Article:

Barra, C., Destefanis,, S., Sena, V. et al. (1 more author) (2024) Students' performance and faculty efficiency. Assessing the role of gender through a metafrontier. Journal of Economic Studies, 51 (9). pp. 323-337. ISSN 0144-3585

https://doi.org/10.1108/JES-01-2024-0037

Students' performance and faculty efficiency. Assessing the role of gender through a metafrontier

Cristian Barra and Sergio Destefanis
Department of Economics and Statistics, University of Salerno, Fisciano, Italy
Vania Sena
The University of Sheffield, Sheffield, UK, and
Roberto Zotti
Department of Economics and Statistics "Cognetti De Martiis", University of Turin, Turin, Italy

Received 24 January 2024
Revised 29 July 2024
14 August 2024
Accepted 15 August 2024

Abstract

Purpose – This paper provides novel evidence on the role of gender in the performance of university students, which is particularly relevant to the debate on the performance of female students in science, technology, engineering and mathematics (STEM) subjects.

Design/methodology/approach – Our approach relies on the metafrontier approach proposed by Huang *et al.* (2014), which measures students' efficiency within a given faculty and the impact of the faculty's technology on students' efficiency. We use a sample of 53,159 first-year students in 8 faculties from a large university in southern Italy from 2002–2003 to 2010–2011.

Findings – Students' efficiency is relatively low, reflecting an essential role of unobserved heterogeneity. The different technologies of somewhat similar faculties have minimal impact on efficiency. There is a performance gap against women in five faculties, which on average is strongest for the faculties in the pure and applied science area. This gap increases with the proportion of female students and decreases with female lecturers.

Practical implications – The metafrontier has the benefit of providing relevant policy information on the drivers of student success by relying on data that universities routinely generate and preserve.

Originality/value – The stochastic metafrontier approach allows us to separate the group-specific frontiers from the metafrontier, yielding a decomposition of the efficiency scores of various faculties into technical efficiency scores and technological gaps.

Keywords Women in STEM fields, University efficiency, Metafrontier

Paper type Research paper

1. Introduction

In his maiden speech to the upper house Senate, the former Italian Prime Minister Mario Draghi stated that "ensuring gender equality also means making sure that everyone has equal access to training in those key skills useful for achieving increasing career success - digital, technological, and environmental . . . so that more and more young women choose to train in the areas on which we intend to relaunch the country". The need to increase investments in technical colleges in Italy to meet the challenges of a new digitalised and

JEL Classification — C67, I20, I23

We thank the Editor and the anonymous referees for their helpful comments. All authors declare that they have no relevant or material financial interests that relate to the research described in this paper. Data and programs are available from the corresponding author upon request.

Journal of Economic Studies
Vol. 51 No. 9, 2024
pp. 323-337
Emerald Publishing Limited
0144-3585
DOI 10.1108/JES-01-2024-0037

technological era is also clear from the National Recovery and Resilience Plan [1]. Mission 4 – Education and Research - aims to improve the educational system by strengthening science, technology, engineering and mathematics (STEM) and digital skills in all school cycles.

It is well-known that women are under-represented in STEM subjects. The gender parity in the number of graduates is eight positions below the European average (see Plan International and Bocconi University (2021) for a more detailed report that analyses the digital gender gap in Italy). Indeed, men make up 59% of STEM graduates, and the percentage increases to 74% and 68.4% when focusing on Engineering and scientific fields, respectively (Almalaurea, 2018a). Educational choices made at an early age, when students may enrol in high school tracks that are more intensive in maths and science, do matter and possibly explain half of the gender gap in STEM graduation rates (Granato, 2020). Women continue to be underrepresented in STEM fields and the workforce. Although over the last twenty years, the participation of women in higher education has rapidly increased with the result that the gender gap in this ambit has reversed in favour of women (OECD, 2010); there are significant and persistent differences between male participation rates and female participation rates at the subject level. These differences are the main factors behind the uneven distribution of women and men across occupations and the gender pay gap (McNabb *et al.*, 2002; Jacobs, 2003; Bertrand, 2018; Harman and Bartůsková, 2023). Indeed, five years after obtaining their STEM master's degree, 85% of women are employed, compared to 92.5% of men, and only 45.1% of women have a stable job, compared to 62.5% of men (Almalaurea, 2018b).

It has been suggested that this gap in performance between secondary education and tertiary education, more generally, the persistent differences in the numbers of male and female students earning a STEM university degree, can be attributed to the teaching methods currently adopted in tertiary education. Blickenstaff (2005) and Kanny *et al.* (2014), who systematically reviewed the literature on the gender gap in STEM in the last 40 years, highlight the importance of teaching-related barriers. Schools, teachers, pedagogy, curriculum and preparation, and classroom structure directly affect the gender gap in STEM. These factors also concur with the determination of psychological factors (such as self-confidence) often evoked as the primary source of this gender gap (Sax and Bryant, 2006). However, as will be apparent from the literature survey in Section 2, few studies address the extent to which the organisation of teaching activities at the university level may affect students' performance independently of their individual characteristics. Quantitative information about this issue is still lacking, especially if non-US data are considered.

In this paper, we attempt to fill this gap and provide novel evidence from Italy about the determinants of differential performance of students in higher education, focusing on gender-related issues and STEM subjects. We use an empirical methodology that attributes students' academic achievement to different factors. We start by separately considering faculties, which allows us to assess the role of each faculty's technology and the role of students' individual characteristics within each faculty (institutional details about the nature and duties of faculties in Italian universities during the period under scrutiny are provided in Section 3). Technology is modelled via a production set based on examining existing literature, while their technical efficiency measures the impact of students' unobserved individual characteristics (such as effort and ability [2]). Next, we consider how the assessment of technology and efficiency changes if we evaluate students' performance *vis-à-vis* a more comprehensive reference set (such as a scientific area). Faculties have access, potentially, to different production technologies, being differently conditioned by several factors, such as lecturers' characteristics and resource availability. In this manner, for instance, we can compare the determinants of female students' relative performance in STEM subjects with those of female students in other subjects.

