple from age 28, that is when couples are formed in the model, until a final age of 57. They choose both labor supply and fertility for the first 12 periods (no children after age 40), then they just decide the amount of labor supply for the remaining periods. The model is estimated on repeated cross-sections holding preference parameters fixed.

Structural estimation offers some important advantages with respect to reduced form approaches. First, it allows to model different sources of endogeneity (ex. self-selection into labor market participation). Second, it provides parameters from a theoretical model that can be used to simulate the effects of policy experiments. We can perform an ex-ante evaluation of policies before implementation, studying both the short term an long term effects of a given set of policies.

In this case we use the estimated model to run the following policy experiment: we simulate the impact of a childbirth transfer similar to the 'bonus bebe' recently introduced by the Italian government. We look at the long run effects of the subsidy in terms of total fertility, wages and labor market experience. The only structural paper making a similar evaluation, up to our knowledge, is Adda et al. (2015).

Italy is an interesting case study because presents an employment rate of women significantly below the European average and, at the same time, a very low fertility rate. In fact, the employment rate of women aged 25-54 was 57.6% in Italy in 2014, around

20 percentage points lower than in France, Germany and the United Kingdom (Marino et al., 2016). The coexistence of very low fertility and low participation to the labor market, makes it necessary to find policies that stimulate fertility without harming female employment. Many women in Italy exit the labor market after the birth of their first child and never go back to work. One possible explanation is the lack of adequate social policies (Del Boca and Sauer, 2009).

Our model builds on the structural literature about labor supply and fertility. The first contribution to this literature (Eckstein and Wolpin, 1989) focuses on the employment choice of married women in their post-fertility stage of life and consider the labor supply behavior of husbands as exogenous.
Van der Klaauw (1996) expands the model making the marital status decision of the woman endogenous, while Keane and Wolpin (1997) model the career decisions of young men. In Francesconi (2002) the joint modeling of fertility and labor supply of women is introduced as well as the distinction between part-time and full-time jobs. Keane and Wolpin (2010) estimate a model in which women make sequential joint decisions about school attendance, work, marriage, fertility and welfare participation.
Some recent contributions (Eckstein and Lifshitz, 2015) take a game-theoretic approach and explain the heterogeneity in married women participation to the labor market as the result of different games within the household.
A strand of the literature (Attanasio et al., 2008; Eckstein and Lifshitz, 2011; Eckstein et al., 2016) takes a cohort perspective to explain some long term trends, like the increased labor market participation of married women.

Say what is the contribution (value-added) of our paper...Probably the first to model labor supply and fertility in Italy? Del Boca for females only, we have a joint model. Check.
One of the very few studies to assess dynamic effects to subsidies to fertility.

The main findings of our paper can be summarized as follows. First, we show that the estimated model is able to replicate quite well the behavior of households in real data.

Estimates of the main parameters are in line with the previous literature. Second, the simulated childbirth transfer is successful in increasing fertility. However, especially for couples that are young when the policy is introduced, the increase in total fertility comes at the expense of a slight decrease in the time spent on the labor market for women.

The paper proceeds as follows. Section 3.2 describes the facts that motivate our paper, in particular the labor supply behavior of married men and women in different Italian cohorts. We also describe the policy that is used as a benchmark for the experiment. Section 3.3 presents the model. The following Section, Section 3.4, explains the methods used to solve and estimate the model. Section 3.5 describes the data used in the empirical analysis, whose results are commented in Section 3.6. Finally, Section 3.7 summarizes the main findings and concludes.

3.2 Motivation

3.2.1 Facts to explain

Italy is still lagging behind most other developed countries in terms of married women participation to the labor force. Surprisingly Italian women also experience one of the lowest fertility rates in the world. In this section we describe, using the SHIW dataset, some changes in the behavior of Italian married people over time. We define here 3 cohorts of married men and women, the 1950 cohort (born between 1947 and 1953), the 1960 cohort (1957-1963) and the 1970 cohort (1967-1973). In figure 3.1a we look at the employment rate for married women in the 3 cohorts between age 28 and 57. Since we have repeated cross-sections for 30 years we are unable to compare the complete life-cycle labor supply behavior of different cohorts. Indeed, we have data covering the whole time interval just for the 1960 cohort. However, it is possible to detect some clear trends. The more striking fact is the large increase in the labor force participation of women above age 50 between the 1950 and 1960 cohorts, with employment growing from less than 40%

to almost 50% for women aged 57 [1]. We also observe a higher labor supply of women below age 40 in the 1970 cohort with respect to the 2 older cohorts.

In figure 3.1b we look at real wages for women in the 3 cohorts. Wages increase significantly moving from the 1950 cohort to the younger cohorts. The increase in the average level of education is partly responsible for this trend. Higher wages offered on the market contribute to explain the increase in the employment rate between cohorts.

We observe the labor supply of men in figure 3.2a. The employment rate of married men is well above 90% until the late 40ies. Labor market participation, as for women, decreases faster after age 50 for the 1950 cohort than for the 1960 cohort. Real wages, in figure 3.2b, follow the same trend seen for women.

3.2.2 The policy experiment

Our goal is to study the effects of a childbirth transfer on household behavior. We simulate the introduction of a subsidy to fertility very similar to the one introduced by the Italian government at the end of 2014 [2].

All couples with children born starting from January 1st, 2015 and with household income below a given threshold [3] are eligible for the transfer. The bonus consists of a payment of 80 euros per month for the first 3 years of the baby's life.

