

---

# Who benefits from online education? Exploring student outcomes in Italian higher education

---

---

Received: 26 November 2025

Accepted: 5 May 2026

Published online: 22 June 2026

Cite this article as: Usala C., Lombardi G. & Priulla A. Who benefits from online education? Exploring student outcomes in Italian higher education. *Genus* (2026). <https://doi.org/10.1186/s41118-026-00292-5>

Cristian Usala, Gabriele Lombardi & Andrea Priulla

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

## Who benefits from online education? exploring student outcomes in italian higher education

This Accepted Manuscript (AM) is a PDF file of the manuscript accepted for publication after peer review, when applicable, but does not reflect post-acceptance improvements, or any corrections. Use of this AM is subject to the publisher's embargo period and AM terms of use. Under no circumstances may this AM be shared or distributed under a Creative Commons or other form of open access license, nor may it be reformatted or enhanced, whether by the Author or third parties. By using this AM (for example, by accessing or downloading) you agree to abide by Springer Nature's terms of use for AM versions of subscription articles: <https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms>

The Version of Record (VOR) of this article, as published and maintained by the publisher, is available online at: <https://doi.org/10.1186/s41118-026-00292-5>. The VOR is the version of the article after copy-editing and typesetting, and connected to open research data, open protocols, and open code where available. Any supplementary information can be found on the journal website, connected to the VOR.

For research integrity purposes it is best practice to cite the published Version of Record (VOR), where available (for example, see ICMJE's guidelines on overlapping publications). Where users do not have access to the VOR, any citation must clearly indicate that the reference is to an Accepted Manuscript (AM) version.

Accepted Manuscript

# Who Benefits from Online Education? Exploring Student Outcomes in Italian Higher Education

Cristian Usala<sup>1,2</sup>, Gabriele Lombardi<sup>3</sup>, Andrea Priulla<sup>4</sup>

## Declarations:

### Availability of data and materials:

The data used in this study have been processed under the research protocol for the study “*From high school to the job placement: analysis of university careers and university mobility from Southern to Northern Italy*” among the Ministry of University and Research, the Ministry of Education and Merit, the University of Palermo as the lead institution, and the INVALSI Institute. For the University of Cagliari, the reference researcher is Mariano Porcu. Due to institutional agreements and privacy regulations, the data are not publicly available outside of this research framework.

### Competing interests:

The authors declare that they have no competing interests

### Acknowledgements:

We acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2.2.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union - NextGenerationEU - Project Title “*From high school to university: Assessing peers' influence in educational inequalities and performance*” - CUP F53D23006150006 - Grant Assignment Decree No. 1060 adopted on 07/17/2023 by the Italian Ministry of University and Research (MUR).

### Authors' contributions:

All authors contributed to the study's conception and design. C.U. performed material preparation, data collection, and defined the empirical framework. G.L. worked on the theoretical background, descriptive statistics, and results discussion. A.P. contributed to the definition of the empirical framework, performed the empirical analysis, and discussed the results. All authors have contributed to the conclusions. All authors read and approved the final manuscript.

---

<sup>1</sup>Corresponding author, e-mail: cristian.usala@unica.it

<sup>2</sup>Department of Social and Political Sciences, University of Cagliari, Cagliari, Italy

<sup>3</sup> Department of Statistics, Informatics, and Applications “G. Parenti”, University of Florence, Florence, Italy

<sup>4</sup> Department of Economics and Law, University of Enna - KORE, Enna, Italy

# Who Benefits from Online Education? Exploring Student Outcomes in Italian Higher Education

Cristian Usala<sup>1,2</sup>, Gabriele Lombardi<sup>3</sup>, Andrea Priulla<sup>4</sup>

## Declarations:

### Availability of data and materials:

The data used in this study have been processed under the research protocol for the study “*From high school to the job placement: analysis of university careers and university mobility from Southern to Northern Italy*” among the Ministry of University and Research, the Ministry of Education and Merit, the University of Palermo as the lead institution, and the INVALSI Institute. For the University of Cagliari, the reference researcher is Mariano Porcu. Due to institutional agreements and privacy regulations, the data are not publicly available outside of this research framework.

### Competing interests:

The authors declare that they have no competing interests

### Acknowledgements:

We acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2.2.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union - NextGenerationEU - Project Title “*From high school to university: Assessing peers' influence in educational inequalities and performance*” - CUP F53D23006150006 - Grant Assignment Decree No. 1060 adopted on 07/17/2023 by the Italian Ministry of University and Research (MUR).

### Authors' contributions:

All authors contributed to the study's conception and design. C.U. performed material preparation, data collection, and defined the empirical framework. G.L. worked on the theoretical background, descriptive statistics, and results discussion. A.P. contributed to the definition of the empirical framework, performed the empirical analysis, and discussed the results. All authors have contributed to the conclusions. All authors read and approved the final manuscript.

---

<sup>1</sup>Corresponding author, e-mail: cristian.usala@unica.it

<sup>2</sup>Department of Social and Political Sciences, University of Cagliari, Cagliari, Italy

<sup>3</sup> Department of Statistics, Informatics, and Applications “G. Parenti”, University of Florence, Florence, Italy

<sup>4</sup> Department of Economics and Law, University of Enna - KORE, Enna, Italy

# Who Benefits from Online Education? Exploring Student Outcomes in Italian Higher Education

## Abstract

In the last decade, the substantial expansion of online education in Italy has been accompanied by a growing public and institutional debate over its academic quality and effectiveness. This study contributes to this debate by investigating the differences in academic outcomes between students enrolled in online institutions and those in traditional universities, with a particular focus on first-year dropout. Using administrative data from the MOBYSU.IT database on the population of students enrolled in an Italian university between 2010 and 2019, we apply Propensity Score Matching to estimate the causal impact of attending an online institution. Our analysis shows that students in online universities are, on average, less likely to drop out after the first year. This retention effect is especially evident among students who enrol later in life, those from technical and vocational schools, and those from Southern Italy. Conversely, students entering university directly after high school are more likely to transfer to traditional institutions. These findings suggest that online universities can play a complementary role in reducing dropout and promoting access to higher education, particularly for non-traditional and underserved student populations.

**Keywords:** higher education, online universities, academic outcomes, dropout rates

## 1 Introduction

In 2000, a seminal paper by [Harasim \(2000\)](#) predicted that the contemporary revolution in computer communications would contribute to increasing educational opportunities through the development of new learning models. By 2006, the follow-up to this work appeared to consider the debate settled, arguing that the rise of online education had replaced the passive, competitive, closed, and individualised traditional higher education system with an active, collaborative, open, and networked online paradigm, destined to become the new mainstream ([Harasim 2006](#)).

Indeed, although online education has not overtaken traditional on-campus programmes even after 20 years, its market share has consistently grown worldwide, reaching approximately 25–30% of the total, depending on different estimates and regions of the world ([Kumar et al. 2017](#); [Cheslock and Jaquette 2022](#); [Palvia et al. 2018](#); [EUROSTAT 2024](#)). In Italy, over the last decade, the educational offerings of online universities have increased by 113%, and the number of enrolled students has grown by more than 400%. As a result, online students now represent over 10% of the total number of university students in Italy ([Marra 2024](#)). Nonetheless, the noticeable expansion of online education in Italy has stimulated an intense debate about the efficacy and efficiency of these institutions. In recent days, Confindustria – the main association representing manufacturing and service enterprises

in Italy – and the CGIL – the oldest and most representative Italian trade union – have accused the Italian online university system of resembling a “degree mill”, calling on the Ministry of University and Research to introduce stricter regulations and limitations.<sup>1</sup> These controversies have found fertile ground following the investigations concerning one of the Italian online universities, which pointed to the possibility that specific categories of students obtained their degrees without sitting any exams.<sup>2</sup>

In this context, the existing gap in the academic literature on the Italian online university system and its impact on students’ choices and outcomes is particularly concerning, given the growing relevance of this sector in the Italian higher education system. Therefore, as online education continues to expand and foster debate among institutions and the broader public, assessing its role has become crucial.

Our article represents an initial attempt to fill this gap in the literature by analysing the differences in academic outcomes between students enrolled in traditional and online universities, focusing on their transition from the first to the second year. To do this, we exploit the information on the population of students enrolled in an Italian university between 2010 and 2019, included in MOBYSU.IT database. This data allows us to distinguish between students who remain enrolled in the same type of institution and those who switch from online to traditional (or vice versa) or drop out of the Italian higher education system. In particular, student dropout is the main outcome explored in our work, given its crucial importance, particularly in Italy, where the share of university graduates and students is significantly lower than the European average (EUROSTAT 2025a,b).

In addition, the database distinguishes between students enrolled in traditional and online universities, precisely identifying not only their institution but also the chosen degree programme. These two elements, along with a detailed set of information on students’ characteristics, enable us to adopt a Propensity Score Matching (PSM) approach to estimate the causal effect of attending an online university on students’ academic outcomes. In detail, PSM is used to remove the selection bias associated with the choice of attending an online university that may affect the results of a simple comparison between the two groups of students. Moreover, the longitudinal nature of the MOBYSU.IT data also allows us to differentiate between direct enrolments (students who enrol immediately after high school) and non-direct enrolments (students who enrol at a later stage). Although we cannot verify with certainty that the latter are working students, this provides a valuable proxy to test the widespread assumption that online universities primarily serve those who enter higher education later in life, who likely lack the opportunity to experience full-time, in-person university life, and may benefit more from the flexibility provided by these institutions.

---

<sup>1</sup>Forte, E. (2025, July 21). Il caso delle Università telematiche: boom di iscritti e accuse di “lauree facili”. *La Stampa*. [https://www.lastampa.it/cronaca/2025/07/21/news/il\\_suk\\_delle\\_universita\\_telematiche\\_boom\\_di\\_iscritti\\_e\\_accuse\\_di\\_lauree\\_facili-15240096/](https://www.lastampa.it/cronaca/2025/07/21/news/il_suk_delle_universita_telematiche_boom_di_iscritti_e_accuse_di_lauree_facili-15240096/).

<sup>2</sup>Mackinson, T. (2025, July 16). Lauree facili alla Link Campus: condannati Vincenzo Scotti e il vertice dell’ateneo. “Esami falsi per i poliziotti”. *Il Fatto Quotidiano*. <https://www.ilfattoquotidiano.it/2025/07/16/link-campus-lauree-facili-condanna-scotti-notizie/8063901/>.

Our results show that, after accounting for students' observable characteristics, attending an online university is associated with a lower probability of dropping out. This effect is particularly pronounced among students who enrol later in life, as well as those from technical and vocational backgrounds or Southern Italy. Conversely, students who enrol directly after high school are more likely to reconsider their initial choice and transfer to a traditional university. These findings highlight the differentiated role that online universities play in supporting student retention across diverse profiles.

The article is structured as follows: Section 2 reviews the Italian debate on online universities, alongside the academic literature on their advantages and limitations, with particular attention to dropout rates, and outlines the theoretical framework underpinning the interpretation of the results. Sections 3 and 4 present the data and methods, describing the strategies adopted to identify the causal effect of attending an online university on students' academic outcomes. Section 5 and 6 present and discuss the implications related to the main results of the analysis. Section 7 concludes and outlines the main limitations of the study by suggesting directions for future research.

## 2 Literature Review

This section introduces the main elements highlighted in the scientific literature and the broader public debate around the establishment of online universities in Italy, their advances and shortcomings, and their role in shaping students' outcomes.

### 2.1 Online universities: establishment, advantages and shortcomings

In the United States, the integration of information technology in education gave rise, as early as the 1990s, to the so-called 'great media debate', which was followed by a substantial body of literature aiming to assess the advantages and disadvantages of this phenomenon (Siemens et al. 2015). Nowadays, many scholars agree that online education offers a flexible educational environment at lower costs, enabling students who cannot afford traditional universities to enrol (Sánchez-Gelabert et al. 2020). However, notable disadvantages persist, such as the costs of training instructors, feelings of isolation, and technological gaps (Siemens et al. 2015).

Nonetheless, Allen and Seaman (2016) report that 71.4% of academic leaders in the US consider online education to be as good as or better than face-to-face learning, particularly in institutions with a high proportion of students located far from traditional service areas or abroad. Contrarily, Zaki (2022) points to some elements of concern, including a lack of motivation and concentration among online students, and the insufficient control by online professors, while also acknowledging

the benefits of financial affordability, flexible schedules, and resource sharing provided by online institutions, particularly when accompanied by self-discipline and self-regulation.

In Italy, the establishment of online universities was formalised by the financial act of 2003. [Pozzi and Conole \(2014\)](#) argue that, although the reform was intended to reduce dropout and promote lifelong learning, its implementation has predominantly relied on rather conventional teaching models with limited peer-to-peer interaction. After 2003, 11 online universities were founded in a short period, leading the government to prohibit new accreditations by 2006. [Piromalli \(2024\)](#) identifies two interrelated sets of concerns regarding the establishment of online universities in Italy: a *governance* issue, which blurs the boundaries between public and private actors, local and global dynamics, and policy-makers and entrepreneurs; and a *pedagogical* issue, which frames these institutions as the primary educational instrument capable of fostering upward social mobility through the production of “flexible” and “adaptable” students (and, consequently, workers).

Nevertheless, it is evident that online universities are perceived as structurally distinct from traditional institutions, also in terms of their positioning and function within the tertiary education market. A Delphi study conducted in 2018 among a group of informed stakeholders found that the two most important indicators for evaluating online universities were clearly related to the quality of teaching and learning. By contrast, institutional reputation and graduate employment outcomes, as well as the quality of research produced, ranked last and fourth from last, respectively, in terms of perceived importance ([Pozzi et al. 2019](#)). This perception finds empirical support in the preliminary evidence provided by [Agasisti and Coppeta \(2025\)](#), which shows that once research output is incorporated into the analysis, public universities display higher levels of efficiency, whereas online institutions experience a marked decline, as if they were being penalised for their predominantly teaching-oriented or access-expanding mission.

Looking at the online students' perspective, [Gillett-Swan \(2017\)](#) emphasises that they face barriers related to isolation and limited peer interactions. In addition, online programmes also require a high degree of responsibility and time management ([Hung et al. 2010](#)). Moreover, while online students generally possess strong digital and quantitative skills, their abilities in areas not related to technology, such as collaborative learning, student-faculty interactions, and discussions with various others, are generally weaker ([Cigdem 2014](#); [Dumford and Miller 2018](#)). Despite this evidence, most academic literature prioritises assessing students' perceptions of online learning rather than its actual effectiveness. Existing studies indicate that satisfaction and persistence in online education are linked to teaching presence, meaningful and targeted teaching styles, as well as perceived usefulness and ease of use ([Eom et al. 2006](#); [Joo et al. 2011](#)).

## 2.2 The Italian public debate around non-traditional students and institutions

In Italy, the ongoing debate concerning Italian online universities includes both criticisms and positive expectations surrounding this phenomenon. For instance, [Palmisano \(2023\)](#) argues that asynchronous teaching, along with a lack of human interactions, leads to a decline in academic performance in the short term. [Miccoli \(2023\)](#), on the other hand, suggests that students in online universities differ from traditional ones and therefore require a distinct learning environment to succeed. Furthermore, [Gavosto \(2023\)](#) highlights how online universities are often perceived as having less rigorous grading standards and a low reputation, attracting students primarily interested in obtaining a degree for public job applications. However, online universities also offer some advantages related to their greater flexibility in time management and the elimination of transfer and accommodation costs.

These contrasting views are also reflected in a recent epistolary debate, where [Simone \(2024\)](#) criticises the profit-driven and market-oriented logic of online universities and the diminished quality of the distance learning experience compared to traditional education. In response, [Caroniti \(2024\)](#) emphasises the inevitability of this phenomenon and the crucial role played by these institutions in offering educational opportunities to students who struggle in the traditional system, such as working individuals. Therefore, their contribution could be particularly significant in the context of Italy's persistent low rate of university graduates.

To this extent, it is useful to look at Italian online universities' student composition. [Palmisano \(2023\)](#) notes that following the COVID-19 pandemic and the resulting cultural shift, the gap between traditional and online universities in terms of student demographics has narrowed. Previously, online universities were primarily chosen by older students who were already employed and seeking to graduate. However, over the past decade, Italian online universities have seen a 20% decline in students over 31 years old, coupled with a nearly 12% increase in students under 24 years old. Along the same lines, [Perla \(2024\)](#) reports that an increasing number of secondary school graduates are now enrolling in online universities immediately after high school completion. One factor that may help explain this shift, as noted by [Marra \(2024\)](#), is the legal equivalence of academic degrees in Italy, which ensures that the qualifications obtained from traditional and online universities are equally recognised by law. A comprehensive picture of this emerging framework is provided by [Minerva et al. \(2024\)](#) and [Di Santo et al. \(2025\)](#), who highlight at least two key dynamics: (i) a marked increase in enrolment in online universities among students from Southern Italy, largely to the detriment of Southern public universities; (ii) an overrepresentation of men in the overall student population, although women have come to account for approximately half of all online students in recent years; and (iii) a greater

attractiveness of online universities on students with technical or vocational secondary education backgrounds and/or lower average grades.