More specifically, we apply a meta-frontier approach recently developed by Huang *et al.* (2014). This stochastic metafrontier approach disentangles the group-specific frontiers and metafrontier, which further decomposes the efficiency scores of various faculties into technical efficiency scores and technology gaps (refer to Section A1 in the Online Appendix for details on the applied procedure). A key advantage of this procedure is that it provides relevant policy information on the determinants of students' performance by relying on data routinely produced and stored by universities.

The analysis uses a sample of 53,159 first-year students from a large university in southern Italy from 2002–2003 to 2010–2011. The metafrontier has been employed to consider the efficiency change among different faculties. To our knowledge, this study is the first work to apply this methodology to the context of Italian higher education. Using student-level data allows us to control for the student's ability and other socio-economic characteristics that influence students' performance at a university. Furthermore, by focusing on data from one university, we reduce the role of unobserved heterogeneity in our data. As explained in Section 3, our statistical information crucially relies on an administrative dataset provided by the statistical office of the university selected for this study.

The paper is organised as follows: Section 2 summarises the main literature on the relevance of the factors affecting the performance of university students, focusing on the role of gender. The empirical approach is described in the Online Appendix (Section A1), while data and the production set are defined in Section 3. The results are presented and commented on in Section 4. Section 5 offers some concluding remarks.

2. University students' outcomes: the role of gender

This section provides a short survey and discusses the gender-related determinants of students' performance. We lay stress on the evidence relating to the dropout rate, as this kind of performance is more pertinent to the notions of efficiency that we shall address in our empirical analysis, where our primary measure of students' output will be the number of credits attained at the end of the first year of university. According to the literature (Boero *et al.*, 2001; Smith and Naylor, 2001; Bratti *et al.*, 2010), the performance of first-year credits is a strong predictor of the probability and speed of conclusive study completion. The policy relevance of gender in students' performance can scarcely be overemphasised due to its long-term impact on gender differences in the labour market.

The empirical findings in this field are mainly based on (single- and cross-university) survey data, generally controlling for student abilities before university enrolment (by final high-school grade and type of institution attended) and showing heterogeneous evidence. On the one hand, men are more likely than women to drop out of university (Montmarquette *et al.*, 2001 for Canada; Arulampalam *et al.*, 2004; Johnes and McNabb, 2004 for the United Kingdom; Di Pietro and Cutillo, 2008; Aina, 2011 for Italy). On the other hand, men's dropout rates are lower than women's (Ishitani, 2003 for the United States; Belloc *et al.*, 2010 for Italy), while other studies find no relevance for gender-related factors in this ambit (Smith and Naylor, 2001). Student drop-out behaviour may differ by gender when job opportunities are affected by a period of recession. When the risk of unemployment is increased and the lifetime earnings are lower, men, mainly from low unemployment areas, drop out more because they are more pessimistic about prospects in the labour market or have higher opportunity costs of remaining in education in the United Kingdom (Bradley and Migali, 2019). Differently, Adamopoulou and Tanzi (2017) find a decrease in the drop-out probability for men and an increase for women during a recession in Italy, mainly explained by the student's field of study, as traditionally, there is a higher concentration of female students in humanities. As far as Italy is concerned, Contini *et al.* (2018) and Contini and Salza (2020) find that gender differences are substantial such that women do better than men, having fewer dropouts and higher graduation probabilities, but these differences reduce and often disappear once we control for prior schooling and fields of study; Ghignoni *et al.* (2018) find those female students that drop out from university without a fallback plan having a higher probability of bad jobs and low earnings.

The proportion of enrolled women significantly impacts the dropout rate for students of either gender, even though the results are highly inconsistent. Using data from Ohio State University, Rogers and Menaghan (1991) find that women's likelihood of persistence in science and technology negatively correlates with the proportion of women in these courses. Using United States cross-university data, Sax (1996) shows that the gender composition of a course does not affect men's and women's grades and persistence; instead, a higher proportion of women has a positive effect on men's persistence and women's satisfaction. On the other hand, Johnson (2010), using data from undergraduate class sections at a single research institution in the United States, finds evidence that the proportion of female students is slightly higher in disciplines where they outperform male students. Johnes and McNabb (2004), relying on cross-university data from the United Kingdom, find that a higher proportion of female students in a higher-education institution is positively related to completion for men but negatively for women. They emphasise that their data do not contain information on gender mix by faculty. Still differently, Mastekaasa and Smeby (2008), using a sample of Norwegian universities, find that men's dropout rates are unrelated to the gender composition of courses, while women drop out of male-dominated programmes more significantly. This finding does not imply that women face particular problems in traditionally male contexts. Instead, women greatly prefer strongly female-dominated programmes.

There is minimal research on the heterogeneous effects of class size on students' outcomes in a higher-education setting. Kokkelenberg *et al.* (2008), using individual data on undergraduate students from a public university in the United States, find that class size negatively correlates with final grades across different departments and sample cuts. Utilising a natural experiment at Stanford Law School, Ho and Kelman (2014) conclude that assignment to small classes reduces the gender gap (against women) in large classes. On the other hand, relying on administrative data for the Economics and Management students at an Italian private university, De Giorgi *et al.* (2012) find both class size and class composition effects. A larger class size is associated with lower student performance, and this effect is more substantial for males and lower-income students. Furthermore, an increase in the proportion of female students increases performance, but only to a given threshold, after which the marginal effect declines (and eventually becomes negative).

A small but growing body of literature for American universities focuses on an instructor's gender role at the higher education level. Bettinger and Long (2005) find that a female instructor has a positive and significant role model in course selection and major choice. However, this finding is invalid for some male-dominated fields, including the sciences. Hoffman and Oreopoulos (2009) find that students taught by a same-gender instructor are less likely to drop a course and more likely to increase their grade performance, with males performing worse when assigned to a female instructor and females performing at the same level. Carrell *et al.* (2010) (in contrast with Bettinger and Long, 2005) find that when female students have female professors in STEM subjects, they face a higher likelihood of graduating. A professor's gender seems to be, instead, irrelevant in humanities courses. Artz and Welsh (2014) find that female students perform better under female lecturers, but this effect weakens as the proportion of female students increases. Solanki and Xu (2018) provide evidence that the gap between female and male students regarding course engagement and attitude toward a STEM subject is reduced when a female instructor teaches a course.