In order to avoid an excessive level of computational complexity we evaluate the effect of a policy that provides the entire cash transfer at birth. We use our model to simulate the effects on women of different ages (28, 33 and 38) when the policy starts.

[1]The observed change in employment is mostly due to the recent pension reforms enacted by the Italian governments.

[2]The policy, called Bonus Bebe, was introduced in December 2014, by law 190/2014, that is the law regulating the public budget for year 2015. 500 millions have been allocated to finance the policy for the first year, 2015, increasing to a peak of 1.5 billions in the following years.

[3]Bonus bebe' is directed to all households with a measure of equivalent household income, called ISEE, below 25000. The threshold is not low for the Italian standards, making a large majority of couples eligible. The bonus doubles if ISEE is below 7000.

3.3 Model

In our model a married couple is formed at age 28, then the spouses make joint decisions about labor supply of both the husband and the wife and fertility until the terminal period, at age 57. Explain better.

The problem of a married couple

We describe here the optimization problem of a married couple. Let t denote the annual time period and j=f,m denote gender. Individuals have 1 unit of time per period. The time is split between work on the market, h, and time at home, l, so:

$$h_t^j + l_t^j = 1 \qquad j = f, m$$

In each period an individual can work full-time, part-time or stay out of the labor market spending all the time endowment at home. Thus:

$$h_t^j \in \{0, 0.5, 1\} \qquad l_t^j \in \{0, 0.5, 1\}$$

Married couples have three choice variables: $\{l_t^m, l_t^f\}$ and pregnancy p_t. We assume that the decision to have a pregnancy leads to childbirth in the next period with probability one.

Let X_t^j denote work experience, and N_t denote the number of children under 18. The laws of motion for the two variables are:

$$X_{t+1}^j = X_t^j + h_t^j \qquad N_{t+1}^j = N_t + p_t - p_{t-18}$$

Preferences and constraints

Married couples have total income :

$$Y_t^M = w_t^m h_t^m + w_t^f h_t^f + b_m I[h_t^m = 0] + b_f I[h_t^f = 0]$$

where w_t^j are wage rates and b_j is the unemployment benefit plus the value of home production.

The household budget constraint is:

$$C_t^M = (1 - \theta(N_t))Y_t^M$$

where N_t is the number of children (under 18) and $\theta(N_t)$ is the fraction of household income spent on children. In every period, the utility of a married individual of age t and gender j is:

$$U_t^{jM}(\Omega_{jt}) = \frac{1}{\alpha}(\psi C_t^M)^\alpha + L_j(l_t^j) + \pi_t^M p_t + A_j^M Q(l_t^f, l_t^m, Y_t^M, N_t)$$

where:

$$L_j(l_t^j) = \frac{\beta_{jt}}{\gamma}(l_t^j)^\gamma + \mu_{jt} l_t^j \qquad \gamma < 1, \alpha < 1$$

The parameter $\psi \in (\frac{1}{2}, 1)$ captures economies of scale in consumption and $\beta_{jt} > 0$ shifts tastes for leisure. For women β_{jt} depends on p_t, for both depends on education and health status. Stochastic variation in the marginal utility of l is captured by $\mu_{jt} l_t^j$. The third term $\pi_t^M p_t$ captures the utility from pregnancy and the fourth term $A_j^M Q(l_t^f, l_t^m, Y_t^M, N_t)$ the utility a couple receives from quality and quantity of children. We assume that the stochastic process for tastes for leisure is:

$$ln(\mu_{jt}) = \tau_{0j} + \tau_{1j} ln(\mu_{j,t-1}) + \tau_{2j} p_{t-1} + \epsilon_{jt}^l l_t^j$$

where $\epsilon_{jt}^l \sim N(0, \sigma_\epsilon^l)$ and $0 < \tau_{1j} < 1$. We expect that the marginal utility of home time will go up when a newborn arrives, especially for women (i.e. $\tau_{2f} > 0$). Now consider the utility from pregnancy:

$$\pi_t = \pi_1 H_{ft} + \pi_2 N_t + \pi_3 p_{t-1} + \epsilon_t^p$$

where $\epsilon_t^p \sim N(0, \sigma_\epsilon^p)$ is a stochastic shock to tastes for pregnancy. The value of π_t is a function of women's health, the number of children and lagged pregnancy. Finally, consider the function Q that gives the utility a couple get from the quality and quantity of children:

$$Q(l_t^f, l_t^m, Y_t^M, N_t) = (a_f(l_t^f)^\rho + a_m(l_t^m)^\rho + a_g(\theta(N_t)Y_t^M)^\rho + (1 - a_f - a_m - a_g)N_t^\rho)^{\frac{1}{\rho}}$$

The first three inputs increase child quality, while A_j^M is a scale parameter.

We can now write the choice-specific value functions for married individuals:

$$V_t^{jM}(l_t^m, l_t^f, p_t|\Omega_{mt}, \Omega_{ft}) = \frac{1}{\alpha}(\psi C_t^M)^\alpha + L(l_t^j) + \pi_t p_t + A_j^M Q(l_t^f, l_t^m, Y_t^M, N_t) + \delta E(V_{t+1}^{jM}(\Omega_{m,t+1}, \Omega_{f,t+1}))$$

Next period state depends on the current state $(\Omega_{mt}, \Omega_{ft})$ and current choices (l_t^m, l_t^f, p_t) through the law of motion of the state variables. The parameter δ is the discount rate.

