Author Query Form

Journal: *Journal of Economic Geography*
Article Doi: 10.1093/jeg/lby044
Article Title: The role of fees in foreign education: evidence from Italy*
First Author: Michel Beine
Corr. Author: Michel Beine

AUTHOR QUERIES – TO BE ANSWERED BY THE CORRESPONDING AUTHOR

The following queries have arisen during the typesetting of your manuscript. Please click on each query number and respond by indicating the change required within the text of the article. If no change is needed please add a note saying "No change."

AQ1: Please check that all names have been spelled correctly and appear in the correct order. Please also check that all initials are present. Please check that the author surnames (family name) have been correctly identified by a pink background. If this is incorrect, please identify the full surname of the relevant authors. Occasionally, the distinction between surnames and forenames can be ambiguous, and this is to ensure that the authors' full surnames and forenames are tagged correctly, for accurate indexing online. Please also check all author affiliations.

AQ2: Please provide missing affiliations for all authors.

AQ3: Please specify the corresponding author and also provide complete postal and email address.
Please provide a Funding statement, detailing any funding received. Remember that any funding used while completing this work should be highlighted in a separate Funding section. Please ensure that you use the full official name of the funding body, and if your paper has received funding from any institution, such as NIH, please inform us of the grant number to go into the funding section. We use the institution names to tag NIH-funded articles so they are deposited at PMC. If we already have this information, we will have tagged it and it will appear as coloured text in the funding paragraph. Please check the information is correct.

AQ4: Please provide complete details for (Foreign Student Flows 2011 and Dal Bianco 2007) in the reference list or delete the citation from the text.

AQ5: There is no mention of (Agasisti & Bianco 2007) in the text. Please insert a citation in the text or delete the reference as appropriate.

AQ6: Please provide conference location for the references.

AQ7: Please provide location of publisher for the references.

AQ8: Please provide editors and publisher details for the references.

AQ9: Please provide notes for "*" in Table 5.

AQ10: Please check if the 'Numbeoo dataset' should be changed to 'Numbeo dataset'.

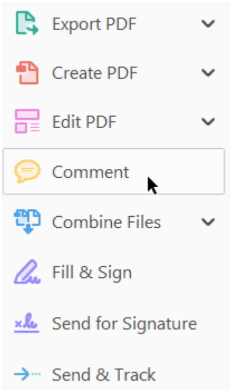
MAKING CORRECTIONS TO YOUR PROOF

These instructions show you how to mark changes or add notes to your proofs using Adobe Acrobat Professional versions 7 and onwards, or Adobe Reader DC. To check what version you are using go to **Help** then **About**. The latest version of Adobe Reader is available for free from get.adobe.com/reader.

DISPLAYING THE TOOLBARS

Adobe Reader DC

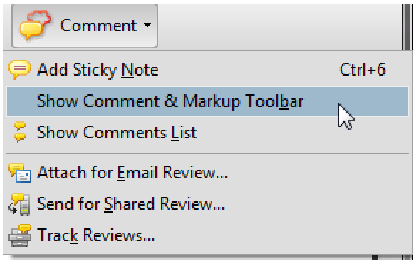
In Adobe Reader DC, the Comment toolbar can be found by clicking 'Comment' in the menu on the right-hand side of the page (shown below).



The toolbar shown below will then display along the top.

Acrobat Professional 7, 8, and 9

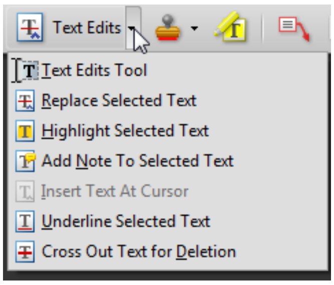
In Adobe Professional, the Comment toolbar can be found by clicking 'Comment(s)' in the top toolbar, and then clicking 'Show Comment & Markup Toolbar' (shown below).



The toolbar shown below will then be displayed along the top.

USING TEXT EDITS AND COMMENTS IN ACROBAT

This is the quickest, simplest and easiest method both to make corrections, and for your corrections to be transferred and checked.



1. Click **Text Edits**
2. Select the text to be annotated or place your cursor at the insertion point and start typing.
3. Click the **Text Edits** drop down arrow and select the required action.

You can also right click on selected text for a range of commenting options, or add sticky notes.