Several mechanisms have been considered to explain this empirical evidence. The literature emphasises that females leave university mainly because of social factors, while for males, the primary cause of quitting is their poor academic performance (Tinto, 1993, 1997; Yorke, 1999). Consequently, a possible explanation for the female dropout probability being much lower in female-dominated programmes than in balanced or male-dominated programmes could be that female-dominated programmes are particularly adept at fostering a positive environment for female students. Rogers and Menaghan (1991) and Johnes and McNabb (2004), who find evidence of a trend in the opposite direction, also explain. According to Rogers and Menaghan, fewer women in a course perform better as they do not constitute a challenge to the norms of gender relations in male-dominated classes and, thus, experience less pressure. Johnes and MacNabb suggest the relevance of a role model distinct from peer effects. Evidence suggesting that an increase in the number of female students may penalise their performance is consistent with a study that suggests that women respond less favourably than men to competitive environments (e.g. Gneezy *et al.*, 2003; Niederle and Vesterlund, 2011). Regarding the consequences of class size on gender-related performance, Weaver and Jiang (2005) argue that females are more sensitive to peer disapproval. Hence, smaller classes reduce fear and heighten participation as students (particularly females) experience less anxiety. Gender stereotypes may influence student perceptions, questioning the possible beneficial effects of the presence of female instructors in male-dominated fields (Bettinger and Long, 2005). Furthermore, the student-instructor gender (mis)match may not be gender exclusive: female instructors may exhibit more empathy toward the learning needs of students. There is ample room for further evidence on these issues, especially in a non-US university setting.

3. Data and production set

3.1 The data

We rely on a unique administrative dataset of 53,159 first-year students enrolled in a large public university in southern Italy for each academic year from 2002–2003 to 2010–2011. The analysis has been carried out only on first-year students to address the critical elements and weaknesses of the Italian higher-education system, such as the high dropout rate at the end of the first year and the consistent number of students who do not take any exams during their first year of study; the performance of first-year students has been considered a proxy of the regularity of the educational path and a suitable predictor of the probability of graduating on time (CNVSU, 2011). More specifically, the empirical investigation has been performed using an administrative dataset provided by the statistical office of the university chosen for this analysis. These data are collected via a student record system that gathers information at the individual level on prior educational qualifications, gender, age, socio-economic status, courses and modules in which students are enrolled, and assessment results. Furthermore, we relied on the University and faculty official reports to construct consistent information in each faculty on the number of registered students, full professors, associate professors, and lecturers. Our overall dataset gathered information on student demographics, such as gender, educational background and pre-enrolment characteristics (type of high school attended and grades on the high-school final exams), range of household's self-declared income, and general information regarding university careers and performance, such as passed exams and acquired credits. There is also information on the structural characteristics of the faculties, such as the number of students, the proportion of female students, the proportion of female lecturers, and the average class size. Data on the following faculties have been utilised: Economics, Law, Political Sciences, Engineering, Mathematical Physical and Natural Sciences, Arts, Languages and Educational Sciences [3]. Descriptive statistics are presented in the Online Appendix (Table A1a).

Remarkably, there is a close correspondence between the proportion of female students in this university and the corresponding data at the country level (Pearson and Spearman correlation coefficients equal to 0.92 and 0.90, respectively). Unfortunately, this comparison is impossible for class size and the proportion of female lecturers. For our analysis, we group the faculties into three main scientific areas following the guidelines of the Italian Ministry of Education: Pure and Applied Sciences (including Engineering, Mathematical Physical and Natural Sciences), Humanities (including Arts, Languages, and Educational Sciences), and

Social Sciences (including Economics, Law, Political Sciences). STEM-related subjects belong to Pure and Applied Sciences, although the two definitions do not precisely correspond.

Furthermore, the faculty's choice of where to enrol after secondary school is a crucial choice that may define the performance at university. Table A1b in the Online Appendix shows the descriptive statistics separated by gender to see patterns that lead to different faculty choices for women and men. The percentage of males from a scientific lyceum is higher in the faculties of Economics, Law, Political Sciences, Arts, Languages and Educational Sciences. Meanwhile, the Faculties of Engineering and Mathematical Physical and Natural Sciences have a higher percentage of female students from a scientific lyceum. Except for the Faculty of Languages, there is a higher percentage of females from a classical lyceum. Moreover, female students in all faculties have higher high-school grades than males. The stratification by income is also interesting as there is a higher percentage of males with high household incomes in all the faculties. Men are, on average, older than females.

To better contextualise the analysis, the university selected for this study is a medium-large size. Approximately 40,000 students, on average, were registered over the sample years. Until 2011, the University was organised into departments and faculties, with departments overseeing research and faculties in charge of teaching provision and management [4]. On average, approximately 900 full professors, associate professors and lecturers were teaching in nine faculties. Regarding the financial commitment of the institution, in the last decade, approximately € 90,000,000 have been invested every year in academic and non-academic resources. The total university turnover fluctuated during the same period at approximately € 100,000,000. The University is state-funded. Fees, dependent on the student's household income, are pretty low (on average, across the period considered in the analysis, the highest fees are generally approximately € 1000 per year). The university's main campus is located a few kilometres east of the central city in the area, to which it is well connected via a motorway. The institution can thus be considered a quasi-urban university, as it is located near a city with slightly above 100,000 inhabitants and a relative mean income of approximately € 22,300 (near the national mean value).

3.2 The production set and other variables

Data includes information on inputs and outputs usually considered in university efficiency studies (Johnes, 2006; Barra and Zotti, 2016). The production set, presented in the Online Appendix (Table A2), includes three measures of inputs and one measure of output.