Household decision making

The partners choose leisure and fertility to maximize:

$$V_t^M(l_t^m, l_t^f, p_t|\Omega_{mt}, \Omega_{ft}) = \lambda V_t^{fM}(l_t^m, l_t^f, p_t|\Omega_{mt}, \Omega_{ft}) + (1-\lambda)V_t^{mM}(l_t^m, l_t^f, p_t|\Omega_{mt}, \Omega_{ft})$$

Here λ and $1 - \lambda$ are Pareto weights. Couples seek a choice vector $\{l_t^m, l_t^f, p_t\}$ that maximize V_t^M. The vector of household choices that maximize V_t^M is:

$$\{l_t^{m*}, l_t^{f*}, p_t^*\} = argmax_{\{l_t^m, l_t^f, p_t\}} V_t^M(l_t^m, l_t^f, p_t|\Omega_{mt}, \Omega_{ft})$$

We define the maximized value function of a married individual in state Ω_{jt}:

$$V_t^{jM}(\Omega_{mt}, \Omega_{ft}) = V_t^{jM}(l_t^{m*}, l_t^{f*}, p_t^*|\Omega_{mt}, \Omega_{ft})$$

The labor market

The wage equations have the standard form:

$$ln(w_t^j) = \omega_{0j} + \omega_{1j}E_t + \omega_{2j}X_t - \omega_{3j}X_t^2 + \epsilon_{jt}^w$$

where E_t is education, X_t is work experience and $\epsilon_{jt}^w \sim$ N(0,σ_ϵ^w). An unemployed receives at most one job offer per period (Full-Time or Part-Time). In each period an individual have three possible choice sets for hours: D = {0},{0,0.5},{0,1}. The probability that each of the three choice sets is offered is determined by a trinomial logit:

$$P_j(D_{kt}) = \begin{cases} \exp(\rho_{jk0} + \rho_{jk1}E_t + \rho_{jk2}X_t + \rho_{jk3}H_t)/IV_{jt} & \text{if} \quad k = 2,3 \\ 1/IV_{jt} \quad \text{if} \quad k = 1 \end{cases}$$

where $IV_{jt} \equiv 1 + \sum_{k=2}^{3} \exp(\rho_{jk0} + \rho_{jk1}E_t + \rho_{jk2}X_t + \rho_{jk3}H_t)$. The probability of each choice set depends on education, work experience and health. The probability of a job loss is a logit function of the same three variables. In each period a person may be unemployed because he/she draws the empty set D = {0} or because has a part-time or full-time offer and rejects it.

Health status

Health status is modeled as a three-state Markov-chain, where $H_{jt} \in \{1, 2, 3\}$ indicates poor, fair and good health. The transition probabilities differ by age, and, as the process is exogenous, the parameters of the health transition matrix are estimated outside the model.

The health transition probability is a multinomial logit:

$$P(H_{jt} = k) = \begin{cases} \exp(\sum_{q=1}^{3} \chi_{jkq} I[H_{j,t-1} = q])/IV_{jt}^H & \text{if} \quad k = 2,3 \\ 1/IV_{jt}^H \quad \text{if} \quad k = 1 \end{cases}$$

where $IV_{jt}^H \equiv 1 + \sum_{k=2}^{3} \exp(\sum_{q=1}^{3} \chi_{jkq} I[H_{j,t-1} = q])$. Health status affects tastes for leisure and the job offer probability. The assumption that health evolves exogenously means that it generates exogenous variation in these decisions.

3.4 Solution and estimation

We solve the model backwards from a final age (we assume 57) to the age when the couple gets married and starts making joint decisions about labor supply and fertility (28).

We list here all the state variables with the number of possible values for each. The variables for a married individual are gender (j=m,f); age (t=28,...,57); education (3 levels, less than high school; high school and university); experience with 5 levels (0-1,2-4,5-9,10-16,17+); children with 4 levels (0,1,2,3+); health with 3 levels (poor, fair and good); taste for leisure with 3 levels; lagged pregnancy (2 levels) plus education, experience, health and taste for leisure of the spouse (with the same number of levels). For example, an individual aged 35 of gender j has 145800 points in the state space.

We back-solve the model from the final age 57 to the starting age 28. In order to reduce the computational burden we exclude childbirth after age 40, that is quite rare in the data. We estimate the model using annual data from the Survey of Household Income and Wealth (SHIW) for the period 1984 to 2014. The sample is composed of married people aged 28 to 57. We focus our attention on the 1960 cohort, i.e. people born between 1957 and 1963, since in the SHIW we have data covering all the ages in the model for this cohort. For the estimation of the model's parameters we use the Method of Simulated Moments, as in McFadden (1989) and Pakes and Pollard (1989), implemented by minimizing the distance between actual data and simulated data from our model [4]. The moments used to fit the model are listed in Table 1, while the list of the parameters estimated is presented in Table ??. After the backward solution of the model, we simulate forward to generate life-cycle choices from the initial age 28 until the terminal period (age 57). We need to draw initial conditions for each person's education and taste for leisure. We assume for now that initial experience equals zero. Then we must draw, for each