The debate can thus be summarised as consisting of two opposing positions: one that recognises the opportunities arising from the growth of online universities, and another that highlights concerns regarding their management practices. In this context, our study contributes to the ongoing debate by providing empirical evidence on the performance and outcomes of students in online universities, helping to inform a discussion that has so far remained largely unexplored in the academic literature.

### 2.3 Students' outcomes: Traditional vs Online universities

In this study, we focus on student dropout as the main outcome of interest. As previously mentioned, this phenomenon is particularly relevant in the Italian context, especially considering that, on average, dropout rates are lower in online universities than in traditional ones (Prenkaj et al. 2020). Therefore, the growth of online education may represent a solution for students who leave traditional programmes, offering an alternative pathway to succeed in their higher education studies (Kang and Wang 2018).

University dropout is a relevant phenomenon from an interconnected and multidimensional perspective: it represents a waste of students' time, as well as public and private resources, and it negatively affects labour market outcomes (Aina et al. 2022). In Italy, efforts to mitigate this phenomenon included a comprehensive reform introduced in 2001, which established a division between undergraduate and graduate programmes. This reform contributed to reducing dropout rates by increasing flexibility in degree structures (Di Pietro and Cuttillo 2008; Belloc et al. 2010). Nonetheless, several contextual factors continue to keep the issue of university retention high on the agenda of policy-makers. While some scholars identify a negative association between regional unemployment rates and university dropout rates in Italy (Di Pietro 2006), suggesting that those who leave university may be *parking* students seeking employment, the Italian academic literature generally agrees that dropout is primarily driven by disadvantaged social contexts, disproportionately affecting more vulnerable students. Indeed, Ghignoni (2017) shows that during the financial crisis, dropout rates also rose, even among the most proficient and academically oriented students. Consistent with this evidence, several studies highlight the role of parental background and financial and economic conditions in shaping the likelihood of dropping out (Aina 2013; Contini et al. 2018; Contini and Zotti 2022). In particular, while parental education and occupation exhibit a moderate effect in shaping the intergenerational transmission of academic outcomes, adverse financial conditions consistently emerge as a key driver of dropout.

Regarding the case of online students, they may have a higher likelihood of dropping out due to several factors such as screen fatigue, isolation, technological issues, and a lack of institutional support

and social interaction (Lee and Choi 2011; Rahmani et al. 2024). Interestingly, it seems that dropout rates in online universities are not significantly associated with specific demographic characteristics. Instead, they identify difficulties in balancing personal life and new study habits as a key factor. Students are not always able to benefit from the greater flexibility, often finding themselves forced to study during weekends, missing deadlines, overwhelmed by the volume of study materials, and unsatisfied with the lack of teacher presence or immediate feedback (Willging and Johnson 2009; De la Varre et al. 2014).

Despite this body of literature, to the best of our knowledge, the differences in academic outcomes of online and traditional students have not been explored. This gap is partly due to the limited availability of administrative data covering the students' academic careers. Moreover, online university students still represent a relatively small share of the overall student population, making it more difficult to conduct robust comparative analyses. Therefore, our contribution to this debate is to exploit the availability of detailed archive data by analysing the impact of attending an online degree programme on students' academic outcomes in Italy.

## 2.4 Theoretical background

The analysis presented in this paper aims to investigate how the decision to enrol in an online university affects the likelihood of dropout compared to enrolment in a traditional, in-person institution. In this section, we outline the main theoretical frameworks that inform our empirical strategy. As highlighted by Aina et al. (2022), decision-making processes and outcomes in education have been extensively studied across different disciplines, each offering relevant insights into the mechanisms underlying students' choices.

A primary theoretical framework for analysing educational choices is Human Capital Theory (Becker 1975). According to this perspective, individuals invest in education in order to acquire skills and knowledge that yield positive lifetime returns, weighing expected costs and benefits. Building on this framework, students evaluate the relative costs and benefits of alternative educational pathways. Online institutions may reduce both direct and indirect costs (e.g., commuting, housing, and opportunity costs), making them particularly attractive for non-traditional students.

However, Robeyns (2006) questions the generalisability of this framework, arguing that it overlooks the social, cultural and non-material dimensions of education, and tends to overemphasise knowledge acquisition while underestimating the broader instrumental role of education in contemporary societies. These limitations are particularly relevant in the context of online universities. Their flexible structure enables access both for individuals pursuing education for intrinsic motivations and for those facing constraints related to socioeconomic disadvantage or time availability, such as working students. Moving beyond a narrow cost-benefit perspective, the Capability Approach (Sen 1997)

provides a more suitable framework, as it explicitly acknowledges that educational attainment may serve as a means to achieve broader life outcomes. In this perspective, obtaining a degree may function as a credential required to access or progress within the labour market, even when the relevant skills have already been acquired. This interpretation is consistent with Social Closure Theory (Weber 1978) and Credentialism (Collins 2011), which emphasise the role of education as a mechanism for regulating access to opportunities through formal qualifications.

At the same time, the implications of this choice for dropout can be interpreted through Tinto (1975) model of student integration. According to this framework, persistence depends on both academic and social integration. While online universities may limit opportunities for social integration, they may enhance academic integration for students who would otherwise face constraints in traditional settings, for instance due to work commitments or limited time availability.

Taken together, these theoretical perspectives suggest potentially competing mechanisms. On the one hand, lower levels of social integration in online settings may increase the likelihood of dropout. On the other hand, reduced costs and greater flexibility may strengthen persistence, particularly among non-traditional students.

Grounding on these considerations, the following analysis investigates the characteristics of students enrolled in online and traditional universities and the effect of online attendance on students' academic outcomes between their first and second year at the university.

### 3 Data

To estimate the causal effect of university online attendance on students' university outcomes, we exploit the information included in the MOBYSU.IT database. MOBYSU.IT contains administrative data on the careers and the characteristics of the population of students enrolled for the first time in an Italian university between the years 2010 and 2019. Due to the effects of the COVID-19 pandemic, we exclude the cohorts of students enrolled after 2019 (i.e., 2020 and 2021), as all students, regardless of institution type, were exposed to some degree of online teaching, potentially confounding the distinction between traditional and online universities.

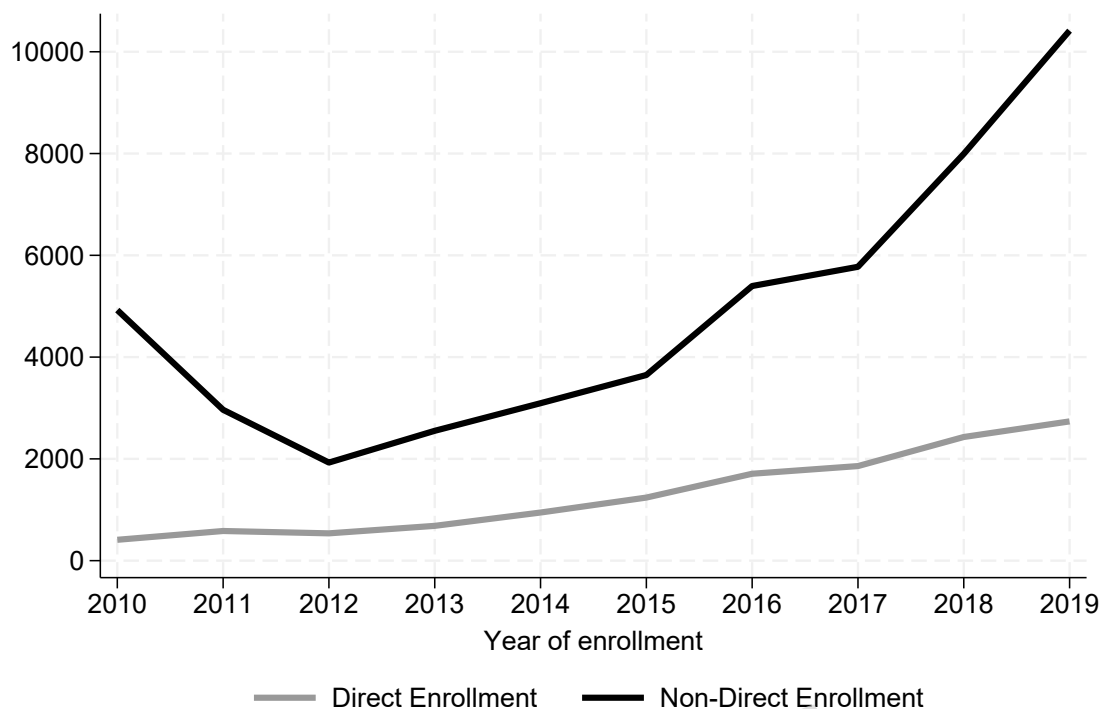
A key information included in MOBYSU.IT regards the type of institution chosen by the students, which allows us to define our variable of interest or *treatment* variable. In particular, we classify students as *treated* if they enrol in an online institution, while *non-treated* or potential *controls* are those enrolled in traditional institutions.

The data on students' careers is used to define their academic outcomes. We classify students into three categories based on the comparison between their enrolment status in the first and second academic years: *dropout*, *churn*, and *ongoing*. Dropouts are all students who are not enrolled in any

Italian university during their second year. As outlined in the Introduction, this is our main outcome of analysis, given its importance in the Italian context, characterised by a lower share of university graduates with respect to the European average (EUROSTAT 2025a,b). Churners are those who decide to change the type of institution between their first and second year. This category includes students who were initially enrolled in online (traditional) universities and switched to traditional (online) universities during the second year. Ongoing students are those who continue their university career without dropping out or changing to a different institution type.

The information on students' characteristics is exploited to account for the elements that may represent potential confounders in estimating the causal effect of online university attendance on students' outcomes. Specifically, MOBYSU.IT includes data on students' sex, age at enrolment, year of enrolment, and high school background. Concerning the high school background, we observe students' high school location, their final mark, and the type of track attended. The location of the high school allows us to account for the heterogeneity in students' choices and performances that may depend on their area of origin. We use the high school macro-region since students may change their residence when enrolling at the university, and we do not observe students' residence before enrolment. Students' final grades and track attended allow us to account for students' performances and for the type of high school track attended, which may influence both their probability of enrolling in an online university and academic outcomes. High school tracks in Italy can be classified into four broad categories: *Licei*, technical high schools, vocational high schools, and a residual category. *Licei*, in general, are more academically-oriented high schools. Technical institutes provide students with a specialisation in a particular subject (e.g., accounting, surveying). Vocational institutes are characterised by laboratory and professional activities and are oriented to train students to enter the workforce directly after graduation (Contini et al. 2017). Moreover, since *Licei* are characterised by a high degree of heterogeneity in terms of disciplinary contents, we decided to further divide this category into three curricula: classical *licei*, scientific *licei*, and other *licei*. An important element related to the high school track is the social stratification associated with these categories. In general, *licei* are considered the most academic-oriented tracks, and students from these schools are more likely to come from a better socioeconomic background, with more financial resources to support their tertiary education career (Impicciatore and Tosi 2019; Barone and Assirelli 2020; Priulla et al. 2025; Usala et al. 2025).

Another key information available in MOBYSU.IT is the year of students' high school graduation and their year of university enrolment. As shown in Figure 1, students who enrol after a gap period following high school graduation (*Non-Direct*) represent the majority of students enrolled in online institutions. Nonetheless, even the number of students who enrol directly after the completion



**Fig. 1:** Number of first-year students enrolled in online programmes: direct vs non-direct

*Notes:* The figure shows the number of first-year students enrolled in online universities by year of enrolment, distinguishing between those who enrolled directly after high school (direct enrolment) and those who enrolled after a gap period (non-direct enrolment).

of secondary education is constantly increasing over time. Therefore, since these two groups of students may have different incentives with respect to online universities, in Section 5 we explore the heterogeneity in the effect of online attendance by analysing each group separately.

Finally, since students in more remote areas may have different incentives with respect to the choice of enrolling in an online institution, we computed the minimal travel distance between each student's city of residence and the nearest university.<sup>3</sup>

Based on this information, we ensure the comparability between online and traditional students by applying a set of eligibility criteria. We start from a total of 2,627,392 students: 2,553,735 enrolled in traditional universities and 73,657 in online ones. To avoid comparing students in very different settings, we first exclude all students enrolled in degree programmes not offered by any online institution (711,098 traditional students). Then, due to our empirical strategy, we do not include students without information on their high school background or city of residence (101,307 traditional students and 11,830 online students). Therefore, our final data includes 61,826 treated students and 1,741,184 potential control units. We lose 16% of treated units and 31.8% of potential control units.

<sup>3</sup>The travel distances have been computed by the National Statistics Institute (ISTAT) by accounting for various impedance factors (e.g., traffic conditions, and other potential delay factors). See <https://www.istat.it/non-categorizzato/matrici-di-contiguita-distanza-e-pendolarismo> for more information.

### 3.1 Descriptive statistics

Table 1 presents the descriptive statistics of the students' characteristics included in our final dataset.

**Table 1:** Descriptive statistics

<i>Variables</i>	<b>Full Sample</b>	<b>Traditional</b>	<b>Online</b>	<b>Online</b>	<b>Online</b>
		<i>Total</i>	<i>Total</i>	<i>Direct</i>	<i>Delayed</i>
No. Observations	1,803,010	1,741,184	61,826	13,134	48,692
<b><i>HS Final Mark</i></b>	78.79	78.99	73.15	73.85	72.96
(st. dev.)	(11.77)	(11.75)	(10.80)	(10.54)	(10.86)
<b><i>HS Regularity</i></b>					
Non-regular	16.02%	15.42%	32.85%	29.81%	36.67%
Regular	83.98%	84.58%	67.15%	70.19%	66.33%
<b><i>HS track</i></b>					
Classical	13.26%	13.60%	3.45%	4.71%	3.12%
Scientific	35.46%	36.23%	13.67%	21.70%	11.51%
Other Licei	20.19%	20.43%	13.56%	20.21%	11.77%
Technical	22.89%	21.97%	48.67%	35.40%	52.25%
Vocational	6.73%	6.55%	11.83%	11.69%	11.87%
Other institutes	1.48%	1.22%	8.81%	6.29%	9.49%
<b><i>HS macro-region</i></b>					
North	39.77%	40.35%	23.23%	23.12%	23.27%
Center	20.09%	20.09%	19.98%	20.38%	19.88%
South & Islands	40.15%	39.56%	56.78%	56.50%	56.86%
<b><i>Minimum distance from nearest university</i></b>					
≤ 30 min	81.12%	81.26%	77.12%	72.18%	78.46%
> 30 min	18.88%	18.74%	22.88%	27.82%	21.54%
<b><i>Sex</i></b>					
Male	45.56%	44.94%	62.83%	55.40%	64.83%
Female	54.44%	55.06%	37.17%	44.60%	35.17%
<b><i>Age</i></b>	20.08	19.73	29.85	20.54	32.36
(st.dev.)	(4.22)	(3.27)	(11.02)	(5.35)	(10.80)
<b><i>Field of study</i></b>					
Agriculture, forestry, fisheries and veterinary	3.29%	3.38%	0.80%	1.04%	0.73%
Arts and humanities	17.67%	18.20%	2.86%	4.32%	2.47%
Business, administration and law	26.36%	26.30%	28.15%	22.60%	29.65%
Education	4.81%	4.68%	8.43%	7.90%	8.57%
Engineering, manufacturing, and construction	18.97%	19.17%	13.11%	10.51%	13.82%
Health and welfare	1.60%	1.65%	0.06%	0.08%	0.06%
Natural sciences, mathematics and statistics	4.12%	4.25%	0.38%	0.94%	0.23%
Social sciences, journalism and information	18.46%	18.09%	28.83%	24.21%	30.08%
Services	4.72%	4.27%	17.37%	28.38%	14.4%
<b><i>Academic outcomes</i></b>					
Dropout	12.20%	12.05%	16.32%	14.15%	16.90%
Churn	0.37%	0.32%	1.82%	6.11%	0.66%
Ongoing	87.43%	87.63%	81.86%	79.74%	82.44%

*Notes:* The table reports the descriptive statistics regarding students' characteristics and outcomes. Column 2 refers to all students, while Column 3 focuses on traditional students. Columns 4-6 include the data regarding, respectively, all online students, online students with direct enrolment after high school, and online students with non-direct or delayed enrolment.

Regarding high school background, notable differences emerge between students enrolled in traditional and online universities. On average, students in online universities obtained their high school diploma with a final mark that is 7.39% lower than that of those enrolled in traditional institutions. Only 67.15% of online students completed secondary education without delay, compared to 84.58%

of their counterparts in traditional universities. Moreover, over 50% of online students come from technical or vocational schools, rather than from *licei*.

As for students' high school macro-region, we observe that more than 50% of online students come from the South & Islands, whereas this area accounts for only one-third of traditional university students. Moreover, less than 80% of online students reside in a town located at most 30 minutes away from the nearest traditional university. As expected, this share is lower than that of students who, despite living far from traditional institutions, still choose to enrol in one of them.