SAVING COMMENTS

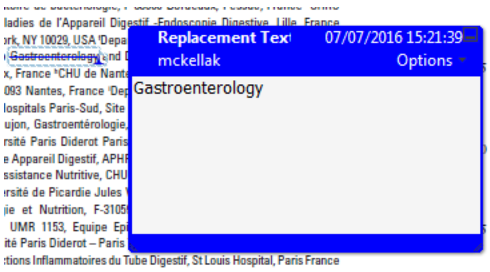
In order to save your comments and notes, you need to save the file (**File, Save**) when you close the document.

USING COMMENTING TOOLS IN ADOBE READER

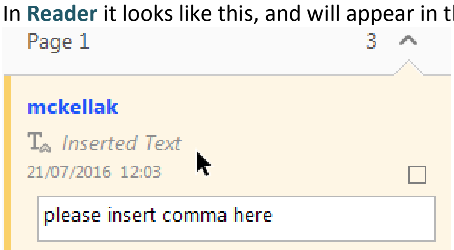
All commenting tools are displayed in the toolbar. You cannot use text edits, however you can still use highlighter, sticky notes, and a variety of insert/replace text options.

POP-UP NOTES

In both Reader and Acrobat, when you insert or edit text a pop-up box will appear. In **Acrobat** it looks like this:



In **Reader** it looks like this, and will appear in the right-hand pane:



DO NOT MAKE ANY EDITS DIRECTLY INTO THE TEXT, USE COMMENTING TOOLS ONLY.

[9.8.2018–8:26pm] [1–31] Paper: OP-JEGJ180045
Copyedited by: MANUSCRIPT CATEGORY: ORIGINAL ARTICLE

Journal of Economic Geography (2018) pp. 1–30 doi:10.1093/jeg/lby044

The role of fees in foreign education: evidence from Italy*

AQ1 AQ2 AQ3
*Michel Beine**,†, Marco Delogu** and Lionel Ragot****

*University of Luxembourg, IZA, CREAM and CES-Ifo
**Università Degli Studi di Sassari DISEA and CRENOS
***University Paris Nanterre, EconomiX and CEPII
†Correspondence to: *email* <michel.beine@uni.lu>

Abstract

This article studies the determinants of international students' mobility at the university level, focusing specifically on the role of tuition fees. We derive a gravity model from a Random Utility Maximization model of location choice for international students in the presence of capacity constraints of the hosting institutions. The last layer of the model is estimated using new data on student migration flows at the university level for Italy. We control for the potential endogeneity of tuition fees through a classical IV approach based on the status of the university. We obtain evidence for a robust and negative effect of fees on international student mobility, with an elasticity around −0.8. The estimations also confirm the positive impact of the quality of the education and support an important role of additional destination-specific variables such as host capacity, the expected return of education, the cost of living and the existence of education programs taught in English.

Keywords: Foreign students, tuition fees, location choice, university quality
JEL classifications: F22, H52, I23, O15
Date submitted: 26 April 2017 **Editorial decision:** 14 July 2018 **Date accepted:** 8 August 2018

1. Introduction

Foreign higher education has become an increasingly important phenomenon these days. The degree of mobility of prospective students wishing to acquire their educational skills abroad has been constantly on the rise for >50 years. While there were 0.6 million international students in 1975, this number rose to 3.5 million in 2005. Despite the turmoil caused by the financial crisis, the global quest for talented workers has pushed these numbers up further, with a 50% increase between 2005 and 2015 (OECD, 2015). Even though these global numbers obviously hide some uneven developments, the number of students emigrating abroad to complete their education has increased in all origin regions of the world. For more than 15 years, foreign students have represented the fastest growing category of international migrants. This article is concerned with the identification of the determinants of the location choice of foreign students at the university level in Italy, with a focus on the role of tuition fees.

For destination countries, the inflow of foreign students yields significant benefits. The benefits drawn by universities from the attraction of foreign students are numerous, which leads them to raise their external attractiveness. The first benefit is obviously financial. In some countries, attracting foreign students instead of domestic ones allows universities to charge higher fees, which might improve their budget situation. In countries in which price discrimination is not possible, to the extent that capacity constraints are not binding and marginal costs of additional students are low, attracting further students from abroad can also improve the financial health of the institutions. A second important advantage is that the inflow of foreign students allows the development of education programs that could not be possible if relying only on domestic students. In disciplines for which skills are easily transferable across countries such as natural sciences (or economics), the attraction of foreign students allows to achieve the critical mass needed to run specific education programs. Finally, foreign students contribute to the international integration of these universities. The attraction of students coming from specific countries can generate spillovers such as research collaboration between the universities of destination and those or the origin countries of these students. Whatever the specific reason, the numerous announcements of the universities toward foreign students reflect in general their desire to raise their external attractiveness. Understanding the factors that shape this attractiveness is, therefore, of utmost importance for the development of appropriate strategies to attract foreign students. This is the main question which is addressed in this article.