[4] For a recent study on how to use the Method of simulated moments for the estimation of dynamic discrete choice models see Eisenhauer et al. (2015)

individual i in each period t, a job offer, a health realization and the shocks.
Conditional on these realizations the model generates simulated choices and outcomes for the endogenous variables: employment, children and wages. We simulate data for 1000 married men and women. We construct a set of statistics for both the simulated and the actual data summarizing key predictions of the model. The statistics include employment rates and wages for men and women. We estimate 57 parameters using 145 moments. Let d_j denote a statistic from the actual data, and let $d_j^s(\theta)$ be the corresponding statistic calculated from the simulated data. We construct moments of the form:

$$m_j^s(\theta) = [d_j - d_j^s(\theta)] \quad for \quad j = 1, ..., J$$

The vector of simulated moments is given by $g'(\theta) = [m_1^s(\theta)...m_J^s(\theta)]$. We minimize the objective function $G(\theta) = g'(\theta)Wg(\theta)$ with respect to θ, where the weighting matrix W is a diagonal matrix consisting of the inverse of the estimated variance of each moment. The variance of the estimator is:

$$\hat{V} = (1 + \frac{1}{NS})(\hat{G}'W\hat{G})^{-1}$$

where NS is the number of households times the number of simulations. $\hat{G}$ is the matrix of the first derivative of every moment with respect to every parameter. The derivatives are approximated numerically.

3.5 Data

Our analysis is mainly based on data from the SHIW, a survey administered by the Bank of Italy (usually) every two years to a sample of approximately 8000 households (20000 individuals) per year. The survey collects detailed information on education, labor market, consumption and income-related issues. Historical data start from year 1977, however precise values for the age of individuals are available just from 1984. For this reason we will focus our attention on waves from 1984 to 2014. In order to be coherent

Table 3.1: List of moments used in the analysis

Moment	Num. of mom. 1960 cohort
men Full Time	10
women Full Time	10
men Part Time	10
women Part Time	10
women with children employment	10
women no children employment	10
women children by age	5
men wage	10
women wage	10
wage by education - women	3x10
employment by education - women	3x10

with the specified model we restrict our sample to married or cohabiting individuals aged 28 to 57. We then use this sample to compute the empirical moments defined in the previous section for the 1960 cohort.

In Table 3.2 we present summary statistics for the main variables used in the analysis. The employment rate is 12.4 percentage point higher for married women without children than for wives with one or more children. The average number of children per married couple is around 1.6. While more than one out of five women work part-time, almost all the men in the sample have a full-time job. Married men tend to earn much more than married women, with a gender gap above 20%. Finally, wages and employment of wives increase dramatically with the level of education. The average wage for a woman with a university degree is more than 50% higher than the wage for an high-school dropout. The same happens for the probability of being employed: while just 35% of women with a low level of education work, the share of university graduates working is close to 83%. In order to compute transition matrices for the health function we use health status information from EU-SILC.

In Table 3.3 we define the values of some parameters that we keep constant in the estimation. The discount factor δ is set at 0.95, very close to values estimated in the previous literature. The Pareto weights λ, reflecting the bargaining power of each of the two spouses, are fixed at 0.5. Some of the recent literature (Browning et al., 2013) found higher weights for women in a marriage, however the evidence is based on Canadian data, for Italy we think that a 50-50 split might be more appropriate. The value for economies of scale in household consumption, ψ, comes from Eckstein et al. (2016).

3.6 Results

In this section we present the main estimation results. The estimates for the model parameters are shown in Tables 3.4-3.9. Table 3.10 presents some measures of model fit and Table 3.11 shows the results of the policy experiment.

3.6.1 Parameter estimates

The estimates for the wage equations are shown in Table 3.6. The return from completing an additional level of education is 0.41 for men and 0.38 for women. This is approximately equivalent to a return from an additional year of education of 8%, a figure in line with the findings of the literature (Belzil, 2007). Experience, as expected, has a positive effect on wages for both men and women. One additional year on the job increases earnings by about 5%, however the effect is decreasing over time. This value is slightly larger than other similar estimates in the recent literature (**?**).

In Table 3.7 we present the estimated parameters of the home time (leisure) equations. We find that the tastes for leisure μ_j are more persistent for women than for men. Furthermore, having a pregnancy in $t-1$ have a much stronger effect on tastes for leisure of wives than husbands. This last finding is not surprising since in Italy many women, especially the low skilled, decide to exit the labor market at childbirth. On the other hand the labor supply of men tend to be independent of the birth of children.

Table 3.8 shows the parameters estimated from the utility from pregnancy and utility from children equations. The utility from a pregnancy in t is negatively affected by the number of children in the household and by having had a pregnancy in $t-1$.

3.6.2 Model fit

In Table 3.10 we present the model fit for employment, wages and fertility in different age groups. The model is able to predict wages quite well for both men and women. However, simulated wages are higher that real wages for older male workers and lower for older employed women. This fact is also evident in Figure 5, where we show the dynamics of wages for man and women. The gap for women in the last years is probably due to the fact that in Italy low productivity women exit the labor market in their early fifties, leaving just the higher ability women on the job. The model is not able to fully account for this positive self-selection. Employment is underestimated for both men and women at younger ages. For elderly workers the model underestimates employment rates for husbands but overestimates them for wives. These moments confirm that the model fails to explain the large rate of early retirement among women. The low estimates for employment rates at young ages is partly due to the fact that we set the years of experience to zero in period 1 as an initial condition. This assumption might be reasonable for university graduates but it is too restrictive for workers with a lower level of education. Finally, due to the fact that we assume women get married when they reach age 28, childbirth happens at older ages in the model than in the data. This explains the slightly lower employment rate for women aged 37-45 with respect to younger women in the model and the fact that fertility is higher than in real data for women between 34 and 40.