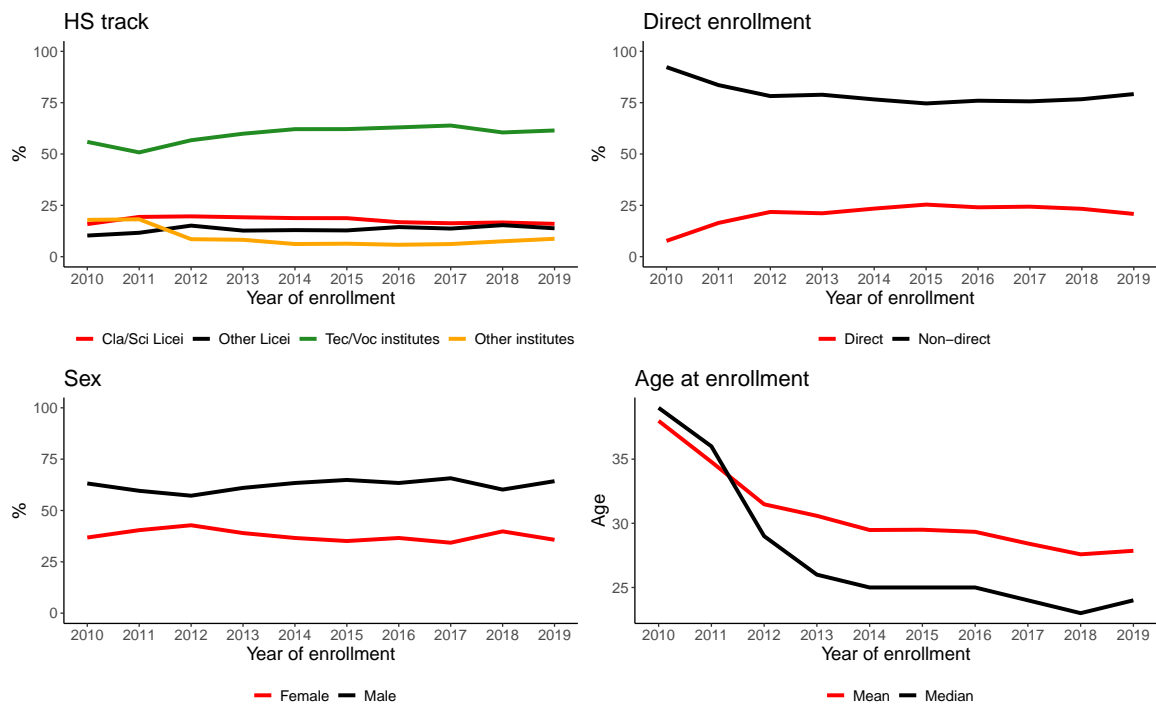
The table also provides information regarding students' characteristics. We can notice that females outnumber male students only in traditional universities, while they represent only around one-third of online students. Unsurprisingly, the average age at enrolment is significantly higher among students at online universities (29.85 years vs 19.73), which aligns with the commonly held view that online institutions are the preferred option for those who choose to pursue higher education later in life.

The share of online students is higher in the following fields of study: Business, Administration and Law; Education; Services; and Social Sciences, Journalism and Information. Instead, the lowest proportion is observed in Health and Welfare. As explained in Section 4, we account for these differences in the probability of choosing online institutions by comparing only students enrolled in the same field of study.

With respect to students' academic outcomes, the table shows that the proportion of online students who either drop out or transfer to a traditional university is higher than the proportion of traditional students making similar transitions.

Table 1 enables us also to distinguish between online students who enrolled immediately after obtaining their high school diploma (direct enrolment) and those who delayed their enrolment (non-direct or delayed enrolment). As explained in Section 2, this distinction is relevant, as online universities are often considered the most suitable option for working students seeking a degree. Indeed, the data show that *delayed* online students are predominantly males, have the lowest average high school final mark, often obtained beyond the standard duration, and are less likely to have attended a *liceo*. It is important to note that, while dropout rates are relatively similar between the two groups, a marked difference emerges in terms of churn: 6.11% of direct online students transfer to traditional universities, compared to only 0.66% observed among non-direct students. This suggests that students who enrol in an online university without delay may be more likely to reconsider their initial choice.

Figure 2 further examines whether the composition of students enrolled in online universities has changed over time, following the expansion of online enrolments shown in Figure 1. The results show that the composition has remained broadly stable across most observable characteristics. The main exception is student age, which has decreased over time, consistent with the growing share of



**Fig. 2:** Online students' characteristics over time

*Notes:* The figure shows the evolution over time of the composition of students enrolled in online universities between 2010 and 2019. Each panel displays the share of students by high school track, enrolment timing (direct vs. non-direct), and sex, as well as the mean and median age at enrolment.

direct enrolments. As discussed in Section 5, we account for this change in students' composition by including age at enrolment in the logit estimation and by distinguishing between direct and non-direct enrolments in the estimation of the causal effect of online attendance.

## 4 Identifying the effect of online university attendance

This section outlines the methodological approach used in this study. Section 4.1 characterises students who enrol in an online university using a standard logit approach. Building on this information, Section 4.2 describes the propensity score matching approach used to identify the causal effect of attending an online university on students' university outcomes between their first and second years.

### 4.1 Who chooses online universities?

Since students self-select into the type of university, a simple comparison between online and traditional students may lead to biased estimates, as the same factors that influence enrolment choices may also affect academic outcomes (Cochran and Rubin 1973).

To account for this selection bias, we exploit the information available on students' characteristics related to their enrolment choice. Specifically, we estimate a binary logit model where the dependent variable indicates whether a student enrolled in an online institution. The model includes individual

characteristics (sex, age at enrolment, and year of enrolment), high school background (final mark, regularity, and high school track attended), geographic characteristics (macro-region and distance from the nearest university), the ISCED field of study, and a binary variable indicating whether the student enrolled at university immediately after completing high school (direct enrolment). The estimated propensity scores are then used in Section 4.2 to account for selection bias in the estimation of the causal effect of online attendance on students' academic outcomes.

Table 2 reports the estimated coefficients from the logit model. The results provide insights into the observable characteristics associated with the choice of enrolling in an online university. Female students are less likely than males to enrol online. As for the high school background, a higher inclination to enrol in online universities is observed among students with a non-*liceo* background, who studied in a southern region, and with a regular career. Conversely, students with higher final high school marks are less likely to enrol in online universities, pointing to a preference for traditional universities among higher-achieving students. A positive effect is also found for distance from the nearest university, suggesting that commuting and housing costs increase the relative attractiveness of online institutions for students in more remote areas.

Further insights arise regarding students' characteristics at the enrolment. Direct enrolment emerges as a crucial deterrent to attending an online university, while the probability of online enrolment increases with age. The results also confirm the positive trend shown in Figure 1, with the magnitude of the coefficients associated with year of enrolment increasing from 2010 to 2019. Finally, substantial variability is observed across fields of study, likely reflecting differences in the availability of online degree programmes.

## 4.2 The propensity score matching approach

As outlined in the previous section, estimating the causal effect of online university attendance on students' academic outcomes requires addressing their non-random selection into online institutions. Indeed, as shown in Table 2, the decision to enrol in an online university is systematically associated with students' sociodemographic characteristics and educational background. Consequently, a simple comparison between treated and control units may yield biased estimates.

To address this issue, we adopt a Propensity Score Matching (PSM) approach which helps in reducing the selection bias by balancing covariates across treated and control groups. It consists of matching treated units with comparable controls based on the probability of enrolling in an online university conditional on students' observed characteristics (the propensity score). The key insight here is that matching units based on the PS is sufficient to eliminate selection bias. A significant advantage of this approach is that it reduces the complexity of matching from multiple covariates to a single scalar value, significantly simplifying the process of matching treated and control units

**Table 2:** Logit results: online versus traditional university enrolment.

Variable	Coefficient	S.E.
Intercept	-6.47***	0.07
<i>Sex</i>		
Female	-0.28***	0.01
<i>HS Track</i>		
Scientific	0.22***	0.03
Other Licei	0.61***	0.03
Technical institutes	0.95***	0.03
Vocational institutes	0.84***	0.03
Other institutes	2.17***	0.03
<i>HS final mark</i>	-0.02***	0.00
<i>HS Macro-region</i>		
Center	0.57***	0.01
South & Islands	0.92***	0.01
<i>HS Regularity</i>		
Regular	0.07***	0.01
<i>Direct enrolment</i>		
Direct	-2.04***	0.01
<i>Age at enrolment</i>		
	0.01***	0.00
<i>Distance from university</i>		
	0.01***	0.00
<i>Year of enrolment</i>		
2011	-0.15***	0.03
2012	-0.31***	0.03
2013	-0.04	0.03
2014	0.25***	0.03
2015	0.45***	0.02
2016	0.72***	0.02
2017	0.78***	0.02
2018	1.13***	0.02
2019	1.33***	0.02
<i>Field of study</i>		
Arts and humanities	-0.14*	0.06
Business, administration and law	1.86***	0.05
Education	1.83***	0.05
Engineering, manufacturing and construction	1.78***	0.05
Health and welfare	-2.16***	0.17
Natural sciences, mathem	-0.45***	0.08
Services	2.80***	0.05
Social sciences, journalism and information	1.97***	0.05
Number of observations	1,803,010	

*Notes:* The table reports the logit estimation results with respect to students' choice to enrol in an online university. For each covariate, it reports the estimated coefficient and the standard error. \*\*\*, \*\*, \* indicate, respectively, a p-value lower than 0, 0.001, 0.01

(Rosenbaum and Rubin 1983). The result of this procedure is a dataset in which treated and control units are comparable in terms of observed characteristics. By improving the covariate balance, the PSM helps us to approximate the conditions of a randomised experiment. Treatment effects can then be estimated by applying regression techniques using only the information on treated and matched control units.

It is important to note that this approach relies on the assumption of unconfoundedness (or conditional independence), which requires that, conditional on the covariates, treatment assignment is independent of potential outcomes. While this assumption is untestable, its plausibility is supported

by the richness of the MOBYSU.IT database, which includes a detailed set of individual and contextual characteristics related to both the choice of enrolling in an online university and students' academic outcomes. The main residual concern relates to students' socioeconomic background and parental education, which are not directly available in our data but are partially proxied by students' high school track. Moreover, while this source of unobserved heterogeneity cannot be ruled out, the direction of the potential bias can be anticipated. As shown in Table 1, online students are substantially overrepresented among those from technical and vocational high school tracks, which in the Italian context are generally associated with lower socioeconomic backgrounds (see, for example, [Barone and Assirelli \(2020\)](#); [Usala et al. \(2025\)](#)). If students from lower socioeconomic backgrounds face higher dropout risks, omitting this variable would likely lead to an underestimation of the true positive effect of online attendance on student retention. In this sense, our estimates can be interpreted as a conservative lower bound of the true effect.

To implement the PSM, we need to define two key elements: the algorithm used to match online and traditional students and the method used to estimate the effect of online attendance on students' academic outcomes.

#### ***The matching algorithm***

Once the propensity score (PS) is estimated, units can be matched using various algorithms. In this study, we combine the elements of different techniques to improve covariate balance while maintaining computational feasibility.

As a baseline, we use the nearest-neighbour matching procedure, which matches each online student with the most similar traditional student in terms of PS. Then, to improve match quality, we introduced two additional constraints. First, a caliper of 0.05 on the PS scale, which restricts matches to students whose estimated probability of online enrolment lies within  $\pm 5$  percentage points of each other ([Arpino and Mealli 2011](#)). Second, we required exact matching on the field of study of enrolment, categorised into nine groups based on the ISCED classification to ensure that matched students are exposed to similar academic content during their first year. It is important to note that this constraint does not affect the estimation of the PS, but it forces the algorithm to match only students with identical values for the selected variable.

#### ***The estimation approach***

To estimate the causal effect of attending an online university on students' outcomes in the transition between the first and second years, we apply a multinomial logistic regression model to the matched data including only the information on treated units and matched control students. This choice

reflects our interest in capturing the three possible outcomes faced by students at this stage: *ongoing*, *churn*, and *dropout*.

Formally, we define the categorical variable  $Y_i \in \{0, 1, 2\}$  as the outcome observed for student  $i$ . In our case,  $Y_i = 0$  corresponds to students who have enrolled in the second year in the same type of institution (*ongoing*),  $Y_i = 1$  indicates a switch between online and traditional universities (*churn*), and  $Y_i = 2$  captures the students that are not enrolled in any Italian university in the second year (*dropouts*). Accordingly, students' probability of being in any of the three states is modelled as:

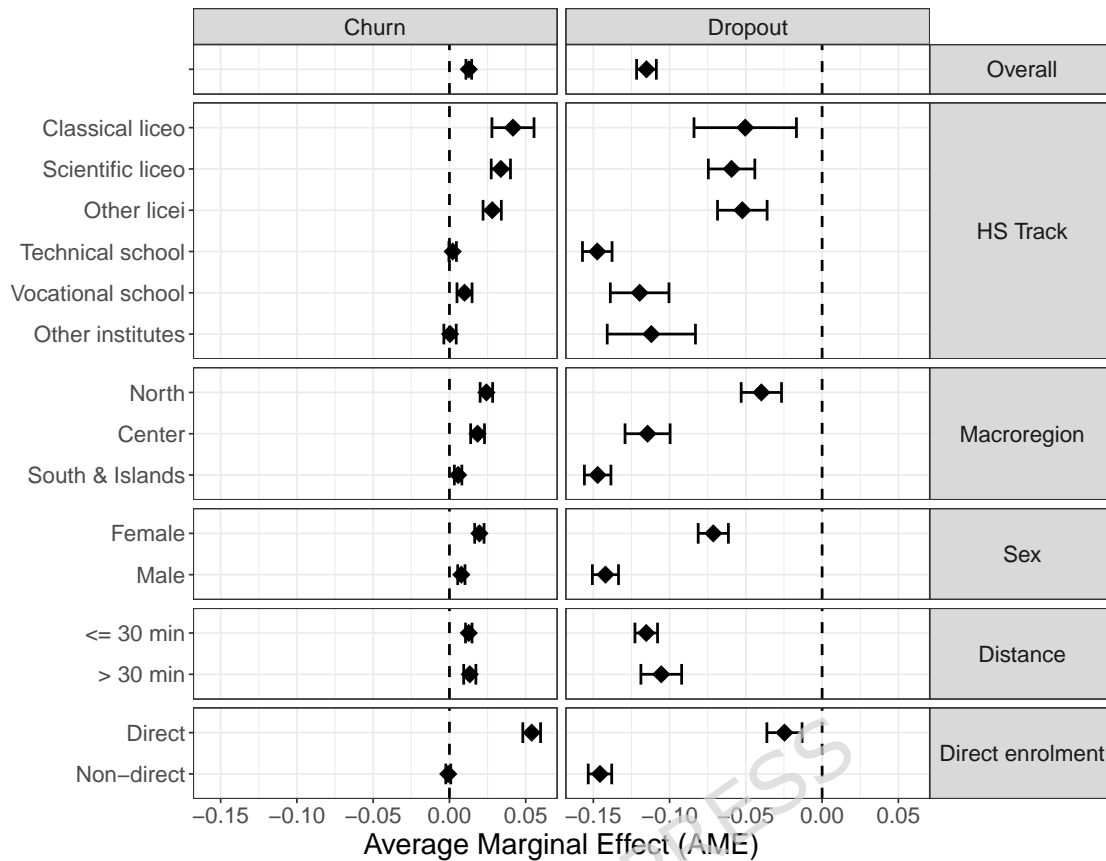
$$\Pr(Y_i = j | T_i) = \begin{cases} \frac{1}{1 + \sum_{k=1}^2 \exp(\alpha_k + \beta_k T_i)} & \text{if } j = 0 \\ \frac{\exp(\alpha_j + \beta_j T_i)}{1 + \sum_{k=1}^2 \exp(\alpha_k + \beta_k T_i)} & \text{if } j = 1, 2 \end{cases} \quad (1)$$

where  $T_i$  is the treatment indicator equal to 1 if student  $i$  enrolled in an online university and 0 otherwise. The baseline category is  $Y_i = 0$  (*ongoing*), which serves as the reference group. The parameters  $\alpha_j$  and  $\beta_j$  capture the intercept and the effect of online attendance on the likelihood of observing outcome  $j$  in contrast to the base category.

Finally, to facilitate results' interpretation, the estimated coefficients are transformed into Average Marginal Effects (AMEs). AMEs measure the change in the predicted probability of observing each outcome caused by the online university attendance. Therefore, they provide a more intuitive understanding of the estimated treatment effect. A key advantage of AMEs is that they allow us to explore heterogeneity in the effect of online attendance by replicating the analysis across different student subgroups, such as those defined by sex, high school track, or geographic origin. This approach allows us to better refine the balance of observable covariates and to investigate the intersectionality of the determinants of university choices and outcomes.