This article contributes to the literature on the identification of factors influencing students' decisions about where to study when they want to study abroad. In contrast with the existing literature, we assess these determinants at the university level, which allows to look at within-country factors. A special attention is devoted to the role of the tuition fees among these factors. Assessing the response of foreign students to variation in tuition fees is important. Under the traditional view, tuition fees are a fundamental component of education costs for students. This view suggests that the response of inflows of students to tuition fees should be negative. Still, the main bulk of the existing literature on the determinants of foreign students using cross-country data fails to find a clear negative response to fees increase. This failure entails several alternative explanations that need to be further assessed. One reason for the non-negative response might be that the impact of fees as a component of the education costs might be offset or mitigated by the signaling content of fees. In the presence of imperfect observation of university quality, higher fees might indeed signal better quality. A second explanation is that fees play a minor role in the decision of the place of study and their effect is dominated by factors such as the quality of the institution or work perspectives after graduation. Another explanation is that cross-country investigations do not allow to address this issue, since the use of an average level of fees across institutions does not allow to capture the variation across institutions. A final reason is mostly econometric. The assumption of exogenous fees is often inconsistent with what is going on in the real world of high education. Failure to properly account for the endogenous nature of fees in the econometric investigation leads to estimation bias towards non-negative effect of tuition fees.

This article revisits the issue of the impact of tuition fees on foreign education and attempts to fill the gaps in the existing literature. First, we use data at the university level for Italy for the academic year 2011–2012. Unlike other countries from Continental Europe such as France, Belgium or Germany, Italian universities (like

UK Universities) show significant variation in the tuition fees across institutions.[1] This in turn allows us to study the role of fees for foreign students when choosing one specific location. This is in addition to other institutional characteristics such as the quality of education, host capacity, expected income, cost of living and the use of the English language in the teaching programs. We compile and use data on foreign student flows between (almost) all countries of the world—the origin country—and each Italian university under investigation—the destination countries. Our econometric investigation, derived from a traditional random utility model (RUM), adapted to student migration, pays special attention to the role of tuition fees. The theoretical framework also clarifies the sources of endogeneity of fees in the empirical framework.

In the econometric investigation, we therefore explicitly take into account the endogenous nature of these fees. Results showing positive or even zero impact of fees might be spurious due to the high degree of endogeneity of fees. Fees are higher when universities succeed in attracting many students, which leads to reverse causality issues. Fee levels might be correlated with factors such as unobserved amenities in the university of destination or with unobserved quotas applied to foreign students. This calls for a causal identification accounting for the possible endogenous status of the observed fees in the econometric regression. We pay specific attention to this issue, using a classical instrumental variable (IV) approach. We instrument the tuition fees by the status of the university (private vs public). Private institutions tend to charge higher fees to cover specific costs and to offset the lower public subsidies compared with public institutions. Our exclusion restriction assumes that students have no particular preferences for private vs public institutions beyond the costs and the quality of education (for which we control in the regression) when choosing a specific university. We further show that the obtained negative impact of fees is robust to reasonable deviations from the strict exclusion restriction by employing methods described in Conley et al. (2012).

Our article is related to the extensive literature on foreign education using aggregate data. At the theoretical level, as reminded by Rosenzweig et al. (2006); Rosenzweig (2008), there are basically two complementary explanations for why students decide to go abroad to complete their higher education. While the education and migration models are about the decision to study abroad, much of the literature has been devoted to the location choice. Our article belongs to this category. Most of the literature makes use of country-level data and combines a multi-origin approach.[2] So far, this literature

1 The current version of this article is the result of a split from a larger working-paper that conducted such an investigation for these two countries, Italy and the UK (Beine et al., 2016). The econometric approach used in the companion paper for the UK (Beine et al., 2017) is nevertheless different. In particular, the way endogeneity is dealt with does not rely on an IV strategy but rather makes use of the institutional caps on fees in place in different regions of the UK. The results are nevertheless qualitatively similar.