3.6.3 Policy experiments

In Table 3.11 we show the results of the experiment. The goal is to study the effect of the permanent implementation of a childbirth transfer on household behavior. We simulated the effect of a subsidy worth 2880 euros given in the first year after a child is born. The evaluation is performed for households aged 28, 33 and 38 when the policy starts. As expected, the effect is largest when the transfer is implemented at the beginning of the household's fertility period (Column 1). We find a 4.8% increase in total fertility in this group, as well as a decrease in both Full Time and Part Time employment of 0.21 years and 0.04 years respectively. In Column 2 we present the results of the experiment for the households aged 33 when the policy starts. The increase in total fertility is still quite large, at 3.7%, and the decrease in employment is lower than before. We observe a reduction of 0.13 years in full time employment and 0.01 years in part time employment. Finally, if the subsidy is implemented when the household's members are 38 years old, i.e. almost at the end of their fertility period, the effect on total fertility is close to zero, and the same is true for the effect on the time spent in employment.

3.7 Concluding remarks

In this work we built and estimated a dynamic life-cycle model with endogenous wages, employment and fertility. Decisions are taken jointly by husbands and wives at the household level. We are able to account for several sources of heterogeneity across individuals and families. The model is estimated using a repeated cross-section of Italian households and the Method of Simulated Moments.

The model can fit the observed behavior of agents quite well. We then use the estimated model to simulate the outcomes of a policy: a childbirth transfer very similar to the one implemented by the Italian government starting from 2015. The results show that the subsidy, if permanent, is successful in increasing fertility. However, especially for young couples, the increase in total fertility comes at the expense of a decrease in the time spent on the labor market for women.

Table 3.2: Summary statistics

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Share wives with children employed	0.47	0.499	0	1
Share wives no children employed	0.594	0.491	0	1
Number of children per couple	1.595	1.006	0	9
Share of employed wives working FT	0.792	0.406	0	1
Share of employed wives working PT	0.208	0.406	0	1
Share of employed husbands working FT	0.973	0.162	0	1
Share of employed husbands working PT	0.027	0.162	0	1
Married women wage	11790	6180	90	95000
Married men wage	15083	7893	100	120000
Married women wage - low education	9669	4752	150	70000
Married women wage - avg education	12275	5317	90	60000
Married women wage - high education	15155	8478	300	95000
Married women employment - low educ.	0.35	0.477	0	1
Married women employment - avg educ.	0.619	0.486	0	1
Married women employment - high educ.	0.828	0.378	0	1

Table 3.3: Fixed parameters (from the literature and Italian data)

	Fixed Value Husband	Fixed value Wife
Discount factor	0.95	
Pareto weight (λ)	0.5	
Economies of scale in HH cons. (ψ)	0.707	
Scale parameter (A_j)	0.25	0.25
Unemployment benefit + housework (b_j)	2000	3000

WE SHOULD FIND THE VALUES FOR b_j FROM DATA ON EMPLOYMENT BENEFITS AND HOURS SPEND WORKING AT HOME. I MADE IT UP NOW.

Table 3.4: Job offers parameters: husbands

	Estimated Value	Standard Error
Job offer parameter - Full Time		
constant	1.5943	0.649
experience	0.2148	0.097
education	0.1518	0.110
health	-0.4147	0.099
Job offer parameter - Part Time		
constant	-0.2857	0.350
experience	0.0806	0.105
education	0.0202	0.119
health	-0.4312	0.256
Job offer parameter - No job		
constant	-0.2860	3.185
experience	0.1621	0.195
education	0.1501	0.361
health	-0.4166	0.345

Table 3.5: Job offers parameters: wives

	Estimated Value	Standard Error
Job offer parameter - Full Time		
constant	1.4325	0.741
experience	0.2009	0.101
education	0.1502	0.125
health	-0.4070	0.135
Job offer parameter - Part Time		
constant	0.1020	0.337
experience	0.1215	0.095
education	0.0809	0.062
health	-0.4076	0.276
Job offer parameter - No job		
constant	-0.5061	0.493
experience	0.1403	0.246
education	0.1498	0.268
health	-0.4218	0.171

Table 3.6: Estimated parameters: wage equations

	Estimated Value	Standard Error
Husband's wage equation		
Constant	8.4011	0.917
Level of education	0.3743	0.085
Experience	0.0495	0.073
Experience^2	-0.001	0.003
Wage error variance	0.4721	1.067
Wife wage equation		
Constant	8.2154	0.304
Level of education	0.3661	0.113
Experience	0.0494	0.051
Experience^2	-0.001	0.003
Wage error variance	0.4702	0.926

Table 3.7: Estimated parameters: home time (leisure) equations

	Estimated Value	Standard Error
Husband home time equation		
Constant	0.0010	0.026
AR coefficient	0.6537	0.152
Pregnancy in $t-1$	0.1475	0.335
Home time shock variance	0.2468	1.447
Wife home time equation		
Constant	0.0010	0.008
AR coefficient	0.8442	0.111
Pregnancy in $t-1$	0.7344	0.326
Home time shock variance	0.2469	0.773

Table 3.8: Estimated parameters: fertility equations

	Estimated Value	Standard Error
Utility from pregnancy		
Health	-0.1460	0.079
Number of kids in the HH	-0.9093	0.828
Pregnancy in $t-1$	-3.1594	4.309
Pregnancy shock variance	0.8936	0.826
Utility from quality-quantity of children		
CES function's parameter (ρ)	-0.8508	0.578
wife leisure (a_f)	0.5623	0.282
husband leisure (a_m)	0.3780	0.285
spending per child (a_g)	0.0004	0.001