## 5 Results

This section presents the results of our analysis. The matching procedure results in a loss of 13.33% of treated students. These students are characterised by atypical profiles for which no comparable counterparts exist among traditional university students, according to the chosen algorithm. These students tend to be considerably older, are almost exclusively enrolled on a non-direct basis, and are more concentrated in specific fields of study, such as Engineering and Services. The covariate balance achieved after matching is satisfactory across all groups, as discussed in detail in Appendix B. Building on the matched sample, we estimated the effect of online attendance for the average student and explored its heterogeneity across subpopulations.



**Fig. 3:** Average Marginal Effects

*Notes:* The figure shows the causal estimates of the effect of online attendance in terms of AMEs. Each AME measures the change in the probability of observing the outcome due to the treatment. To improve visual clarity, we show results for churn and dropout only, omitting the ongoing category (the baseline). Table C4 in the Appendix reports the numeric values associated with each AME. The results are reported for the overall population and each subgroup.

Figure 3 presents the results of the multinomial logistic regression models in terms of AMEs. As explained in Section 4.2, we explore the heterogeneity in the impact of online attendance by applying the procedure to both the overall population of online students and to subgroups of the same population stratified by high school track, macro-region, sex, distance from the nearest university, and whether students enrolled directly at the university after completing high school.

The results provide valuable insights into the effects of online attendance on students' outcomes. In the overall population, the results indicate that online students are significantly less likely to drop out ( $AME = -0.12$ ), while the effect on the churn probability seems marginal ( $AME = 0.01$ ). This finding contrasts with the descriptive statistics in Table 1, which indicate a higher dropout rate among online students. This suggests that the higher dropout rate observed in the data is mostly attributable to differences in student composition rather than to the type of institution.

A more detailed picture emerges from the stratified analysis. Although the deterrent effect on dropout remains visible across all subgroups, it is stronger for students who typically face greater

challenges in traditional universities. Indeed, these students appear to benefit the most from enrolling in an online institution.

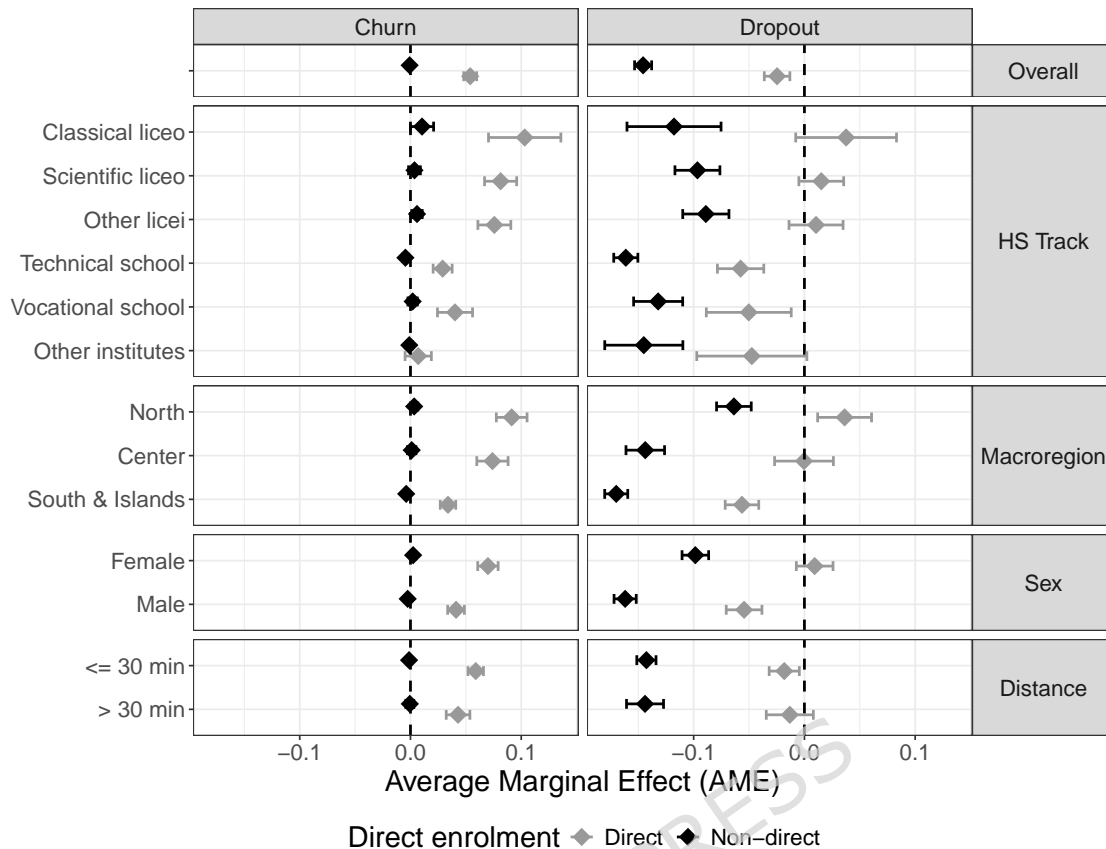
In terms of high school background, students from technical and vocational tracks, who generally show higher dropout rates in traditional settings (Priulla and Attanasio 2024), are substantially less likely to drop out when enrolled in online universities (-0.15 and -0.12 for technical and vocational students, respectively). A similar pattern is observed among students from Southern Italy, who not only have higher dropout rates in traditional universities but also represent the majority of all students in online institutions.

With respect to individual characteristics, male students also appear to benefit more from online university enrolment. Interestingly, the effect of attending an online university does not vary significantly with respect to the distance from the nearest university. The most striking difference is found in the comparison of students who enrol directly at university after high school with those who delay university entry. In this case, the latter shows a considerably lower probability of dropping out when enrolled in online universities (-0.02 and -0.15 for direct and non-direct students, respectively). Conversely, the effect of online attendance appears to be in the opposite direction when considering switching between university types. Indeed, while we observe a general increase, despite minimal, in the probability of churn, the effect is stronger for students who typically encounter fewer difficulties in traditional universities (e.g., those with a *liceo* diploma or who enrolled immediately after high school).

As a robustness check, we replicated the analysis using Inverse Probability Weighting (IPW) as an alternative estimation strategy. Rather than considering only the pairs of matched treated and control units, IPW reweights observations by the inverse of their estimated propensity score. This strategy allows to weight more the observations that are closer to the treated units. Results are reported in section A in the Appendix. The results are virtually identical to those obtained with the nearest-neighbour matching procedure, confirming that our findings are not sensitive to the choice of algorithm.

### *Direct enrolment and online attendance*

The results presented in Figure 3 highlight the differing sensitivity to the effect of online attendance based on whether students enrolled at university directly after completing high school or after one or more years. Accordingly, we extended our analysis by dividing the population into *direct* and *non-direct* enrolled. This further investigation aligns with our aim to offer meaningful insights into the intersectionality of educational choices and the different roles that online attendance may have based on students' characteristics and needs.



**Fig. 4:** Average Marginal Effects: direct enrolment

*Notes:* The figure shows the causal estimates of the effect of online attendance in terms of AME. Grey dots refer to students enrolled directly after high school, while black dots indicate students enrolled with a delay. Therefore, the results for the classical *liceo* are estimated considering either the direct (grey dots) or non-direct students (black dots) who have attended a classical *liceo* track. Table C4 in the Appendix reports the numeric values associated with each AME.

Figure 4 shows the AMEs with respect to these two groups of students. Section B in the Appendix presents the results of the matching procedure.

Considering students' dropout, we can notice that, while the overall pattern is qualitatively similar to the one observed in Figure 3, the positive effect of online attendance is much stronger for students who delayed university enrolment. Within this group, AMEs indicate a notable reduction in the probability of dropping out, indicating that the general positive effect of online attendance on students' dropout is mainly driven by this category of students. In general, the effect of online attendance on the dropout probabilities of direct students remains significant only for specific subgroups, such as those from technical and vocational tracks, students from Southern Italy, and male students.

A different picture emerges when considering the probability of churn, where online attendance seems to affect only direct students. Indeed, AMEs reveal a marked increase in the likelihood of switching to a traditional university in almost all the considered groups. In contrast, the effect on non-direct students appears almost negligible.

## 6 Discussion of the results

This article analyses the differences in academic outcomes between students enrolled in traditional and online Italian universities. To the best of our knowledge, this is not only one of the first Italian studies to quantitatively examine the effects of online universities on students' academic outcomes, but also one of the very few in the international literature to benefit from data that allow for a direct comparison between traditional and online students. The decision to initiate this line of research by focusing specifically on the probability of dropping out is not coincidental. Student retention has become a particularly urgent issue in the Italian higher education system, given the country's low graduation rate compared to the European average. In this context, a key argument in favour of the expansion of online universities lies in the idea that these new institutions may contribute to retaining students who tend to struggle most in the traditional education system. Moreover, the number of online students has steadily increased over the past decade, including those who choose this type of institution immediately after high school. This recent trend runs counter to the commonly held view that online universities primarily serve working students who pursue higher education at a later stage in life.

### 6.1 Online university students' profile

Both the descriptive statistics and the results of the logit model on the propensity to enrol in online universities confirm that the online students are more likely to come from non-*licei*, have lower final marks, and originate from Southern Italy or towns located far from the nearest university. The literature on the Italian university student population has extensively shown that these characteristics are associated with groups from lower socioeconomic backgrounds, with fewer financial resources, for whom the economic, social, and psychological costs of transitioning to the tertiary education system are higher (Impicciatore and Tosi 2019; Lombardi et al. 2019; Barone and Assirelli 2020; D'Agostino et al. 2021; Usala et al. 2025).

In addition to these socio-educational and geographic patterns, gender differences also emerge. Despite women representing more than half of the overall student population in the Italian university system, online universities host a particularly high proportion of male students. In particular, the data show that females in Italy are more likely to enrol in university, have higher graduation rates, and perform better overall, even in response to the challenges of transitioning to the university system (D'Agostino et al. 2022; Priulla and Attanasio 2024; Tocchioni et al. 2025).

Finally, the strong negative association between immediate enrolment in higher education after secondary school and enrolment in online universities represents an important finding that warrants further investigation. This pattern can be interpreted in light of two theoretical perspectives, which

may be seen as either complementary or competing. On the one hand, delayed enrolment in online universities may reflect a rational cost–benefit decision, consistent with the framework proposed by Becker (1975). On the other hand, this choice may also be examined through a credentialist perspective, as advanced by Collins (2011), whereby the primary objective is not the acquisition of skills and knowledge, but rather the attainment of a formal qualification.

## 6.2 Online universities and students' outcomes

The substantial difference in student composition between traditional and online universities required the implementation of a Propensity Score Matching. In the Italian context, where causal analyses based on administrative data are still relatively rare, this approach represents a significant methodological advancement. It enables a comparison of academic outcomes based solely on individuals who are most similar to each other in terms of observable characteristics. In this way, it reduces the bias arising from the fact that the same characteristics influencing students' decision to enrol in an online university may also affect their likelihood of dropping out or changing the type of institution. Therefore, it avoids misleading comparisons based on individuals with very different likelihoods of enrolling in an online university, and allows us to isolate the effect of the institutional type from the influence of student composition. Of course, this represents a preliminary exploration of this causal effect in the Italian context, which we hope will serve as a stepping stone for more robust evidence in future research.

Surprisingly, the multinomial logit analysis on the matched sample generally shows that students enrolled in online universities have significantly higher probabilities of continuing their studies and lower probabilities of dropping out. Although transitions between online and traditional universities are very rare, students in online institutions are slightly more likely to move to a traditional university than the reverse.

These findings become even more relevant when interpreted in light of students' characteristics. Indeed, the stratified analysis shows that the effect of online attendance varies across subgroups. While students from *licei* are more likely to transfer from online to traditional universities, those from technical and vocational schools benefit more from online programmes, showing significantly higher probabilities of continuing their studies and lower probabilities of dropping out. A similar pattern applies to male students, those from Southern Italy, and those who did not enrol immediately after completing secondary education. From this perspective, the results seem to confirm that online universities may have a beneficial effect on student retention, particularly within the most disadvantaged segments of the university population.

The contrast highlighted above is even more pronounced when comparing the sub-samples of students who enrolled in university immediately after completing secondary education and those who

chose to enrol at a later stage in life. It reinforces the notion that online universities provide a crucial service for individuals at a point in their lives (and likely their careers) where pursuing a degree through a traditional academic path would be excessively challenging. When interpreted through the lens of [Tinto \(1975\)](#)'s model of student retention, this result suggests that online universities may offset their intrinsically lower levels of social integration by fostering stronger academic integration, especially among students who tend to struggle in traditional institutions.

From this perspective, it is worth briefly focusing on the case of *online churners*, i.e., students initially enrolled in online universities who later transfer to traditional institutions, especially those who enrolled immediately after completing secondary school. According to our data, 67% of these students attended a *liceo*, the high school track most commonly associated with progression to university. Their average final mark at high school is 74.65/100 (73.91 among those from *licei*), slightly above the general average for online university students. Moreover, 60.52% of these *liceo* churners come from Central or Northern Italy, and 22.16% reside over 30 minutes away from the nearest traditional university. Both figures are higher than the overall averages in our dataset. These students thus appear to have profiles in line with those who typically choose traditional universities. The evident mismatch they experience, enrolling initially in an online institution and transferring after the first year, may reflect a failure in the guidance policies within their local contexts and/or the fact that they originate from disadvantaged backgrounds. Such circumstances could have made a direct transition to traditional higher education too demanding economically, socially, or psychologically.

## 7 Conclusions

The findings summarised thus far carry significant policy implications. As outlined in [Section 2](#), Italy is currently experiencing an intense debate regarding the relevance and continued presence of online universities within the tertiary education system, a debate further fuelled by recent judicial investigations involving some of these institutions. However, our analysis does not allow us to determine whether the apparent positive effect of online universities on reducing dropout rates is due to the greater flexibility they offer or to more lenient standards and lighter workloads imposed by their academic leaders, as suggested by critics.

What is certain, as argued by [Harasim \(2000\)](#) and [Caroniti \(2024\)](#), is that the integration of technology and online modalities into education is an inevitable and irreversible phenomenon, with the potential to widen participation in higher education to those who may otherwise be excluded from traditional pathways ([Miccoli 2023](#)). Nevertheless, in Italy, online universities currently suffer from a relatively low reputation ([Gavosto 2023](#)), partly sustained by the perception that predominantly

private institutions are primarily incentivised to award degrees with the same legal value as those from traditional universities, rather than to ensure academic rigour (Marra 2024; Simone 2024).

Bringing these issues together, we argue that online universities are already playing a meaningful role within the Italian higher education landscape by retaining a significant segment of students who are more likely to require support. On this basis, it is the responsibility of policymakers to implement robust oversight measures to ensure that these positive outcomes are not the result of lower academic standards or excessive leniency, as critics often contend.

While our analysis provides valuable insights into the role of online universities in reducing dropout rates, it does not aim to settle the debate on the causal effect of online education on the propensity to drop out. Rather, it contributes new evidence to this issue. At the same time, our study is subject to several limitations, which point to promising avenues for future research. First, although the Propensity Score Matching approach helps mitigate selection bias, our analysis does not directly account for students' socioeconomic background or parental education. These factors may still influence both the choice of institution and academic outcomes, and their inclusion could further refine causal estimates. Second, our focus on dropout between the first and second year captures only a short-run dimension of academic success. Future studies could extend the analysis to other outcomes, such as degree completion, credit accumulation over time, or time-to-degree, to better understand long-term effects. Third, while our study concentrates on academic outcomes, future research could benefit from integrating additional information on students' employment status during their studies. Controlling for this element would help to better identify the role of online universities in supporting non-traditional learners, such as working students or those with care-giving responsibilities.

Finally, it is important to remark that, although the analysis is based on population-level administrative data, the use of propensity score matching restricts the estimation to the region of common support. As a result, the findings should be interpreted as applying to the subpopulation of students for whom enrolment in online versus traditional universities is empirically comparable. While this restriction strengthens the internal validity of the estimates, it may limit their generalisability to individuals with more extreme characteristics. In addition, differences in higher education systems and in the organisation of online provision suggest caution when extending these results beyond the Italian context.