Knowledge and skills when entering tertiary education are important determinants of students' performance as higher ability lowers their educational costs and increases their motivation. As indicated in Section 2, the type of high school and final grades correlate with students' university performance. Therefore, the first input (*High school grade*) is the grade each student has obtained on the national exam held at the end of the high school cycle. Following the same argument, the second input (*High school track*) is a categorical variable for the type of secondary school track (academic-oriented or lyceum vs technical high school and professional high school). The third input (*Income*) is the level of self-reported household income. More specifically, the income variable measures a student's equivalent income based on household income, wealth, and size (similar to the variable used by Belloc *et al.*, 2010). A student's equivalent income is identified by the class of tuition fees paid when enrolling at a university. We constructed only three categories because a finer disaggregation was not stable throughout the study years: (1) low-household income if a student has an equivalent income from € 0 to € 12,000, (2) middle-household income if a student has an equivalent income from €12,001 to € 32,000, and (3) high-household income if a student has an equivalent income greater than € 32,000.

Our baseline output (*Credits*) measure is the number of credits obtained at the end of the first year, like Agasisti and Murtinu (2016). They measure students' academic performance at the end of the first year. Various studies (Boero *et al.*, 2001; Smith and Naylor, 2001; Bratti *et al.*, 2010) point out that the number of first-year credits strongly predicts the probability and speed of conclusive study completion. The choice of this output variable is also strictly related to the institutional setup of Italian universities. Credits obtained at the end of the first year matter for a single university more than any other measure of students' academic achievements, as funding from the central government is primarily conditional upon them. Quantitative indicators were developed to evaluate teaching productivity because of the reforms implemented in the Italian higher education system in the 1990s and at the beginning of the 2000s. Universities were urged to give students a certain number of credits as close as possible to those theoretically obtainable each year and to graduate within the timeframe established by the degree protocol. The transition between the first and second years was considered the main checkpoint to evaluate the regularity of the educational path, and the credits obtained at the end of the first year assumed a crucial role [5]. It could be argued that, particularly in the Italian university system, students can choose whether to register to attend the exam. This decision can depend on risk aversion, and most part of the economic literature agrees on a higher risk aversion for the female population (Croson and Gneezy, 2009). Therefore, women can postpone the exam date until they feel very prepared, while men can attempt the exam even if they are less ready. This may affect first-year performance but not final performance. Tables A1c and A1d in the Online Appendix compare first-year and final performance by gender. More specifically, for each faculty, categories of faculty and gender, we compare first-year performances (credits, credits weighted and grades) with three potential measures of the final performances: percentage of graduates, final grade and time to get the degree. Regarding the number of credits and grades obtained at the end of the first year, female students perform better (Table A1c). This trend is also confirmed concerning the final performances. Indeed, female students have a higher probability of graduating, obtaining a higher final grade and obtaining a degree in a shorter time (Table A1d) [6]. This evidence confirms that the first-year performance is a reasonable estimate of the final performances by gender and could be safely used as the main output.

Given our interest in gender issues, a dummy variable (*Female*) with a value of 1 if a student is female and 0 otherwise is always included in the estimates. We also interact *Female* with the inputs (*High school grade*, *High school track* and *Income*) and several control variables available at the faculty level, such as the average class size (*Average class size*), measured as the total number of students divided by the number of classrooms within the faculty; the proportion of female students (*Proportion female students*), measured as the total number of female students divided by the total number of students within the faculty; and the proportion of female lecturers (*Proportion female lecturers*), measured as the total number of female academic staff divided by the total number of academic staff within the faculty.

4. The empirical results

4.1 Main results

We start our empirical exercise by testing whether the faculties share the same technology. If a common underlying technology generated the faculties' data, we would not be justified in adopting the metafrontier production function approach. Following Huang *et al.* (2014), we apply a likelihood ratio test for the null hypothesis that the production frontiers are the same for different faculties. The null is rejected across the three scientific areas and for the whole university (results are available upon request). As this evidence indicates that faculties are operating under heterogeneous technologies, potential production technology gaps exist, and adopting the metafrontier production function approach is vindicated.

In our empirical analysis, regressions of the first and second stages (as described in Section A1 of the Online Appendix) are run for four different models. First-stage regressions relate to the performance of students belonging to a given faculty vis-à-vis a frontier, including all students within a given scientific area. In contrast, second-stage regressions relate to the performance of faculties belonging to a given area vis-à-vis a frontier, including all the University's areas. We start with a simple model, which includes our measure of output, the basic inputs (high school score, high school track, and income), and a gender dummy (equal to 1 for female students). Next, we ran a model with individual interactions, where high school score, high school track, and income interacted with the female dummy. We employ a model with faculty interactions, where the female dummy has interacted with the average size of classes in a faculty, the female proportion of students in a faculty, and the female proportion of lecturers in a faculty. Finally, a model with both individual interactions and faculty interactions is estimated. A distinctive trait of all these estimates is that convergence of the maximum likelihood procedures is never achieved when positing for the u_{it} term the truncated-normal distribution needed to make it a function of a vector of exogenous variables. On the other hand, the half-normal distribution from (4), see Section A1 in the Online Appendix, performs rather well. A consequence of this condition is that the most remarkable differences among the faculties' estimates arise because of differences in their inputs' coefficients. We thus chose a format for presenting the evidence that highlights these differences, particularly for the gender-related variables. We give results for the simple model, which provides a convenient baseline, and for a restricted version with individual and faculty interactions, which maximises the log-likelihood ratio test vis-à-vis the baseline [7].

First, we present the individual faculty estimates by scientific areas: Social Sciences (Table 1a, Online Appendix), Pure and Applied Sciences (Table 1b, Online Appendix) and Humanities (Table 1c, Online Appendix). Second, we present the aggregate second-stage estimates for each scientific area (Table 1d, Online Appendix). High school grades and track are always positive and statistically significant among the primary inputs. At the same time, income (mostly positive) is not significant for Political Sciences, Engineering, Mathematical Physical and Natural Sciences, and Arts. There is evidence of a substantial gender gap (penalising females) in the measure of faculty technology for Law, Engineering, Mathematical Physical and Natural Sciences, and Arts. In Economics and Languages, the female dummy is negative but insignificant; in Political Sciences and Educational Sciences, the dummy is positive (and insignificant). The strongest gender gap is evident in the Pure and Applied Sciences area. This conclusion is also confirmed by the second stage (area) estimates, where the female dummy is always negative and significant. Still, its absolute value is much higher for the Pure and Applied Sciences area.