Table 3.9: Estimated parameters: utility function

	Estimated Value	Standard Error
CRRA consumption parameter (α)	0.1449	0.044
CRRA leisure parameter (γ)	0.9151	0.961
leisure when pregnant (wives only)	0.0507	0.255
leisure by education (wives)	0.0399	0.176
leisure by education (husbands)	0.0204	0.141
leisure by health (wives)	0.0506	0.179
leisure by health (husbands)	0.0503	0.300

Table 3.10: Model fit

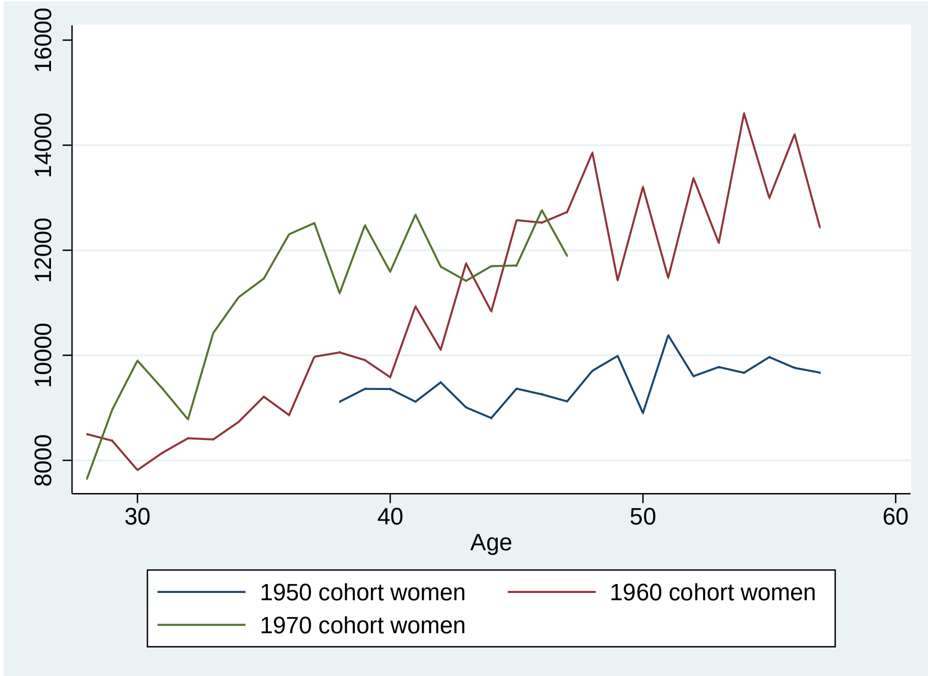
	Simulated Data	Real Data
employment 28-36 males	0.710	0.936
employment 28-36 females	0.483	0.495
employment 37-45 males	0.761	0.949
employment 37-45 females	0.503	0.512
employment 46-57 males	0.782	0.874
employment 46-57 females	0.567	0.519
wages 28-35 males	12634	12780
wages 28-35 females	10556	10163
wages 36-45 males	15759	15700
wages 36-45 females	12041	12178
wages 46-57 males	18306	17039
wages 46-57 females	13695	14327
Number of children 29-33	1.18	1.44
Number of children 34+	1.90	1.65

Table 3.11: Experimental results: effects of the fertility subsidy

	Age at start of policy		
	28	33	38
Change in total fertility	4.8%	3.7%	0.7%
Change in female years working FT	-0.208	-0.132	-0.040
Change in female years working PT	-0.041	-0.011	-0.003