## References

- Agasisti T, Coppeta GF (2025) De-efficiency of italian universities: A comparative analysis of public, private, and online institutions. Tech. rep.
- Aina C (2013) Parental background and university dropout in italy. *Higher Education* 65(4):437–456

- Aina C, Baici E, Casalone G, et al (2022) The determinants of university dropout: A review of the socio-economic literature. *Socio-Economic Planning Sciences* 79:101102
- Allen IE, Seaman J (2016) Online report card: Tracking online education in the United States. ERIC
- Arpino B, Mealli F (2011) The specification of the propensity score in multilevel observational studies. *Computational Statistics & Data Analysis* 55(4):1770–1780
- Barone C, Assirelli G (2020) Gender segregation in higher education: an empirical test of seven explanations. *High Educ* 79:55–78. <https://doi.org/10.1007/s10734-019-00396-2>
- Becker GS (1975) *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, 2nd edn. National Bureau of Economic Research, URL <https://www.nber.org/books-and-chapters/human-capital-theoretical-and-empirical-analysis-special-reference-education-second-edition>
- Becker SO, Caliendo M (2007) Sensitivity analysis for average treatment effects. *The Stata Journal* 7(1):71–83. <https://doi.org/10.1177/1536867X0700700104>
- Belloc F, Maruotti A, Petrella L (2010) University drop-out: An Italian experience. *Higher Education* 60(2):127–138
- Caroniti D (2024) Non ci sono atenei di serie b: anche le università telematiche aiutano lo sviluppo del paese. URL <https://www.editorialedomani.it/idee/cultura/il-pregiudizio-sugli-atenei-di-serie-b-in-difesa-delle-universita-telematiche-s1p53dpx>
- Cheslock JJ, Jaquette O (2022) Concentrated or fragmented? the US market for online higher education. *Research in Higher Education* 63(1):33–59
- Cigdem H (2014) Effects of students characteristics on online learning readiness: A vocational college example. *Turkish Online Journal of Distance Education* 15(3):80–93
- Cochran WG, Rubin DB (1973) Controlling bias in observational studies: A review. *Sankhyā: The Indian Journal of Statistics, Series A* pp 417–446
- Collins R (2011) Credential inflation and the future of universities. *Italian Journal of Sociology of Education* 3(Italian Journal of Sociology of Education 3/2):228–251
- Contini D, Zotti R (2022) Do financial conditions play a role in university dropout? new evidence from administrative data. In: *Teaching, research and academic careers: An analysis of the interrelations and impacts*. Springer, p 39–70

- Contini D, Tommaso MLD, Mendolia S (2017) The gender gap in mathematics achievement: Evidence from Italian data. *Econ Educ Rev* 58:32–42. <https://doi.org/10.1016/j.econedurev.2017.03.001>
- Contini D, Cugnata F, Scagni A (2018) Social selection in higher education. enrolment, dropout and timely degree attainment in Italy. *Higher Education* 75(5):785–808
- Database MOBYSU.IT (Mobilità degli Studi Universitari in Italia) Research protocol MUR - Universities of Cagliari, Palermo, Siena, Torino, Sassari, Firenze, Cattolica and Napoli Federico II, Scientific Coordinator Massimo Attanasio (UNIPA), Data Source ANS-MUR/CINECA
- Di Pietro G (2006) Regional labour market conditions and university dropout rates: Evidence from Italy. *Regional Studies* 40(6):617–630
- Di Pietro G, Cutillo A (2008) Degree flexibility and university drop-out: The Italian experience. *Economics of Education Review* 27(5):546–555
- Di Santo V, Trancossi S, Turri M (2025) University online teaching in Italy: Diffusion and characteristics of students. *European Journal of Education* 60(3):e70065
- Dumford AD, Miller AL (2018) Online learning in higher education: exploring advantages and disadvantages for engagement. *Journal of Computing in Higher Education* 30(3):452–465
- D’Agostino A, Ghellini G, Lombardi G (2021) Movers and stayers in STEM enrollment in Italy: who performs better? *Genus* 77(1):31
- D’Agostino A, Ghellini G, Lombardi G (2022) Gender effect at the beginning of higher education careers in STEM studies: Does female recover better than male? In: *Methodology for Multilevel Modeling in Educational Research: Concepts and Applications*. Springer, p 363–382
- Eom SB, Wen HJ, Ashill N (2006) The determinants of students’ perceived learning outcomes and satisfaction in university online education: An empirical investigation. *Decision Sciences Journal of Innovative Education* 4(2):215–235
- EUROSTAT (2024) Increase in online education in the EU in 2023. <https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20240124-2> [Accessed: 2025-07-24]
- EUROSTAT (2025a) Educational attainment statistics. URL [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Educational\\_attainment\\_statistics](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Educational_attainment_statistics)
- EUROSTAT (2025b) Tertiary education statistics. URL [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Tertiary\\_education\\_statistics](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Tertiary_education_statistics)

- Gavosto A (2023) Rapporto anvur: la fotografia degli iscritti all'università. lavoceinfo <https://lavoce.info/archives/101835/rapporto-anvur-la-fotografia/-degli-iscritti-alluniversita/>, Retrieved on November 11, 2024
- Ghignoni E (2017) Family background and university dropouts during the crisis: The case of Italy. *Higher Education* 73(1):127–151
- Gillett-Swan J (2017) The challenges of online learning: Supporting and engaging the isolated learner. *Journal of Learning Design* 10(1):20–30
- Harasim L (2000) Shift happens: Online education as a new paradigm in learning. *The Internet and higher education* 3(1-2):41–61
- Harasim L (2006) A history of e-learning: Shift happened. In: *The international handbook of virtual learning environments*. Springer, p 59–94
- Hung M, Chou C, Chen C, et al (2010) Learner readiness for online learning: Scale development and student perceptions. *Computers & Education* 55:1080–1090
- Impicciatore R, Tosi F (2019) Student mobility in Italy: The increasing role of family background during the expansion of higher education supply. *Res Soc Strat Mobil* 62:100409. <https://doi.org/10.1016/j.rssm.2019.100409>
- Joo YJ, Lim KY, Kim EK (2011) Online university students' satisfaction and persistence: Examining perceived level of presence, usefulness and ease of use as predictors in a structural model. *Computers & Education* 57(2):1654–1664
- Kang K, Wang S (2018) Analyze and predict student dropout from online programs. In: *Proceedings of the 2nd International Conference on Compute and Data Analysis*, pp 6–12
- Kumar A, Kumar P, Palvia SCJ, et al (2017) Online education worldwide: Current status and emerging trends. *Journal of Information Technology Case and Application Research* 19(1):3–9
- Lee Y, Choi J (2011) A review of online course dropout research: Implications for practice and future research. *Educational Technology Research and Development* 59:593–618
- Lombardi G, Ghellini G, et al (2019) The effect of grading policies on Italian universities' attractiveness: A conditional multinomial logit approach. *Electronic Journal of Applied Statistical Analysis* 12(04):801–825

- Marra A (2024) Università telematiche: una partita aperta. Rivista Il Mulino URL <https://www.rivistailmulino.it/a/universit-telematiche-una-partita-aperta>, <https://www.rivistailmulino.it/a/universit-telematiche-una-partita-aperta>, Retrieved on November 10, 2024
- Miccoli P (2023) Università telematiche: un successo annunciato? lavoceinfo <https://lavoce.info/archives/101638/sale-in-cattedra-luniversita-telematica/>, Retrieved on November 11, 2024
- Minerva T, De Santis A, Bellini C, et al (2024) Students in italian online universities: enrollments time series analysis from 2005 to 2021. JE-LKS JOURNAL OF E-LEARNING AND KNOWLEDGE SOCIETY 19(4):50–75
- Palmisano F (2023) Sale in cattedra l'università telematica. lavoceinfo <https://lavoce.info/archives/101638/sale-in-cattedra-luniversita-telematica/>, Retrieved on November 11, 2024
- Palvia S, Aeron P, Gupta P, et al (2018) Online education: Worldwide status, challenges, trends, and implications. Journal of global information technology management 21(4):233–241
- Perla L (2024) La qualità (perduta) della formazione dello studente. Rivista Il Mulino URL <https://www.rivistailmulino.it/a/la-qualit-perduta-della-formazione-dello-studente>, <https://www.rivistailmulino.it/a/la-qualit-perduta-della-formazione-dello-studente>, Retrieved on July 07, 2025
- Piromalli L (2024) Making markets in higher education: State reform, global edtech, and virtual universities in italy. Italian Journal of Sociology of Education 2(16):109–130
- Pozzi F, Conole G (2014) What is the future for moocs in italy? Italian Journal of Educational Technology 22(3):173–182
- Pozzi F, Manganello F, Passarelli M, et al (2019) Ranking meets distance education: Defining relevant criteria and indicators for online universities. International Review of Research in Open and Distributed Learning 20(5):42–63
- Prenkaj B, Velardi P, Stilo G, et al (2020) A survey of machine learning approaches for student dropout prediction in online courses. ACM Computing Surveys (CSUR) 53(3):1–34
- Priulla A, Attanasio M (2024) Unveiling gender disparities in university pathways: insights from Italy's master's level. European Journal of Higher Education 14(4):574–597
- Priulla A, Miaci E, Panichella N (2025) Social selection in student mobility: The interplay of origin, achievement and track in the italian north-south divide. Research in Social Stratification and

Mobility p 101105

Rahmani AM, Groot W, Rahmani H (2024) Dropout in online higher education: a systematic literature review. *International Journal of Educational Technology in Higher Education* 21(1):19

Robeyns I (2006) Three models of education: Rights, capabilities and human capital. *Theory and research in education* 4(1):69–84

Rosenbaum PR (2002) Observational studies. In: *Observational studies*. Springer, p 1–17

Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1):41–55

Sánchez-Gelabert A, Valente R, Duarte JM (2020) Profiles of online students and the impact of their university experience. *International Review of Research in Open and Distributed Learning* 21(3):230–249

Sen AK (1997) Human capital and human capability. *World development* 25(12):1959–1961

Siemens G, Gasevic D, Dawson S (2015) Preparing for the digital university: A review of the history and current state of distance, blended and online learning [https://researchmgt.monash.edu/ws/portalfiles/portal/256525723/256524746\\_oa.pdf](https://researchmgt.monash.edu/ws/portalfiles/portal/256525723/256524746_oa.pdf), Retrieved on March 27, 2025

Simone R (2024) Il business opaco delle università telematiche: il topolino diventato una montagna. URL <https://www.editorialedomani.it/fatti/il-business-opaco-delle-universita-telematiche-il-topolino-diventato-una-montagna-snbe3fsy>

Tinto V (1975) Dropout from higher education: A theoretical synthesis of recent research. *Review of educational research* 45(1):89–125

Tocchioni V, Milone S, Lombardi G, et al (2025) Gender disparities in school-to-university transition in Italy: the role played by the socio-economic condition and the type of high school. *RIVISTA ITALIANA DI ECONOMIA, DEMOGRAFIA E STATISTICA* pp 67–78

Usala C, Sulis I, Porcu M (2025) Exploring the impact of high schools, socioeconomic factors, and degree programs on higher education success in Italy. *Big Data Research* p 100539

De la Varre C, Irvin MJ, Jordan AW, et al (2014) Reasons for student dropout in an online course in a rural k–12 setting. *Distance Education* 35(3):324–344

Weber M (1978) *Economy and society: An outline of interpretive sociology*, vol 2. University of California press

Willing PA, Johnson SD (2009) Factors that influence students' decision to dropout of online courses. *Journal of Asynchronous Learning Networks* 13(3):115–127

Zaki MS (2022) Advantages and disadvantages of online learning. *Journal of International Social Research* 15(92)

Accepted manuscript

## Appendix A Robustness and sensitivity analysis

This section assesses the robustness of our main findings through a series of complementary analyses. First, we verify whether our results are sensitive to the choice of estimation method by replicating the analysis using Inverse Probability Weighting (IPW) as an alternative to the nearest-neighbour matching algorithm used in the main text. Second, we conduct a Rosenbaum bounds analysis to evaluate the sensitivity of our findings to potential unobserved confounding. Third, we implement an instrumental variable approach as an additional strategy to address the endogeneity of online university enrolment. Taken together, these analyses provide further support for the validity and robustness of our main results.

### A.1 Alternative estimation strategy: Inverse Probability Weighting

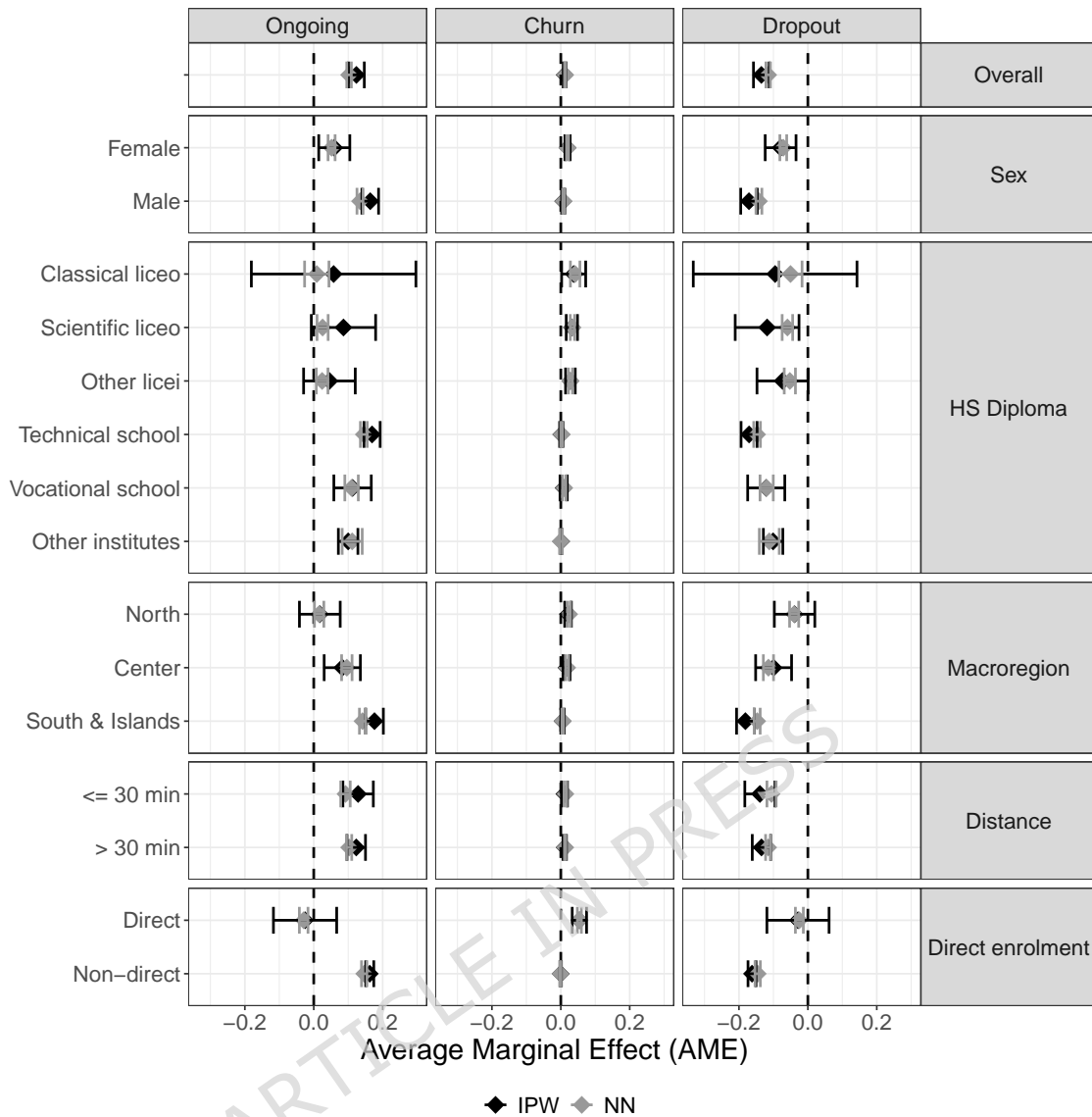
As a first robustness check, we replicate our analysis using Inverse Probability Weighting (IPW) as an alternative to the nearest neighbour matching algorithm. Rather than matching each treated unit to a comparable control, IPW reweights observations by the inverse of their estimated propensity score, thereby assigning greater weight to observations that are closer to the treated units and constructing a pseudo-population in which the distribution of observed covariates is balanced across treated and control units.

Figure A1 compares the AMEs obtained with the two methods. The results are virtually identical across all outcomes and subgroups, confirming that our findings are not sensitive to the choice of the estimation algorithm. The main difference concerns the precision of the estimated effect. Indeed, IPW estimates results in a wider confidence interval with respect to the estimation strategy used in the main paper. Overall, this robustness check provides further support for the validity of our main conclusions.

### A.2 Sensitivity to unobserved confounding: Rosenbaum Bounds

As explained in Section 4.2, our results rely on the conditional independence assumption, which requires that, conditional on the observed covariates, treatment assignment is independent of potential outcomes. While we argued in Section 4.2 that this assumption is plausible given the richness of the MOBYSU.IT database, we cannot rule out the presence of unobserved confounders. To assess the sensitivity of our results to this concern, we conduct a Rosenbaum bounds analysis (Rosenbaum 2002; Becker and Caliendo 2007).

The analysis is based on the matched pairs produced by the PSM procedure, each consisting of one online and one traditional student who are similar with respect to all observable characteristics. If our identifying assumption holds, two students who are identical on all observables should have the



**Fig. A1:** Average Marginal Effects: Inverse Probability Weighting vs Nearest Neighbour

*Notes:* The figure reports the Average Marginal Effects estimated using Inverse Probability Weighting (IPW) and Nearest Neighbour (NN) matching. Each AME measures the change in the probability of observing the outcome due to online university attendance. Results are reported for the overall population and each subgroup.

same probability of enrolling in an online university. However, if an unobserved confounder exists, this may no longer be the case. The Rosenbaum bounds analysis asks: how large would this unobserved confounder need to be before our conclusions would no longer hold?