Regarding the efficiency scores (Table 1e, Online Appendix), they are substantially low across all faculties, with averages ranging from 0.3849 to 0.2657, mainly because of the high number of students that conclude the first year without receiving any credit. It is worth repeating that these scores cannot be associated with either gender or any other covariate. Accordingly, they can be taken as being driven by students' unobserved individual characteristics. On the other hand, the TGRs are high for all faculties, with average values from 0.9997 to 0.8442. Particularly high values are obtained in the Social Sciences area. As a result, considering that the technologies of relatively similar faculties do not entail significant differences in performance and efficiency, the technical efficiency concerning the metafrontier production technology (MTE) is similar to the technical efficiency concerning the faculty technology (TE). The lowest TGR (and hence the most significant difference between TE and MTE) is obtained for the faculty of Engineering.

The interactions with the female dummy reported in Tables A3a (Social Sciences), A3b (Pure and Applied Sciences), A3c (Humanities) and A3d (2nd stage estimates) in the Online Appendix highlight some interesting differences in the faculties' technologies. In three faculties (Political Sciences, Mathematical Physical and Natural Sciences, and Arts), a significant negative interaction exists between the female dummy and the proportion of female students. An increase in the proportion of female students worsens the gender gap against women (like the findings of Rogers and Menaghan, 1991; Johnes and McNabb, 2004). As shown in Figure A1 in the Online Appendix, these faculties exhibit a relatively big relative change (either positive or negative) in the proportion of female students. For other faculties, the proportion is more stationary (and the interaction with the female dummy is insignificant in estimation). Two faculties (Economics and Political Sciences) have a significant positive interaction between the female dummy and the proportion of female lecturers. An increase in the proportion of female lecturers reduces the gender gap against women (which echoes broadly similar results obtained for American universities). Once again, the ratio of female lecturers in these faculties shows a sizeable upward change (refer to Figure A1 in the Online Appendix). This proportion is more stationary for other faculties. There is an exception to this rule, the faculty of Educational Sciences, where the proportion of female lecturers considerably increases to no avail. However, there is no gender divide in Educational Sciences, and the proportion of female students consistently exceeds 90% (refer to Figure A2 in the Online Appendix). Hence, this evidence is consistent with the results of Artz and Welsch (2104). In Mathematical Physical and Natural Sciences, the rise in the proportion of female lectures is relatively large but still considerably less than the large increase in the proportion of female students. Further evidence (from different universities) seems to be needed if it must be concluded that this faculty is impervious to the increase in female lecturers due to strong gender stereotypes (Bettinger and Long, 2005). Given this caveat, this joint evidence has at least a tentative policy implication. If the proportion of female students increases without impairing their relative performance, the ratio of female lecturers should also be increased [8]. There are also significant interactions in the female dummy with average class size, high school grade and income, but they are somewhat sporadic, and no pattern seems to emerge.

A final point worth emphasising is that a female term that comprises the gender dummy and all its significant interactions can be constructed. We estimate this female term using the mean values of the interacted variables. These values are significant for Law, Economics (Table A3a), Engineering, Mathematical Physical and Natural Sciences (Table A3b), and Arts (Table A3c). In addition, the second stage (area) estimates show that the strongest gender gap is again observed in the Pure and Applied Sciences (Table A3d). The results concerning the efficiency levels are almost the same in the estimates that include interactions with the female dummy (reported in Table A3e). There are only notable changes in the TGRs for Engineering, which increase compared to the baseline model, and for Mathematical Physical and Natural Sciences and Languages, which (to a lesser extent) decrease compared to the baseline model. For a comprehensive list of robustness checks, refer to Tables A4a-A4d, A5, A6 e A7, Sections A2.1 and A2.2, in the Online Appendix.

5. Concluding remarks

While the participation of women in higher education has increased over the last twenty years, there are concerns over the proportion of female students and their performance, especially relating to STEM subjects. It has been argued that there is a need to re-evaluate the pedagogy of higher-education institutions, especially concerning STEM subjects. In doing so, however, it is not easy to disentangle the contribution to the performance of female students in the organisation of teaching at the faculty level from the role of students' individual characteristics.

In this paper, we have provided an empirical approach by which data routinely produced and stored by higher-education institutions can be used to monitor the performance of students at the end of their first year, to identify the efficiency of students achieved within a given faculty, and to determine whether the technologies of relatively similar faculties entail significant differences in performance and efficiency. This procedure considers both faculty-specific frontiers and a metafrontier for rather similar faculties, allowing the calculation of a student's technical efficiency compared to any faculty's production technology, a student's technical efficiency compared to the metafrontier production technology, and a technology gap ratio that measures the impact of a given faculty on the student's performance. Using this procedure, we provide evidence for a large Italian university, a field of analysis with some novel content.

We find that the mean students' efficiency measured via faculty-specific frontiers is low and does not seem to be related to any available covariates, including gender. This finding likely reflects the critical role of unobserved heterogeneity, possibly driven by students' effort and ability. On the other hand, the mean efficiency of faculties vis-à-vis the metafrontier, although equally unrelated to observables, is relatively high, sometimes approaching the unity bound. This result means that the different technologies of similar faculties do not considerably impact efficiency. However, within any given faculty, technology affects the performance of any single student via the production function parameters. In this paper, we have devoted particular attention to the parameters that model the role of gender, about which there is still considerable uncertainty in the literature.