Figure 3.1: Employment rate by cohort

(a) Employment rate women



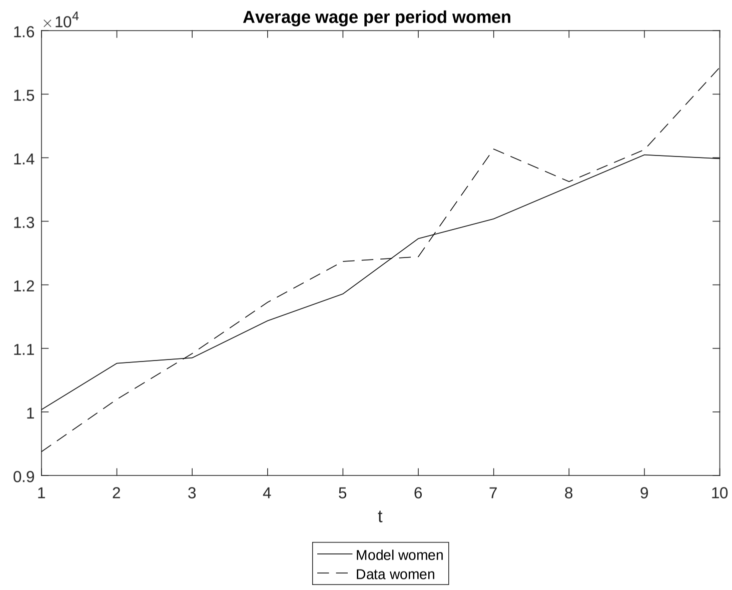
(b) Real wage women

Figure 3.2: Real wage by cohort

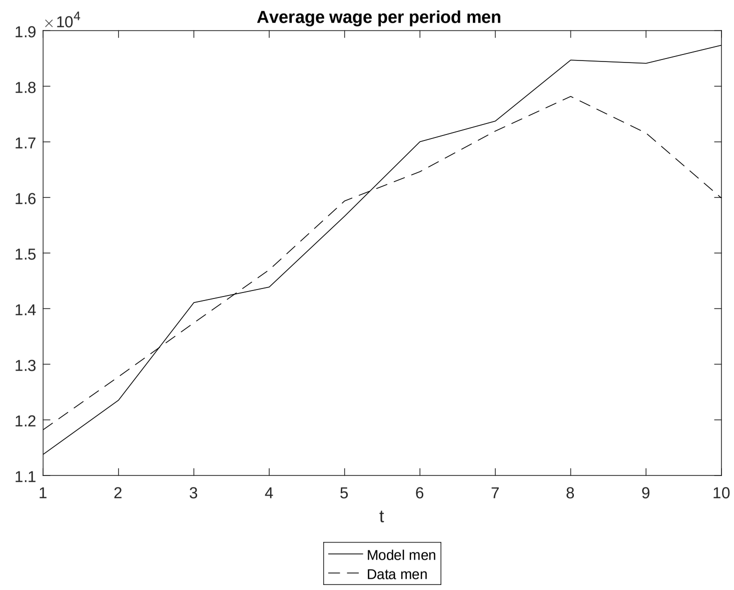
(a) Employment rate men

(b) Real wage men

Figure 3.3: Model fit: wages



(a) Model fit: women wages



(b) Model fit: men wages

Bibliography

Adda, J., Dustmann, C., Stevens, K., 2015. The career costs of children. Journal of Political Economy (forthcoming).

Altonji, J. G., Pierret, C. R., 1997. Employer learning and statistical discrimination. Tech. rep., National Bureau of Economic Research.

ANVUR, 2013. Valutazione della qualit della ricerca 2004-2010, rapporto finale.

Arulampalam, W., Booth, A. L., Bryan, M. L., 2007. Is there a glass ceiling over europe? exploring the gender pay gap across the wage distribution. Industrial & Labor Relations Review 60 (2), 163–186.

Attanasio, O., Low, H., Sánchez-Marcos, V., 2008. Explaining changes in female labor supply in a life-cycle model. The American Economic Review 98 (4), 1517–1552.

Auranen, O., Nieminen, M., 2010. University research funding and publication performance–An international comparison. Research Policy 39 (6), 822–834.

Babcock, P., 2010. Real costs of nominal grade inflation? new evidence from student course evaluations. Economic Inquiry 48 (4), 983–996.

Bagues, M., Labini, M. S., Zinovyeva, N., 2008. Differential grading standards and university funding: Evidence from italy. CESifo Economic Studies 54 (2), 149–176.

Bailey, M. A., Rosenthal, J. S., Yoon, A. H., 2014. Grades and incentives: assessing competing grade point average measures and postgraduate outcomes. Studies in Higher Education, 1–15.

Belzil, C., 2007. The return to schooling in structural dynamic models: a survey. European Economic Review 51 (5), 1059–1105.

Betts, J. R., Grogger, J., 2003. The impact of grading standards on student achievement, educational attainment, and entry-level earnings. Economics of Education Review 22 (4), 343–352.

Bonaccorsi, A., Cicero, T., 2015. Distributed or concentrated research excellence? evidence from a large-scale research assessment exercise. Journal of the Association for Information Science and Technology.

Bowman, N. A., Bastedo, M. N., 2009. Getting on the front page: Organizational reputation, status signals, and the impact of us news and world report on student decisions. Research in Higher Education 50 (5), 415–436.

Braga, M., Paccagnella, M., Pellizzari, M., 2014. Evaluating students evaluations of professors. Economics of Education Review 41, 71–88.

Bratti, M., Checchi, D., De Blasio, G., 2008. Does the expansion of higher education increase the equality of educational opportunities? evidence from italy. Labour 22 (s1), 53–88.

Bratti, M., Checchi, D., Salini, S., Verzillo, S., Serlenga, L., Viesti, G., dUggento, A., 2015. Il sistema universitario italiano: unanalisi regionale.

Broecke, S., 2015. University rankings: do they matter in the uk? Education Economics 23 (2), 137–161.

Browning, M., Chiappori, P.-A., Lewbel, A., 2013. Estimating consumption economies of scale, adult equivalence scales, and household bargaining power. The Review of Economic Studies, rdt019.

Brunello, G., Cappellari, L., 2008. The labour market effects of alma mater: Evidence from italy. Economics of Education Review 27 (5), 564–574.

Butcher, K. F., McEwan, P. J., Weerapana, A., 2014. The effects of an anti-grade-inflation policy at wellesley college. The Journal of Economic Perspectives 28 (3), 189–204.

Cappellari, L., Lucifora, C., 2009. The "Bologna Process" and college enrollment decisions. Labour Economics 16 (6), 638–647.

Cattaneo, M., Meoli, M., Signori, A., 2016. Performance-based funding and university research productivity: the moderating effect of university legitimacy. The Journal of Technology Transfer 41 (1), 85–104.

Checchi, D., Ichino, A., Rustichini, A., 1999. More equal but less mobile?: Education financing and intergenerational mobility in italy and in the us. Journal of public economics 74 (3), 351–393.

Chevalier, A., Jia, X., 2015. Subject-specific league tables and students' application decisions. The Manchester School.

De Angelis, I., Mariani, V., Modena, F., Montanaro, P., et al., 2016. Academic enrolment, careers and student mobility in italy. Tech. rep., Bank of Italy, Economic Research and International Relations Area.

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

Del Boca, D., Sauer, R. M., 2009. Life cycle employment and fertility across institutional environments. European Economic Review 53 (3), 274–292.