The size of this unobserved confounder is captured by the parameter  $\gamma$ . For example, a value of  $\gamma = 1$  means that no hidden bias exists, i.e., two observationally identical students have the same probability of treatment. Instead, a value of  $\gamma = 2$  would indicate that an unobserved factor could make one student twice as likely to enrol in an online university as the other. For each value of  $\gamma$ , we compute two p-values based on the Wilcoxon signed-rank test applied to the discordant matched pairs (pairs in which online and traditional students differ in their outcome). The null hypothesis of

the test is that online attendance does not affect students' outcomes. That is, among discordant pairs, it is equally likely that the treated or the control student has the worse outcome. The two p-values correspond to two opposite assumptions about the direction in which the unobserved confounder may affect our results. The lower bound p-value represents the best-case scenario: imagine that the unobserved confounder is being a full-time worker. Students who work full-time are more likely to enrol in an online university but also more likely to drop out, as they have less time to dedicate to their studies. In this case, the confounder reinforces our estimates, as online universities help precisely those students who face the greatest obstacles, and the true effect of online attendance would be even larger than our estimate. The upper bound p-value represents the worst-case scenario: imagine that the unobserved confounder is intrinsic motivation and that it is positively related to the probability of choosing an online university. Then, students who are more motivated are more likely to choose an online university and less likely to drop out, regardless of the type of institution they attend. In this case, our estimated effect would capture both the effect of online attendance and that of students' motivation. In this context, the larger the value of  $\gamma$  at which the upper bound p-value remains below 0.05, the more confident we can be that our results are not driven by unobserved confounding.

Since our outcome of interest is multinomial, we apply the analysis separately to each of the three binary outcomes. Table A1 reports the results.

**Table A1:** Rosenbaum bounds sensitivity analysis.

$\gamma$	Ongoing		Churn		Dropout	
	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
1	0	0	0	0	0	0
1.1	0	0	0	0	0	0
1.2	0	0	0	0	0	0
1.3	0	0	0	0	0	0
1.4	0	0	0	0	0	0
1.5	0	0	0	0	0	0
1.6	0	0	0	0	0	0
1.7	0	0.002	0	0	0	0
1.8	0	0.840	0	0	0	0
1.9	0	1	0	0	0	0.001
2	0	1	0	0	0	0.652
2.1	0	1	0	0	0	0.999
2.2	0	1	0	0.002	0	1
2.3	0	1	0	0.019	0	1
2.4	0	1	0	0.097	0	1

*Notes:* The table reports the lower and upper bound p-values of the Rosenbaum bounds sensitivity analysis for each value of  $\gamma$  and each outcome. The p-values are derived from the Wilcoxon signed-rank test applied to the discordant matched pairs. For each value of  $\gamma$ , the test evaluates whether the observed distribution of discordant pairs is consistent with the null hypothesis of no treatment effect, under the assumption that an unobserved confounder of size  $\gamma$  exists.

The results for the lower bound, as expected, equal zero for all values of  $\gamma$  and all outcomes, indicating that under the most favourable scenario, the results remain valid. Instead, the upper bound p-values suggest that the estimated effects remain significant at the 5% level up to  $\gamma = 1.7$  for ongoing,  $\gamma = 2.3$  for churn, and  $\gamma = 1.9$  for dropout. Taking the most conservative threshold, an unobserved confounder would need to make one student in a matched pair at least 1.7 times more likely to enrol in an online university than the other before invalidating our conclusions. While these results cannot rule out the presence of unobserved confounding entirely, they provide reassuring evidence that our findings are unlikely to be driven by hidden bias alone.

Accepted Article  
ARTICLE IN PRESS  
manuscript

## Appendix B Matching results

Table B2 presents the results of the matching procedure, both for the overall sample and for the various subgroups defined by the stratifying variables. The most important element is given by the percentage of treated units for which no suitable control could be identified (unmatched). As we can notice, this proportion greatly depends on the specific group considered.

**Table B2:** Results of the matching procedure by subgroups of students.

Subgroup	Controls	Treated			
	All N	All N	Matched N	Unmatched N	%
<b>Overall</b>	1,741,184	61,826	53,587	8,239	13.33
<b>Direct enrolment</b>					
Non-direct	227,509	48,692	40,072	8,620	17.70
Direct	1,513,675	13,134	12,918	216	1.64
<b>HS Track</b>					
Classical liceo	236,864	2,136	1,825	311	14.56
Scientific liceo	630,856	8,454	7,716	738	8.73
Other licei	355,699	8,382	7,779	603	7.19
Technical school	382,525	30,093	24,747	5,346	17.76
Vocational school	114,038	7,314	6,579	735	10.05
Other institutes	21,202	5,447	2,617	2,830	51.96
<b>Sex</b>					
Female	958,661	22,981	21,709	1,272	5.54
Male	782,523	38,845	31,392	7,453	19.19
<b>Macro-region</b>					
North	702,543	14,365	12,677	1,688	11.75
Center	349,805	12,356	10,906	1,450	11.74
South & Islands	688,836	35,105	28,854	6,251	17.81
<b>Distance</b>					
> 30 min	326,113	14,143	12,006	2,137	15.11
≤ 30 min	1,415,071	47,683	41,385	6,298	13.21

*Notes:* The table reports the results of the matching procedure with respect to the number of treated students successfully matched. For each subgroup of students, we show the number of potential controls, treated units, matched treated units, and the number and percentage of unmatched treated units.

The overall matching procedure, carried out using the data on all units, results in a loss of 13.33% of treated students. This value is higher when considering students with delays between the high school diploma and enrolment (17.7%), whereas it is much lower for direct enrolments (1.64%). With respect to high school tracks, we lose more than one-half of all students enrolled in other institutes (51.96%). Further difficulties in identifying suitable matches are observed among male students (19.19%) and those who obtained their high school diploma in Southern Italy (17.81%).

Table B3 reports the characteristics of unmatched treated units. These students tend to be considerably older, are almost exclusively enrolled on a non-direct basis, and are more concentrated in specific fields of study, such as Engineering and Services. These elements suggest that these students have profiles for which no comparable counterpart exists among traditional university students.

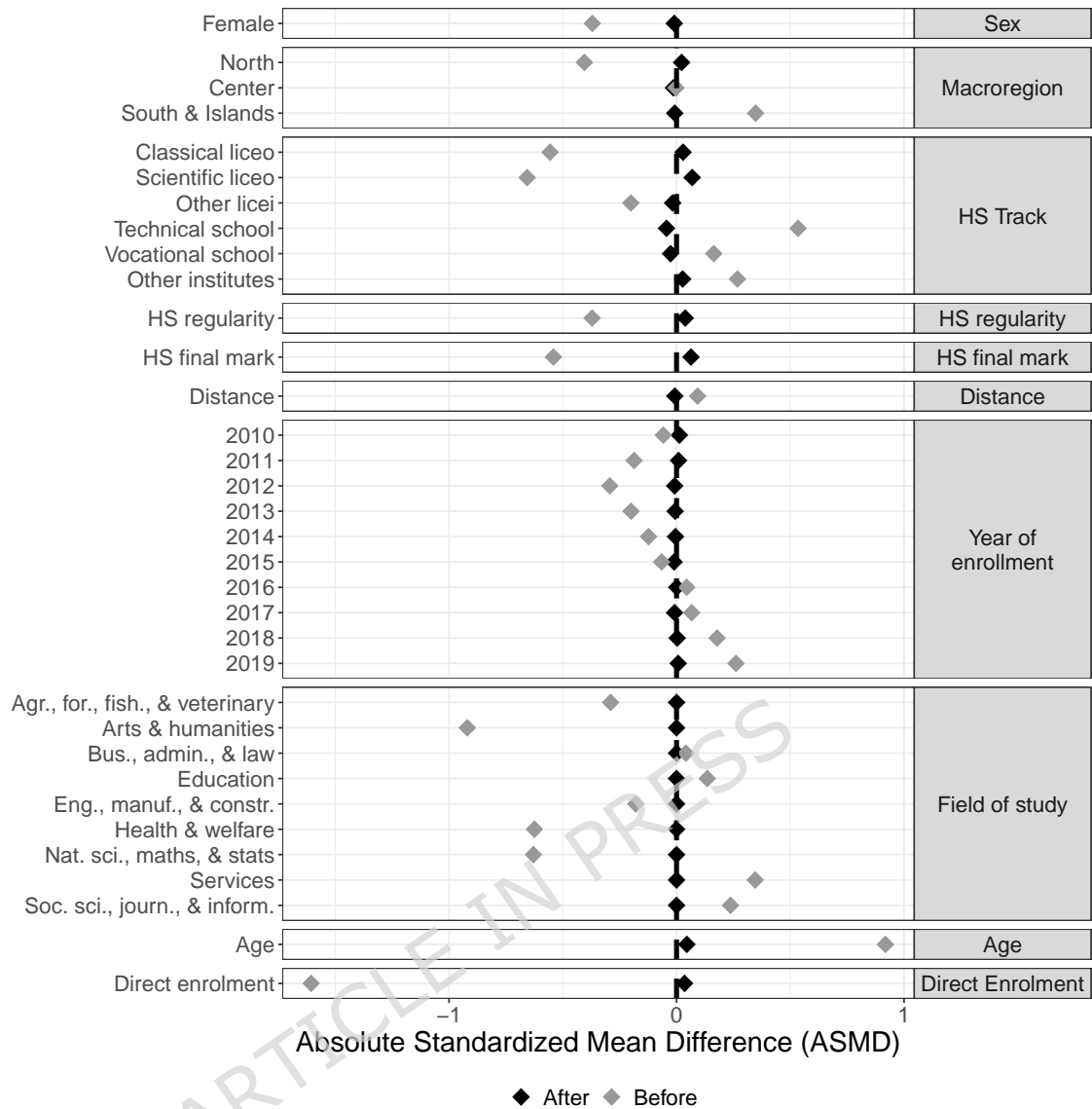
**Table B3:** Comparison between matched and unmatched students

<b>Variables</b>	<b>Unmatched</b>	<b>Matched</b>
<b>No. observations</b>	8239	53587
<b><i>HS track</i></b>		
Classic	1,0%	3,8%
Scientific	3,6%	15,2%
Other Licei	4,6%	14,9%
Technical institutes	62,4%	46,6%
Vocational institutes	8,5%	12,3%
Other institutes	20,0%	7,1%
<b><i>Age</i></b>		
(mean)	41,0	28,1
<b><i>HS final mark</i></b>		
(mean)	71,2	73,5
<b><i>Sex</i></b>		
Female	19,9%	39,8%
<b><i>Field of study</i></b>		
Agr., for., fish., & veterinary	0,1%	0,9%
Arts & humanities	0,0%	3,3%
Bus., admin., & law	33,9%	27,3%
Education	1,3%	9,5%
Health & welfare	-	0,1%
Eng., manuf., & constr.	30,7%	10,4%
Nat. sci., maths, & stats	0,0%	0,4%
Services	18,3%	17,2%
Soc. sci., journ., & inform.	15,7%	30,8%
<b><i>macro-region</i></b>		
North	15,9%	24,4%
Center	18,4%	20,2%
South & Islands	65,7%	55,4%
<b><i>HS regularity</i></b>		
Regular	56,7%	68,8%
<b><i>Direct enrolment</i></b>		
Direct	0,3%	24,5%
<b><i>Minimum distance from university</i></b>		
> 30	76,5%	77,2%

*Notes:* The table reports the average characteristics of matched and unmatched students. Unmatched students are those for whom no suitable control unit could be identified within the imposed caliper of 0.05 on the propensity score scale. Matched students are those for whom a suitable control was found among traditional university students.

Regarding the balance between the covariates of treated and control units, Figure B2 shows the absolute standardised mean difference (ASMD) between the two student groups, before and after matching. The ASMD is defined as the absolute difference between the weighted means of the treatment and control groups, divided by the unweighted standard deviation of the pooled sample. It allows us to investigate the residual differences between the groups after the matching procedure by using the same unit of measure across the various covariates.

The ASMDs before the PSM indicate a strong imbalance between the two groups, providing evidence in favour of the application of a balancing procedure. In detail, most of the unbalancing was due to the age at enrolment, which was substantially higher among online students, and the proportion of direct enrolments. Other important differences are found regarding students' sex, with male participation being higher in online universities than females, field of study, and high school



**Fig. B2:** Absolute standardised mean differences before and after the matching.

*Notes:* For each covariate, the figure shows the absolute standardised mean difference between online and traditional students before and after the PSM procedure.

background. Indeed, online universities have a larger share of students with a non-*liceo* background, who are non-regular during high school and have a lower high school final mark.

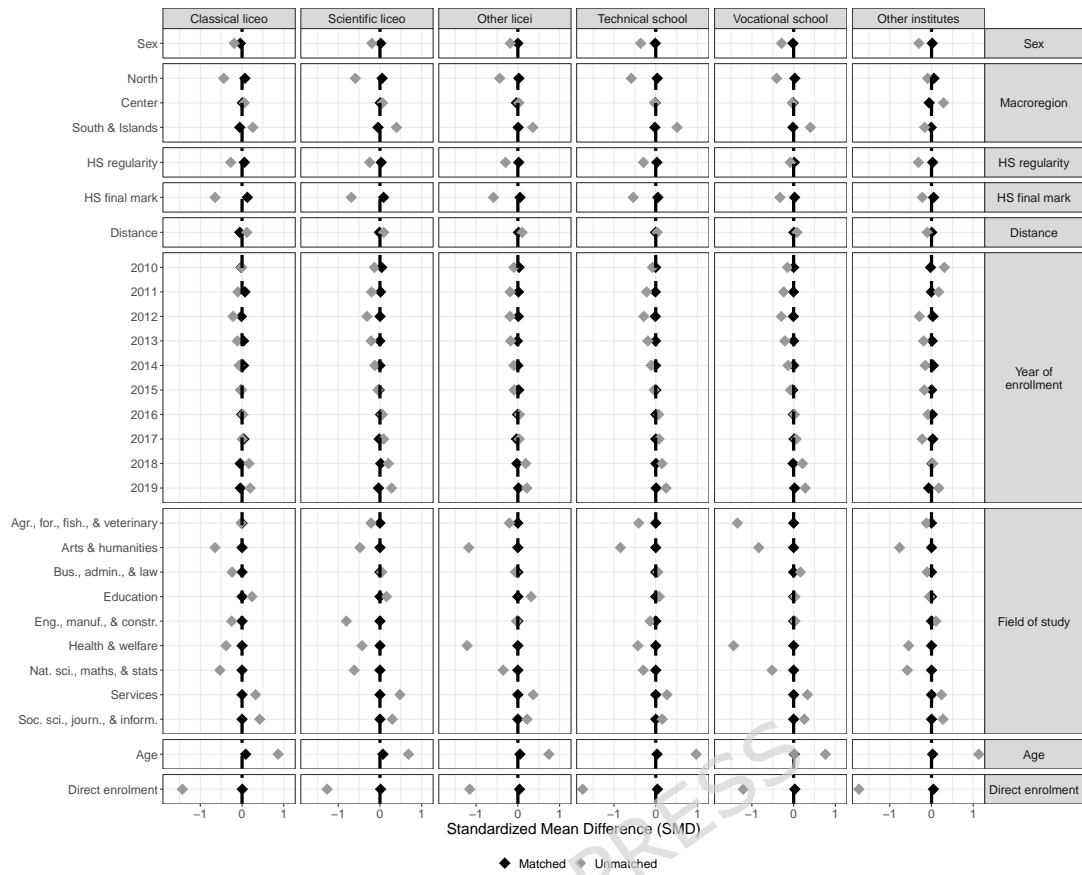
Nonetheless, the ASMD after the matching procedure shows very promising results, with a very good balance achieved between the two groups, with the only exception being the high school track. However, we adjust for these differences in the following sections by analysing the effect of online attendance among students who attended the same high school track.

Similar results are obtained looking at the different subgroups defined according to students' high school track, sex, and distance from the nearest university, as shown in Figure B3.