Whenever significant, the gender gap in performance in our data is always against women. We find five faculties with a significant gender gap: Economics, Law, Engineering, Arts, Mathematical Physical and Natural Sciences. These faculties are liberal profession-oriented faculties (except for Arts). The gender gap is strongest in the Pure and Applied Science areas, where STEM-related subjects belong, although the two definitions do not precisely correspond. In four faculties (Political Sciences, Arts, Languages, Mathematical Physical and Natural Sciences), the divide against women increases with a larger proportion of female students (which is reminiscent of the results of Rogers and Menaghan, 1991; Johnes and McNabb, 2004; and is consistent with a broader literature that suggests that women react poorly to the competition: e.g. Gneezy *et al.*, 2003; Niederle and Vesterlund, 2011). These four faculties share the feature of having a relatively strong trend (positive or negative) in the proportion of female students. In other faculties, this proportion is stationary. In two faculties (Political Sciences and, to a slighter extent, Economics), the divide against women is reduced by a larger proportion of female lecturers. In these faculties, the variable of interest (proportion of female lecturers in this case) has a particularly strong trend. In Section 4, we have explained the lack of impact on the proportion of female lecturers in Mathematical Physical and Natural Sciences and Educational Sciences (although the former case may require further evidence). The core of this evidence remains valid when a robustness check is carried out, changing the output indicator in the production set. Weighting the credits obtained at the end of the first year with the grades associated with each exam reduces the strength and pervasiveness of the gender divide, which remains strongest in the faculty of Mathematical Physical and Natural Sciences. The association of the gender divide with the proportions of female students and female lecturers also remains virtually untouched.

We find strong evidence of lower performance among female students in STEM subjects. As this lower performance is related to lower probability and speed of final studies completion, it can be at least part of the explanation for the still low numbers of female students earning a degree in these subjects. Furthermore, our evidence is consistent with the proposition that increasing the proportion of female students in a faculty without a proportionate rise in the proportion of female lecturers is likely to worsen the relative performance of female students. Yet note that both proportions change very little in our data for Engineering. In contrast, in our data for Mathematical Physical and Natural Sciences, the rise in female lectures is relatively significant but still considerably less than the large increase in female students. Consequently, and especially for STEM subjects, further validation of our tentative policy implication would be gained from datasets with different patterns of variations in these proportions across faculties. In future work, we also plan to extend our empirical set-up, examining the different organisational structures within faculties that can explain the differences in the performance of males and females. Investigating the performance variation over time may also be essential to assess the impact of eventual institutional reforms on our variable of interest.

Notes

1. The National Recovery and Resilience Plan is part of the that the European Union negotiated in response to the pandemic crisis. For more details, see https://www.governo.it/it/approfondimento/pnrr-gli-obiettivi-e-la-struttura/16702.
2. As pointed out by a referee, effort could in principle be measured through attendance and hours of study. Unfortunately, we do not have this kind of information.
3. Note that, due changes in their credit regime, faculties of Law were excluded from our administrative dataset from 2007 onwards.
4. This institutional setup was ended in 2011, which also led to a stop in the construction of the dataset used in this study.
5. Formally, the main parameters applied by the Italian Ministry of Education in the teaching quality assessment include (1) the proportion of students enrolled in the second year, having already obtained a given number of credits in the first year, and (2) the proportion of students who do not obtain any credits or pass any exam (i.e. inactive students) at the end of the first year (Ministerial Decree 18 October 2007, n. 506; see also CNVSU, 2009).
6. The only exceptions are the faculties of Mathematical Physical and Natural Sciences, Art and Educational Sciences where male students have a lower time to get the degree.
7. The full set of estimates is available upon request.
8. If female students react poorly to competition, their attitudes can be influenced by the provision of positive role models, such as female lecturers. Mechanisms of this kind have recently been analysed in Leibbrandt *et al.* (2018).

References

Adamopoulou, E. and Tanzi, G.M. (2017), "Academic drop-out and the great recession", *Journal of Human Capital*, Vol. 11 No. 1, pp. 35-71, doi: 10.1086/690650.

Agasisti, T. and Murtinu, S. (2016), "Grants in Italian university: a look at the heterogeneity of their impact on students'performances", *Studies in Higher Education*, Vol. 41 No. 6, pp. 1106-1132, doi: 10.1080/03075079.2014.966670.

Aina, C. (2011), "The determinants of success and failure of Italian university students. Evidence from administrative data", *Rivista Internazionale di Scienze Sociali*, Vol. 119 No. 2, pp. 85-108.

Almalaurea (2018a), "XX Indagine - Profilo dei Laureati 2017", available at: https://www.almalaurea.it/universita/profilo/profilo2017

Almalaurea (2018b), "XX Indagine - Condizione occupazionale dei Laureati", available at: https://www.almalaurea.it/universita/occupazione/occupazione16

Artz, B. and Welsh, D.M. (2014), "The Effect of peer and professor gender on college student performance", *Southern Economic Journal*, Vol. 80 No. 3, pp. 816-838, doi: 10.4284/0038-4038-2012.158.

Arulampalam, W., Naylor, R., Robin, A. and Smith, J.P. (2004), "Hazard model of the probability of medical school drop-out in the UK", *Journal of the Royal Statistical Society*, Vol. 167 No. 1, pp. 157-178, doi: 10.1046/j.0964-1998.2003.00717.x.

Barra, C. and Zotti, R. (2016), "Managerial efficiency in higher education using individual versus aggregate level data. Does the choice of decision making unit count?", *Managerial and Decision Economics*, Vol. 37 No. 2, pp. 106-126, doi: 10.1002/mde.2697.

Belloc, F., Maruotti, A. and Petrella, L. (2010), "University drop-out: an Italian experience", *Higher Education*, Vol. 60 No. 2, pp. 127-138, doi: 10.1007/s10734-009-9290-1.

Bertrand, M. (2018), "Coase lecture – the glass ceiling", *Economica*, Vol. 85 No. 338, pp. 205-231, doi: 10.1111/ecca.12264.

Bettinger, E. and Long, B. (2005), "Do faculty serve as role models? The impact of instructor gender on female students", *American Economic Review*, Vol. 95 No. 2, pp. 152-157, doi: 10.1257/000282805774670149.

Blickenstaff, J.C. (2005), "Women and science careers", *Gender and Education*, Vol. 17 No. 4, pp. 369-386, doi: 10.1080/09540250500145072.

Boero, G., McNight, A., Naylor, R. and Smith, J. (2001), "Graduates and graduate labour markets in the UK and Italy", *Lavoro e Relazioni industriali*, Vol. 2, pp. 131-172.

Bradley, S. and Migali, G. (2019), "The effects of the 2006 tuition fee reform and the Great Recession on university student dropout behaviour in the UK", *Journal of Economic Behavior and Organization*, Vol. 164, pp. 331-356, doi: 10.1016/j.jebo.2019.06.002.