Di Pietro, G., Cutillo, A., 2008. Degree flexibility and university drop-out: The Italian experience. Economics of Education Review 27 (5), 546–555.

Dotti, N. F., Fratesi, U., Lenzi, C., Percoco, M., March 2013. Local Labour Markets and the Interregional Mobility of Italian University Students. Spatial Economic Analysis 8 (4), 443–468.

Eckstein, Z., Keane, M., Lifshitz, O., 2016. Sources of change in the life-cycle decisions of american men and women: 1962-2014.

Eckstein, Z., Lifshitz, O., 2011. Dynamic female labor supply. Econometrica 79 (6), 1675–1726.

Eckstein, Z., Lifshitz, O., 2015. Household interaction and the labor supply of married women. International Economic Review 56 (2), 427–455.

Eckstein, Z., Wolpin, K. I., 1989. Dynamic labour force participation of married women and endogenous work experience. The Review of Economic Studies 56 (3), 375–390.

Eisenhauer, P., Heckman, J. J., Mosso, S., 2015. Estimation of dynamic discrete choice models by maximum likelihood and the simulated method of moments. International economic review 56 (2), 331–357.

Figlio, D. N., Lucas, M. E., 2004. Do high grading standards affect student performance? Journal of Public Economics 88 (9), 1815–1834.

Fondazione RES, 2016. Universit in declino. Un'indagine sugli atenei da Nord a Sud. Donzelli Editore, Rome.

Franceschet, M., Costantini, A., 2011. The first italian research assessment exercise: A bibliometric perspective. Journal of informetrics 5 (2), 275–291.

Francesconi, M., 2002. A joint dynamic model of fertility and work of married women. Journal of Labor Economics 20 (2), 336–380.

Geuna, A., Piolatto, M., 2016. Research assessment in the UK and Italy: Costly and difficult, but probably worth it (at least for a while). Research Policy 45 (1), 260–271.

Gibbons, S., Neumayer, E., Perkins, R., 2015. Student satisfaction, league tables and university applications: evidence from britain. Economics of Education Review 48, 148–164.

Goldin, C., 2013. Can yellen effectattract young women to economics? BloombergView, October 14.

Griffith, A., Rask, K., 2007. The influence of the us news and world report collegiate rankings on the matriculation decision of high-ability students: 1995–2004. Economics of Education Review 26 (2), 244–255.

Horstschräer, J., 2012. University rankings in action? the importance of rankings and an excellence competition for university choice of high-ability students. Economics of Education Review 31 (6), 1162–1176.

Jiménez-Contreras, E., de Moya Anegón, F., López-Cózar, E. D., 2003. The evolution of research activity in Spain: The impact of the National Commission for the Evaluation of Research Activity (CNEAI). Research Policy 32 (1), 123–142.

Keane, M. P., Wolpin, K. I., 1997. The career decisions of young men. Journal of political Economy 105 (3), 473–522.

Keane, M. P., Wolpin, K. I., 2010. The role of labor and marriage markets, preference heterogeneity, and the welfare system in the life cycle decisions of black, hispanic, and white women. International Economic Review 51 (3), 851–892.

Kohler, H.-P., Billari, F. C., Ortega, J. A., 2002. The emergence of lowest-low fertility in europe during the 1990s. Population and development review 28 (4), 641–680.

Luca, M., Smith, J., 2013. Salience in quality disclosure: evidence from the us news college rankings. Journal of Economics & Management Strategy 22 (1), 58–77.

Marino, M. R., Romanelli, M., Tasso, M., et al., 2016. Women at work: the impact of welfare and fiscal policies in a dynamic labor supply model. Tech. rep., Bank of Italy, Economic Research and International Relations Area.

Meredith, M., 2004. Why do universities compete in the ratings game? an empirical analysis of the effects of the us news and world report college rankings. Research in Higher Education 45 (5), 443–461.

Monks, J., Ehrenberg, R. G., 1999. The impact of us news & world report college rankings on admissions outcomes and pricing policies at selective private institutions (cornell higher education research institute, 1999).

Montanaro, P., 2008. I divari territoriali nella preparazione degli studenti italiani: evidenze dalle indagini nazionali e internazionali. Occasional papers (14).

Oppedisano, V., 2011. The (adverse) effects of expanding higher education: Evidence from italy. Economics of Education Review 30 (5), 997–1008.

Pigini, C., Staffolani, S., 2015. Beyond participation: do the cost and quality of higher education shape the enrollment composition? the case of italy. Higher Education First online: 05 May 2015, 1–24.

Rebora, G., Turri, M., 2013. The uk and italian research assessment exercises face to face. Research policy 42 (9), 1657–1666.

Sabot, R., Wakeman-Linn, J., 1991. Grade inflation and course choice. The Journal of Economic Perspectives 5 (1), 159–170.

Sauder, M., Lancaster, R., 2006. Do rankings matter? the effects of us news & world report rankings on the admissions process of law schools. Law & Society Review 40 (1), 105–134.

Soo, K. T., 2013. Does anyone use information from university rankings? Education Economics 21 (2), 176–190.

Tutterow, C., Evans, J. A., 2016. Reconciling the small effect of rankings on university performance with the transformational cost of conformity. The University Under Pressure (Research in the Sociology of Organizations, Volume 46) Emerald Group Publishing Limited 46, 265–301.

Van der Klaauw, W., 1996. Female labour supply and marital status decisions: A life-cycle model. The Review of Economic Studies 63 (2), 199–235.