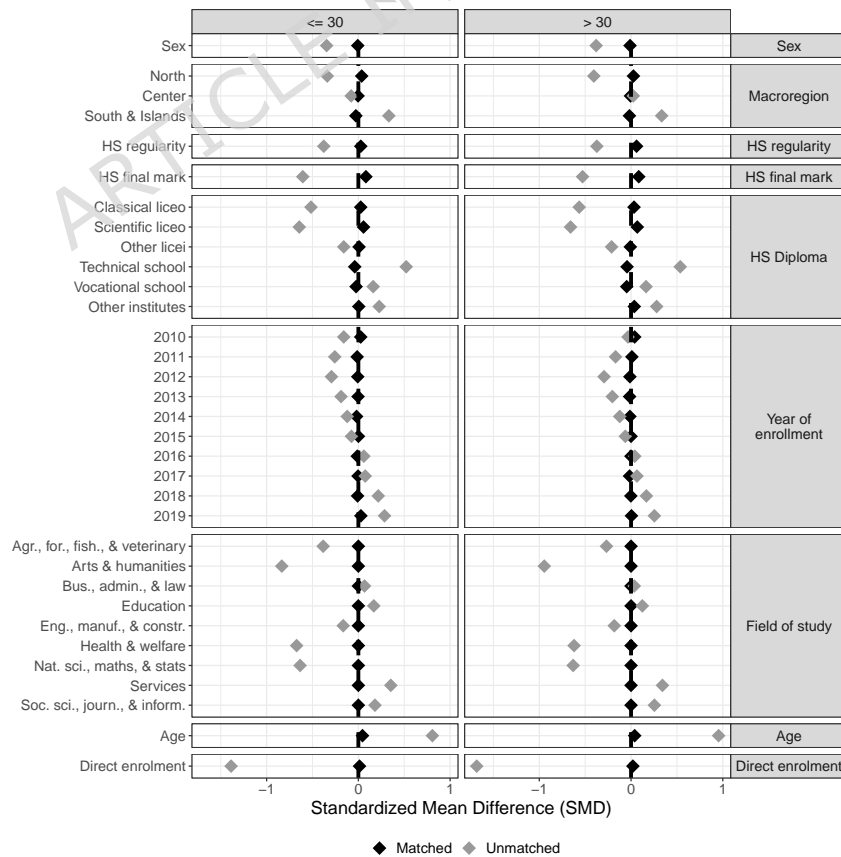
Finally, Figure B4 reports the covariate balance before and after matching separately for *direct* and *non-direct* students, while Figures B5 and B6 show, respectively, the absolute mean differences

computed before and after the matching for the subgroups of students defined according to their macro-region, sex, distance from the nearest university, and high school track. Overall, the results confirm that a satisfactory level of balance is achieved in all subgroups.

ARTICLE IN PRESS  
Accepted Manuscript



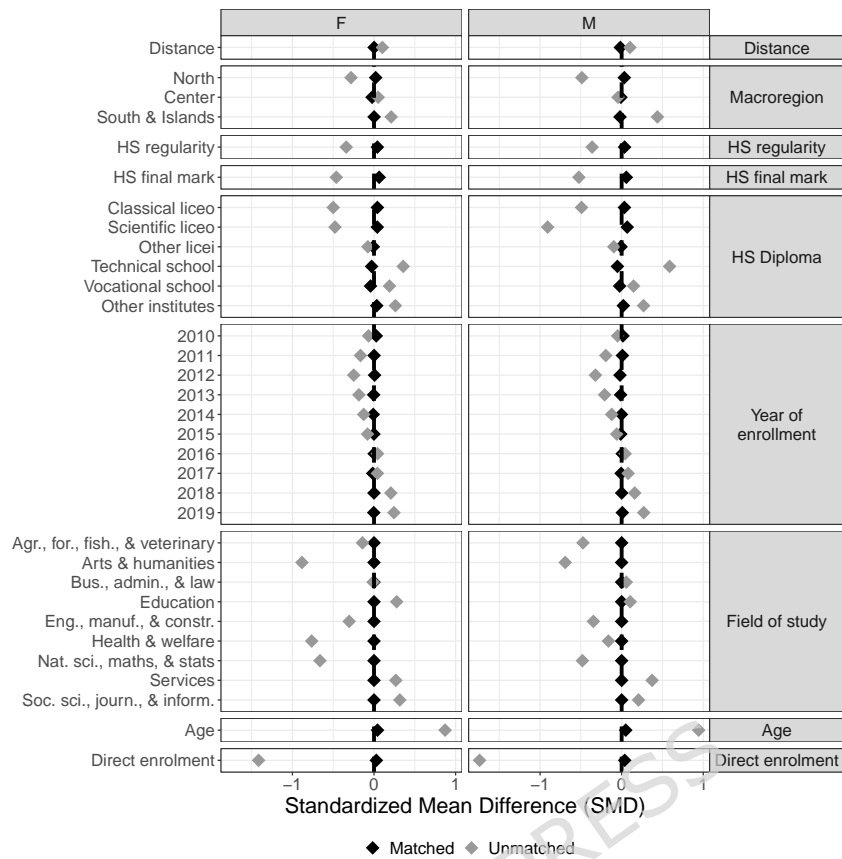
(a) High school track



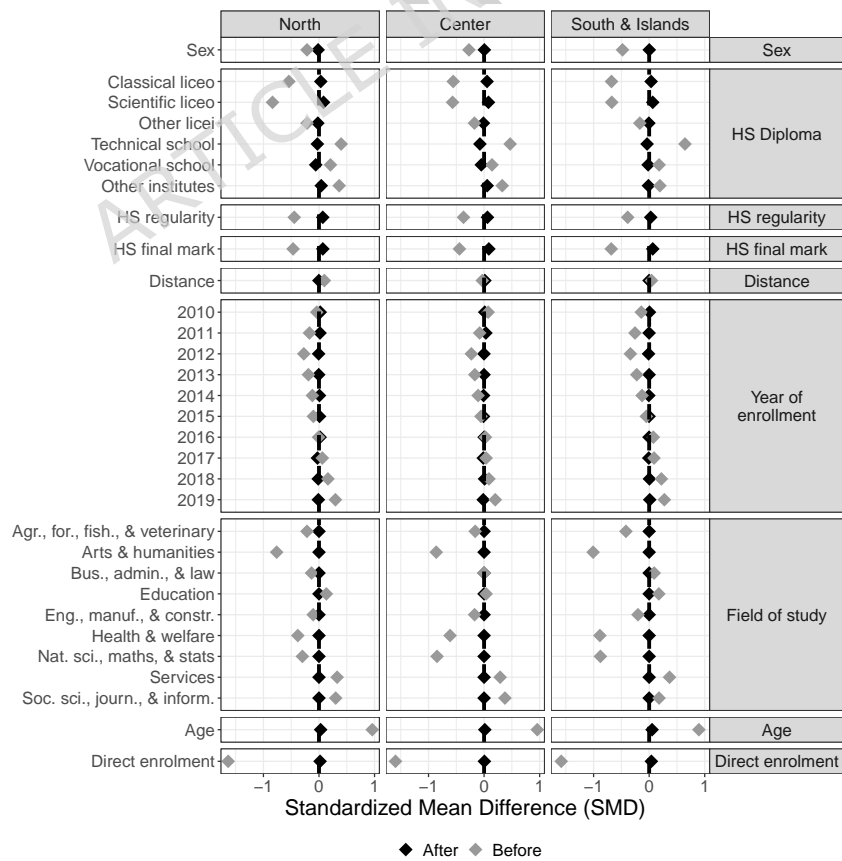
(b) Distance from the nearest university

**Fig. B3:** Absolute standardised mean differences before and after the matching; subgroups

*Notes:* For each covariate, the figure shows the absolute standardised mean difference between online and traditional students before and after the PSM procedure with respect to each considered subgroup.



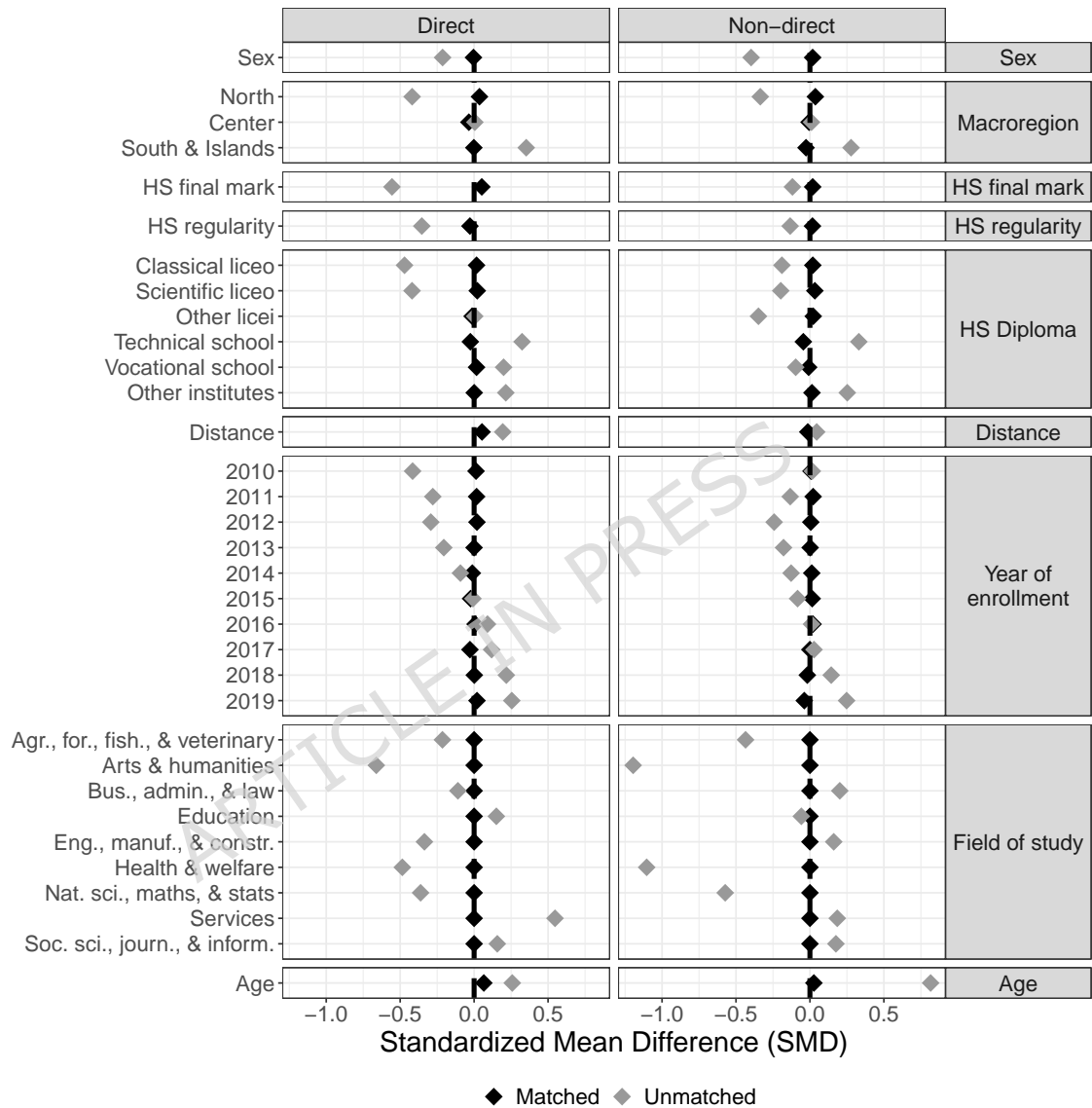
(c) Sex



(d) Macro region

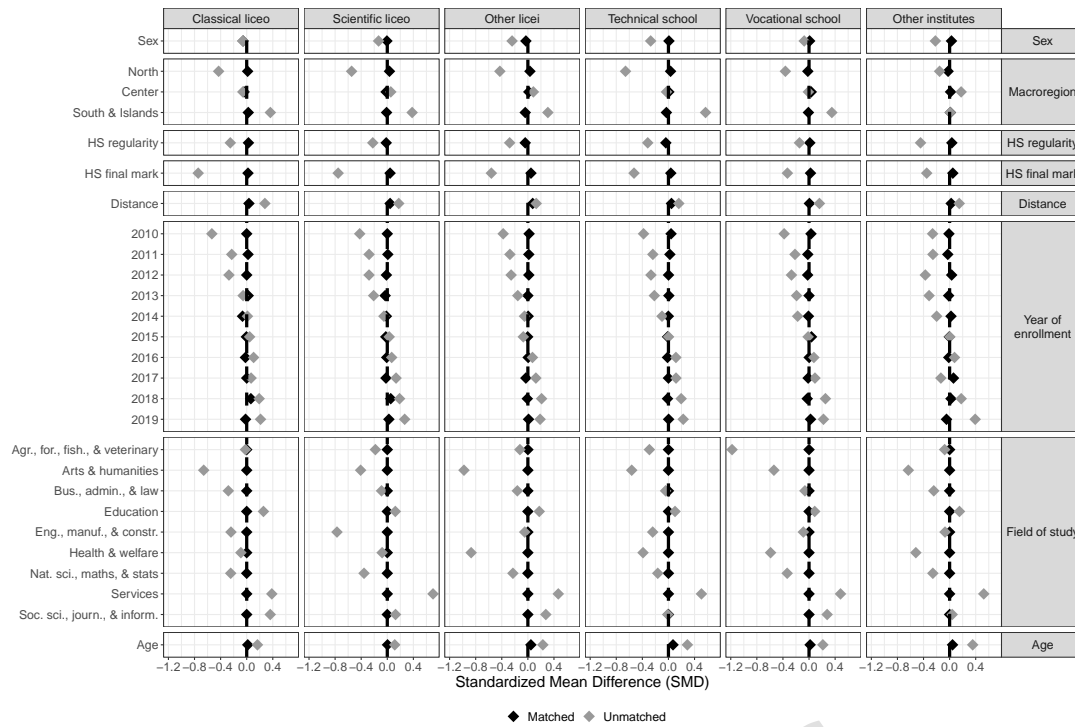
**Fig. B3:** Absolute standardised mean differences before and after the matching: subgroups (continued)

*Notes:* For each covariate, the figure shows the absolute standardised mean difference between online and traditional students before and after the PSM procedure with respect to each considered subgroup.

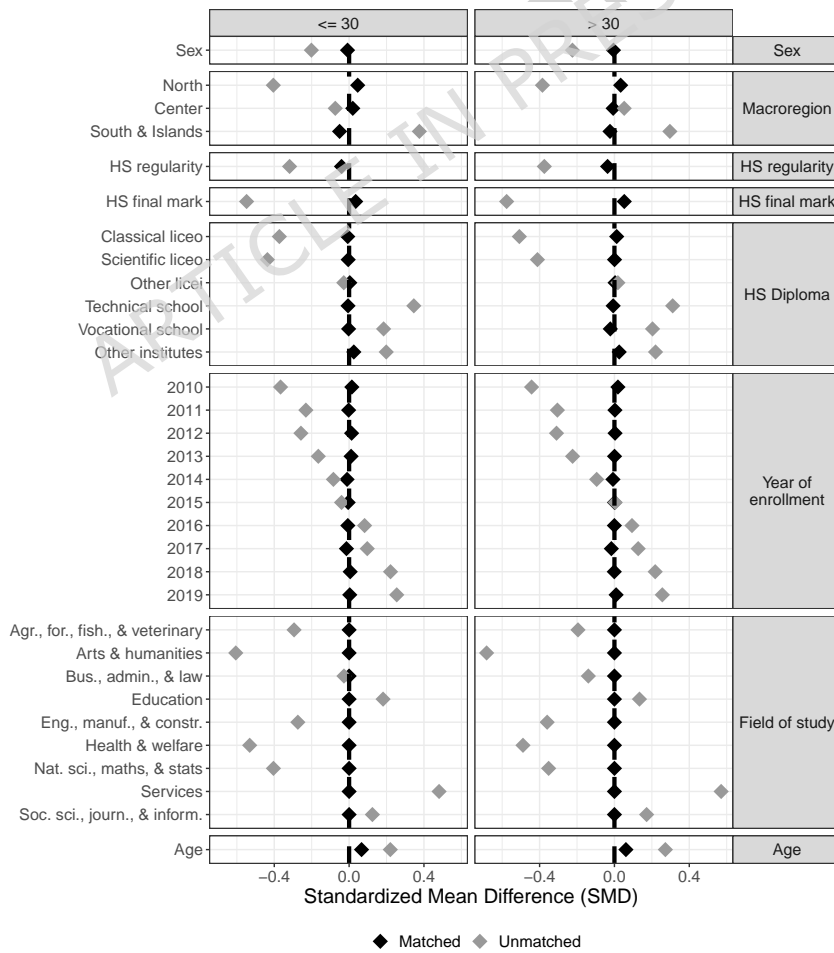


**Fig. B4:** Absolute standardised mean differences before and after the matching: direct vs non-direct

*Notes:* For each covariate, the figure shows the absolute standardised mean difference between online and traditional students before and after the PSM procedure with respect to *direct* and *non-direct* enrolled.