Bratti, M., Broccolini, C. and Staffolani, S. (2010), "Higher education reform, student time allocation and academic performance in Italy: evidence from a Faculty of Economics", *Rivista Italiana Degli Economisti*, Vol. 15 No. 2, pp. 275-304.

Carrell, S., Page, M. and West, J. (2010), "Sex and science: how professor gender perpetuates the gender gap", *Quarterly Journal of Economics*, Vol. 125 No. 3, pp. 1101-1144, doi: 10.1162/qjec.2010.125.3.1101.

CNVSU (2009), "Comitato Nazionale per la Valutazione del Sistema Universitario", Indicatori per la ripartizione del Fondo di cui all'art. 2 della Legge 1/2009, DOC 07/2009, MIUR, Roma.

CNVSU (2011), *Undicesimo Rapporto sullo Stato del Sistema Universitario*, MIUR, Roma.

Contini, D. and Salza, G. (2020), "Too few university graduates. Inclusiveness and effectiveness of the Italian higher education system", *Socio-Economic Planning Sciences*, Vol. 71, 100803, doi: 10.1016/j.seps.2020.100803.

Contini, D., Cugnata, F. and Scagni, A. (2018), "Social selection in higher education. Enrolment, dropout, and timely degree attainment in Italy", *Higher Education*, Vol. 75 No. 5, pp. 785-808, doi: 10.1007/s10734-017-0170-9.

Croson, R. and Gneezy, U. (2009), "Gender differences in preferences", *Journal of Economic Literature*, Vol. 47 No. 2, pp. 1-27, doi: 10.1257/jel.47.2.448.

De Giorgi, G., Pellizzari, M. and Woolston, W.G. (2012), "Class size and class heterogeneity", *Journal of the European Economic Association*, Vol. 10 No. 4, pp. 795-830, doi: 10.1111/j.1542-4774.2012.01073.x.

Di Pietro, G. and Cutillo, A. (2008), "Degree flexibility and university drop-out: the Italian experience", *Economics of Education Review*, Vol. 27 No. 5, pp. 546-555, doi: 10.1016/j.econedurev.2007.06.002.

Ghignoni, E., Croce, G. and D'Ambrosio, A. (2018), "University dropouts vs high school graduates in the school-to-work transition. Who is doing better?", *International Journal of Manpower*, Vol. 40 No. 3, pp. 449-472, doi: 10.1108/ijm-02-2018-0051.

Gneezy, U., Niederle, M. and Rustichini, A. (2003), "Performance in competitive environments: gender differences", *Quarterly Journal of Economics*, Vol. 118 No. 3, pp. 1049-1074, doi: 10.1162/00335530360698496.

Granato, S. (2020), "Early influences and the gender Gap in STEM", *Granato*, SSRN, doi: 10.2139/ssrn.3845391.

Harman, J. and Bartůsková, L. (2023), "The gender pay in the Visegrad Groups", *Journal of Economic Studies*, Vol. 51 No. 4, pp. 733-763, doi: 10.1108/JES-02-2023-0072.

Ho, D.E. and Kelman, M.J. (2014), "Does class size affect the gender gap? A natural experiment in law", *Journal of Legal Studies*, Vol. 43 No. 2, pp. 291-321, doi: 10.1086/676953.

Hoffman, F. and Oreopoulos, P. (2009), "A professor like me: the influence of instructor gender on university achievement", *Journal of Human Resources*, Vol. 44 No. 2, pp. 479-494, doi: 10.1353/jhr.2009.0024.

Huang, C.J., Huang, T.H. and Liu, N.H. (2014), "A new approach to estimating the metafrontier production function based on a stochastic Frontier framework", *Journal of Productivity Analysis*, Vol. 42 No. 3, pp. 241-254, doi: 10.1007/s11123-014-0402-2.

Ishitani, T.T. (2003), "A longitudinal approach to assessing attrition behavior among first-generation students: time-varying effects of pre-college characteristics", *Research in Higher Education*, Vol. 44 No. 4, pp. 433-449, doi: 10.1023/a:1024284932709.

Jacobs, J.A. (2003), "Detours on the road to equality: women, work and higher education", *Contexts*, Vol. 2 No. 1, pp. 32-41, doi: 10.1525/ctx.2003.2.1.32.

Johnes, J. (2006), "Measuring efficiency: a comparison of multilevel modelling and data envelopment analysis in the context of higher education", *Bulletin of Economic Research*, Vol. 58 No. 2, pp. 75-104, doi: 10.1111/j.0307-3378.2006.00238.x.

Johnes, G. and McNabb, R. (2004), "Never give up on the good times: student attrition in the UK", *Oxford Bulletin of Economics and Statistics*, Vol. 66 No. 1, pp. 23-47, doi: 10.1111/j.1468-0084.2004.00068.x.

Johnson, I.Y. (2010), "Class size and student performance at a public research university: a cross-classified model", *Research in Higher Education*, Vol. 51 No. 8, pp. 701-723, doi: 10.1007/s11162-010-9179-y.

Kanny, M.A., Sax, L.J. and Riggers-Pieh, T.A. (2014), "Investigating forty years of stem research: how explanations for the gender gap have evolved over time", *Journal of Women and Minorities in Science and Engineering*, Vol. 20 No. 2, pp. 127-148, doi: 10.1615/jwomenminorscieneng.2014007246.

Kokkelenberg, E.C., Dillon, M. and Christy, S.M. (2008), "The effects of class size on student grades at a public university", *Economics of Education Review*, Vol. 27 No. 2, pp. 221-233, doi: 10.1016/j.econedurev.2006.09.011.

Leibbrandt, A., Wang, L.C. and Foo, C. (2018), "Gender quotas, competitions, and peer review: experimental evidence on the backlash against women", *Management Science*, Vol. 64 No. 8, pp. 3469-3970, doi: 10.1287/mnsc.2017.2772.

Mastekaasa, A. and Smeby, J.C. (2008), "Educational choice and persistence in male-and female-dominated fields", *Higher Education*, Vol. 55 No. 2, pp. 189-202, doi: 10.1007/s10734-006-9042-4.