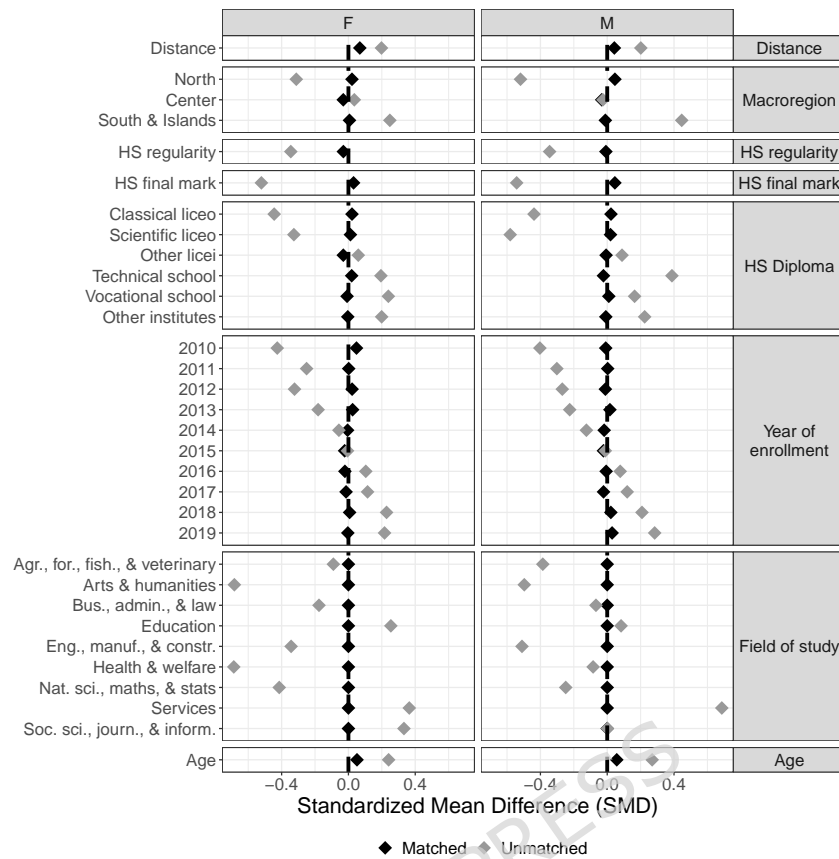
(a) High school track



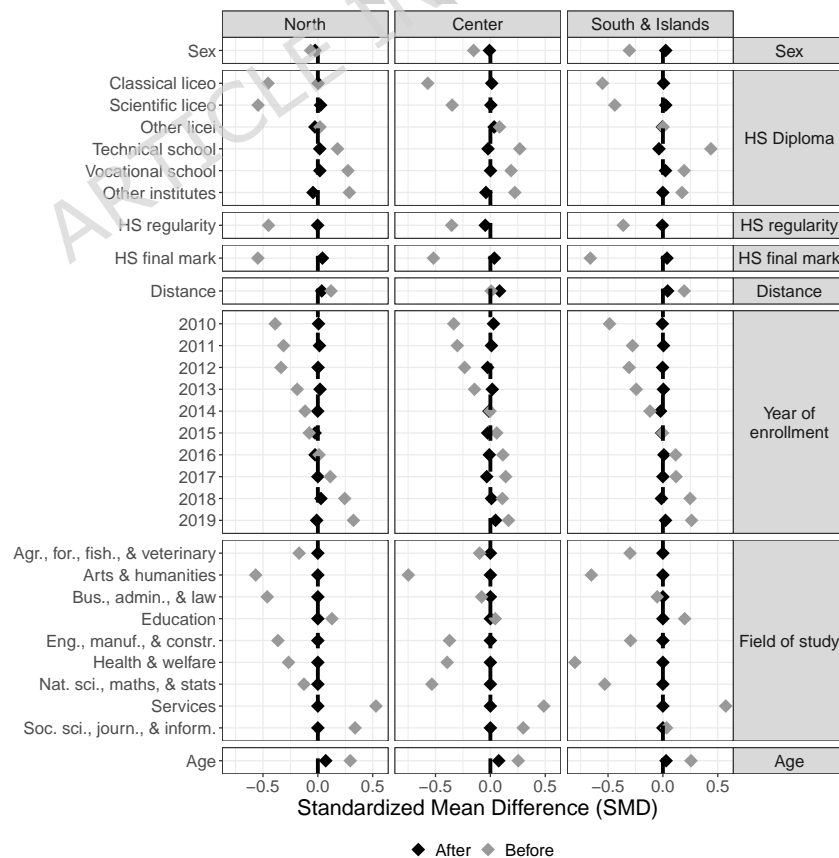
(b) Distance from the nearest university

**Fig. B5:** Absolute standardised mean differences before and after the matching: subgroups of direct students

*Notes:* For each covariate, the figure shows the absolute standardised mean difference between online and traditional students before and after the PSM procedure with respect to each considered subgroup.



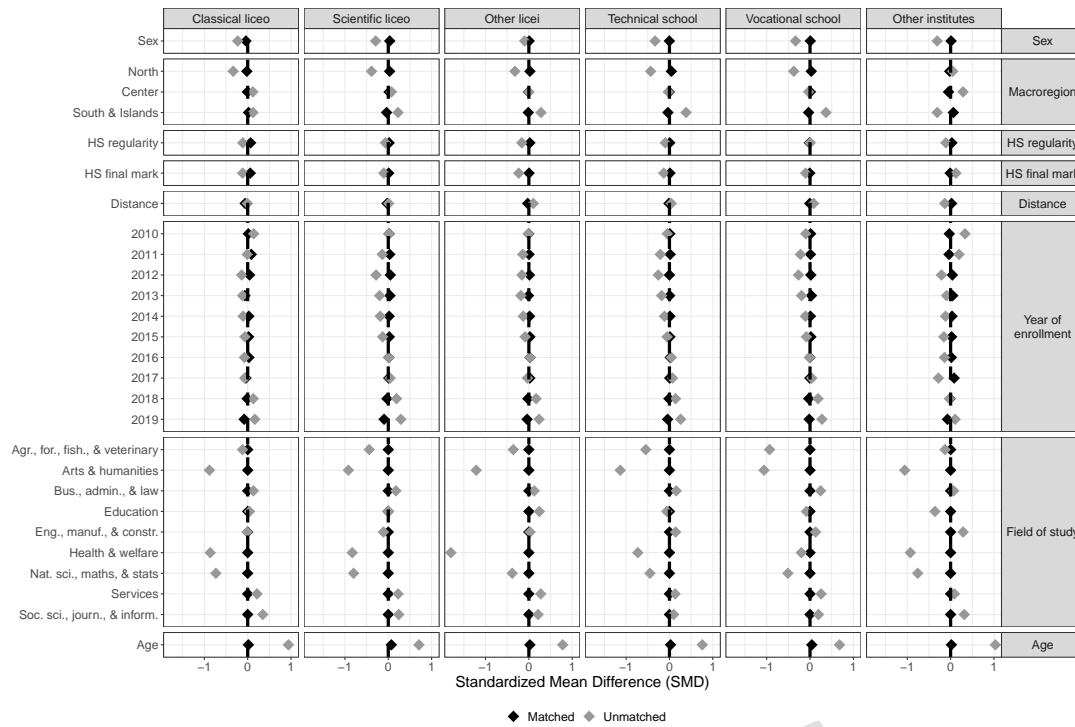
(c) Sex



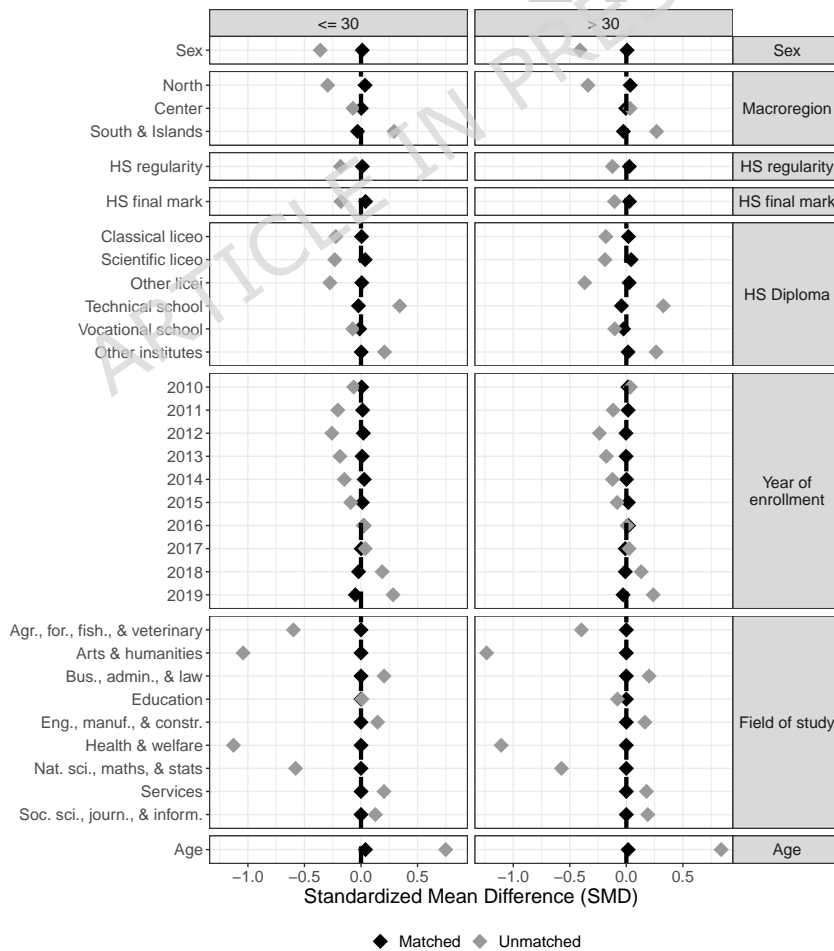
(d) Macro region

**Fig. B5:** Absolute standardised mean differences before and after the matching: subgroups of direct students (continued)

*Notes:* For each covariate, the figure shows the absolute standardised mean difference between online and traditional students before and after the PSM procedure with respect to each considered subgroup.



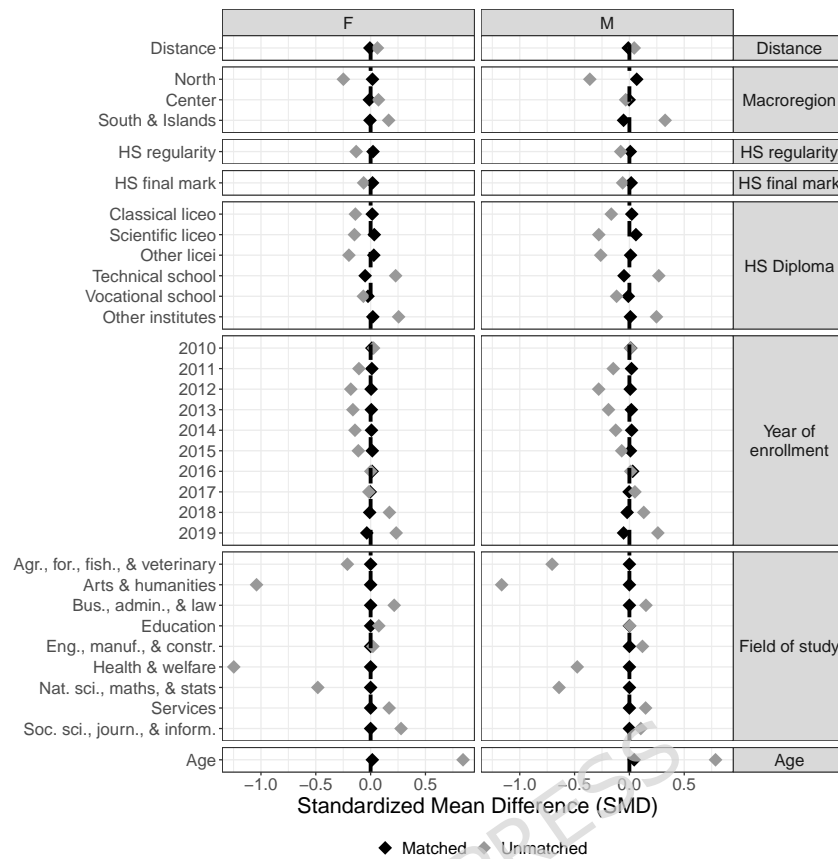
(a) High school track



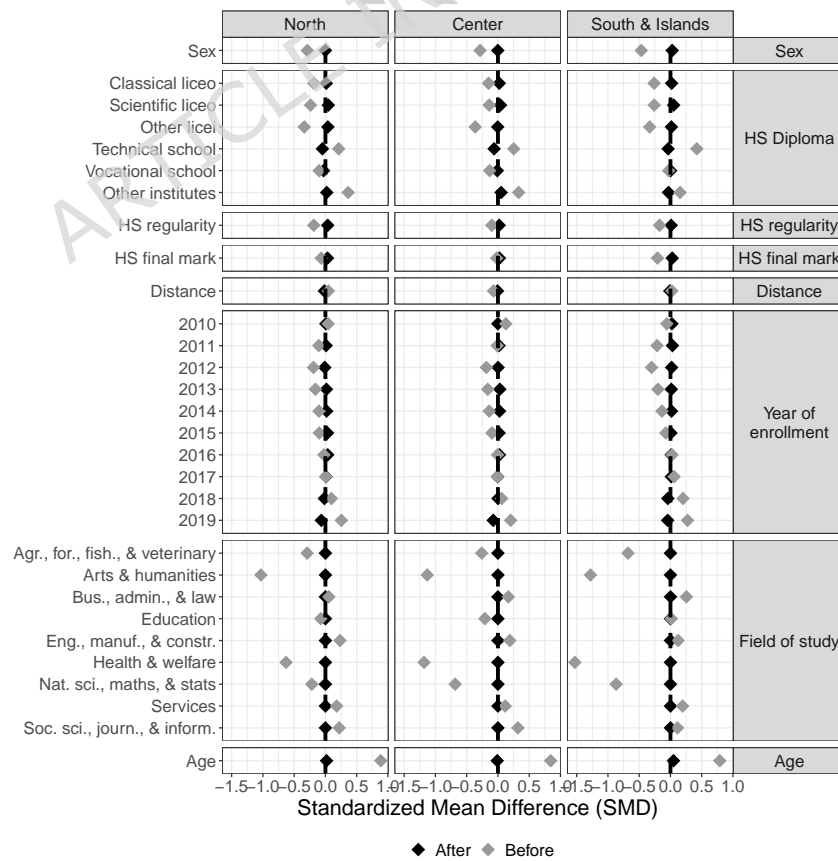
(b) Distance from the nearest university

**Fig. B6:** Absolute standardised mean differences before and after the matching: subgroups of non-direct students

*Notes:* For each covariate, the figure shows the absolute standardised mean difference between online and traditional students before and after the PSM procedure with respect to each considered subgroup.



(c) Sex



(d) Macro region

**Fig. B6:** Absolute standardised mean differences before and after the matching: subgroups of non-direct students (continued)

*Notes:* For each covariate, the figure shows the absolute standardised mean difference between online and traditional students before and after the PSM procedure with respect to each considered subgroup.

## Appendix C Details on average marginal effects

Table C4: Average Marginal Effects.

Subgroup	Overall			Direct			Non-direct		
	Ongoing	Churn	Dropout	Ongoing	Churn	Dropout	Ongoing	Churn	Dropout
<i>Overall</i>	0,10	0,01	-0,12	-0,03	0,05	-0,02	0,15	0,00	-0,15
<i>HS Track</i>									
Classical liceo	0,01	0,04	-0,05	-0,14	0,10	0,04	0,11	0,01	-0,12
Scientific liceo	0,03	0,03	-0,06	-0,10	0,08	0,02	0,09	0,00	-0,10
Other licei	0,02	0,03	-0,05	-0,09	0,08	0,01	0,08	0,01	-0,09
Technical school	0,15	0,00	-0,15	0,03	0,03	-0,06	0,17	0,00	-0,16
Vocational school	0,11	0,01	-0,12	0,01	0,04	-0,05	0,13	0,00	-0,13
Other institutes	0,11	0,00	-0,11	0,04	0,01	-0,05	0,15	0,00	-0,15
<i>Macro-region</i>									
North	0,02	0,02	-0,04	-0,13	0,09	0,04	0,06	0,00	-0,06
Center	0,10	0,02	-0,11	-0,07	0,07	0,00	0,14	0,00	-0,14
South & Islands	0,14	0,01	-0,15	0,02	0,03	-0,06	0,17	0,00	-0,17
<i>Sex</i>									
Female	0,05	0,02	-0,07	-0,08	0,07	0,01	0,10	0,00	-0,10
Male	0,13	0,01	-0,14	0,01	0,04	-0,05	0,16	0,00	-0,16
<i>Distance</i>									
> 30 min	0,09	0,01	-0,11	-0,03	0,04	-0,01	0,14	0,00	-0,14
≤ 30 min	0,10	0,01	-0,12	-0,04	0,06	-0,02	0,14	0,00	-0,14
<i>Direct enrolment</i>									
Direct	-0,03	0,05	-0,02	-	-	-	-	-	-
Non-direct	0,15	0,00	-0,15	-	-	-	-	-	-

The table shows the causal estimates of the effect of online attendance in terms of AMEs. Each AME measures the change in the probability of observing the outcome due to the treatment within each subgroup. The table reports the results considering the overall population (cols. 2-4), students enrolled directly after completing the high school (cols. 5-7), and those enrolled after one year or more (Cols. 8-10). The results regarding the coefficients estimated with the logit model are available from the authors upon request.