McNabb, R., Pal, S. and Sloane, P. (2002), "Gender differences in educational attainment: the case of university students in England and Wales", *Economica*, Vol. 69 No. 275, pp. 481-503, doi: 10.1111/1468-0335.00295.

Montmarquette, C., Mahseredjian, S. and Houle, R. (2001), "The determinants of university dropouts: a bivariate probability model with sample selection", *Economics of Education Review*, Vol. 20 No. 5, pp. 475-484, doi: 10.1016/s0272-7757(00)00029-7.

Niederle, M. and Vesterlund, L. (2011), "Gender and competition", *Annual Review of Economics*, Vol. 3 No. 1, pp. 601-630, doi: 10.1146/annurev-economics-111809-125122.

OECD (2010), *Education at a Glance*, OECD Publishing, Paris.

Plan International and Bocconi University (2021), "Current and future challenges for gender equality in Italy: the gender digital divide", available at: https://www.dondena.unibocconi.eu/

Rogers, S.J. and Menaghan, E.G. (1991), "Women's persistence in undergraduate majors - the effects of gender-disproportionate representation", *Gender & Society*, Vol. 5 No. 4, pp. 549-564, doi: 10.1177/089124391005004007.

Sax, L.J. (1996), "The dynamics of "tokenism"': how college students are affected by the proportion of women in their major", *Research in Higher Education*, Vol. 37 No. 4, pp. 389-425, doi: 10.1007/bf01730108.

Sax, L.J. and Bryant, A.N. (2006), "The impact of college on sex-atypical career choices of men and women", *Journal of Vocational Behavior*, Vol. 68 No. 1, pp. 52-63, doi: 10.1016/j.jvb.2005.01.004.

Smith, J.P. and Naylor, R.A. (2001), "Dropping out of university: a statistical analysis of the probability of withdrawal for UK university students", *Journal of the Royal Statistical Society-Series A*, Vol. 164 No. 2, pp. 389-405, doi: 10.1111/1467-985x.00209.

Solanki, S.M. and Xu, D. (2018), "Looking beyond academic performance: the influence of instructor gender on student motivation in STEM fields", *American Educational Research Journal*, Vol. 55 No. 4, pp. 801-835, doi: 10.3102/0002831218759034.

Tinto, V. (1993), *Leaving College. Rethinking the Causes and Cures of Student Attrition*, University of Chicago Press, Chicago.

Tinto, V. (1997), "Colleges as communities: exploring the educational character of student persistence", *Journal of Higher Education*, Vol. 68 No. 6, pp. 599-623, doi: 10.2307/2959965.

Weaver, R.R. and Jiang, Q. (2005), "Classroom organization and participation: college students' perceptions", *Journal of Higher Education*, Vol. 76 No. 5, pp. 570-601.

Yorke, M. (1999), *Leaving Early. Undergraduate Non-completion in Higher Education*, Falmer, London.

Further reading

Agasisti, T. and Berbegal-Mirabent, J. (2021), "Cross-country analysis of higher education institutions' efficiency: the role of strategic positioning", *Science and Public Policy*, Vol. 48 No. 1, pp. 66-79, doi: 10.1093/scipol/scaa058.

Atzeni, G., Deidda, L.G., Delogu, M. and Paolini, D. (2022), *Drop-out Decisions in a Cohort of Italian Universities Teaching, Research and Academic Careers: An Analysis of the Interrelations and Impacts*, Springer International Publishing, Cham, pp. 71-103.

Battese, G.E. and Coelli, T.J. (1995), "A model for technical inefficiency effects in a stochastic Frontier production function for panel data", *Empirical Economics*, Vol. 20 No. 2, pp. 325-332, doi: 10.1007/bf01205442.

Battese, G.E., Rao, D.S.P. and O'Donnell, C.J. (2004), "A metafrontier production function for estimation of technical efficiencies and technology gaps for firms operating under different technologies", *Journal of Productivity Analysis*, Vol. 21 No. 1, pp. 91-103, doi: 10.1023/b:prod.0000012454.06094.29.

Byrnes, J.P., Miller, D.C. and Schafer, W.D. (1999), "Gender differences in risk taking: a meta-analysis", *Psychological Bulletin*, Vol. 75 No. 3, pp. 367-383, doi: 10.1037//0033-2909.125.3.367.

Fadigan, K.A. and Hammrich, P.L. (2004), "A longitudinal study of the educational and career trajectories of female participants of an urban informal science education program", *Journal of Research in Science Teaching*, Vol. 41 No. 8, pp. 835-860, doi: 10.1002/tea.20026.

Huang, C.J. and Liu, J.T. (1994), "Estimation of non-neutral stochastic Frontier production function", *Journal of Productivity Analysis*, Vol. 5 No. 2, pp. 171-180, doi: 10.1007/bf01073853.

OECD (2017), "The under-representation of women in STEM fields", in *The Pursuit of Gender Equality: An Uphill Battle*, OECD Publishing, Paris, doi: 10.1787/9789264281318-10-en.

Ong, M., Wright, C., Espinosa, L.L. and Orfield, G. (2011), "Inside the double bind: a synthesis of empirical research on undergraduate and graduate women of color in science, technology, engineering, and mathematics", *Harvard Educational Review*, Vol. 81 No. 2, pp. 172-209, doi: 10.17763/haer.81.2.t022245n7x4752v2.

O'Donnell, C.J., Rao, D.S.P. and Battese, G.E. (2008), "Metafrontier frameworks for the study of firm-level efficiencies and technology ratios", *Empirical Economics*, Vol. 34 No. 2, pp. 231-255, doi: 10.1007/s00181-007-0119-4.

Sulis, I., Barbieri, B., Salaris, L., Melis, G. and Porcu, M. (2023), "Gender bias in university student mobility: a cohort analysis in Italy", *International Journal of Manpower*, Vol. 45 No. 5, pp. 1076-1092, doi: 10.1108/IJM-02-2023-0101.

Supplementary material

The supplementary material for this article can be found online.

Corresponding author

Cristian Barra can be contacted at: cbarra@unisa.it