2 Bessey (2012) focuses on foreign students in Germany, finding that the stock and the flow of students of the same nationality are positively correlated. Dreher and Poutvaara (2008) and Rosenzweig et al. (2006) look at the determinants of foreign education in the USA. The papers stress the importance of networks (Dreher and Poutvaara, 2008 36) and skill premium (Rosenzweig et al., 2006 36). Other studies combine various origins and destinations, carrying out estimations with a gravity model. Perkins and Neumayer (2014) consider many origin (151) and destination countries (105) over a couple of years and evaluate the role of geographic factors. Van Bouwel and Veugelers (2013) look at student migration among 18 European countries and assess the role of university quality, which was evaluated through the number of institutions appearing in the most widely known international university rankings. They show that quality matters but tend to find a positive impact of tuition fees. Beine et al. (2014) derive a gravity specification and focus on the 13 main destinations for foreign education. They estimate the role of determinants such

has focused on factors observed at the country level. One of the main value-added of the article is that we conduct our analysis with universities as the destination. While a cross-country analysis is important to understand the reasons for the uneven distribution of students across destination countries, information at the country level conceals significant variation among universities of the same country. For instance, the average national quality of universities might not accurately reflect the attractiveness of the country as a provider of tertiary education. Foreign students might concentrate, for instance, on the upper tier of universities in the country. The distribution of foreign students across Italian Universities confirms that it is definitely the case. Therefore, the fact that a country hosts many universities of relatively modest quality might not be an important factor, at least for explaining inflows of foreign students to that country. This in turn stresses the need to use information at the university level. The same applies to fees. The average level of fees might not mean anything for students since they might end up relatively good universities charging relatively higher fees. To overcome this limitation, we study the role of these factors, observed at the university level. Our econometric analysis allows to identify very precisely the various university-specific factors that lead students to choose among Italian institutions.[3] Such an investigation is unique in the literature in that respect.

Our article is also related to the literature investigating the response of students at the higher education to variation in tuition fees. While there are few studies investigating the impact of fees on inflows of foreign students using university data, some papers look at the behavior of native students. Alecke and Mitze (2013) study how an increase in the level of tuition fees charged in Germany affected the internal mobility of students. Bruckmeier and Wigger (2015) address the same increase, focusing on how it relates to the time of graduation. More closely related to our purpose, another strand of the literature looks at the impact of tuition fees on the enrollment of native students. Using individual data on enrollment and data on fees by institution and subject, Wales (2013) looks at the effect of tuition fees on the demand for postgraduate education in the UK. He finds that on average a 10% increase in tuition fees is associated to a reduction in the probability of continuing to a postgraduate degree of between 1.7% and 4.5%. Dearden et al. (2011) also find a negative impact on undergraduate enrollment. These findings are mirrored in the US case. Leslie and Brinkman (1987) find that students respond to price signals in higher education institutions. Carneiro and Heckman (2002) show that credit constraints affect higher education participation in the US. Nevertheless, some papers find contrasting results of the impact of education costs. Flannery and O'Donoghue (2009) find no evidence that tuition fees reduce undergraduate participation in Ireland. Chapman and Ryan (2005) find similar results for Canada. Christofides et al. (2010) find that tuition fees have a small negative impact on

as networks, quality and fees in explaining the extent of the bilateral flows of foreign students. Regarding fees, while they fail to identify a negative impact of tuition fees, they do show that the positive impact of fees obtained in '*naïve*' regressions might be due to endogeneity. Other interesting papers of the literature using dyadic flows include Abbott and Silles (2015), Jena and Reilly (2013), González et al. (2011), Kahanec and Králiková (2011). Gravity models have also been used to explain student mobility between regions of the same country. See for instance Agasisti and Dal Bianco (2007) for Italy. Alecke and Mitze (2013) and Bruckmeier et al. (2013) exploited German data and give a special attention to the role of tuition fees.

3 These estimated factors could be further used in the estimation of the full decision tree involving the choice of the destination country and the decision to study abroad.

student demand in Australia. More directly related to our work, Soo and Elliott (2010) find a minor role of fees for application from international undergraduate students in a selection of British universities. Our article aims at revisiting this issue in the specific case of foreign students, using aggregate flow data varying across universities and across countries of origin. Importantly, our data span the whole set of Italian universities. Our results provide results that complement the mixed evidence coming from the use of micro data. The emphasis on foreign students allows to assess to what extent the response of international students compares with the one of native students.

Our results shed some new light on the literature on the determinants of foreign students. Regarding the role of tuition fees, we first stress the need to deal with the endogeneity of these fees by isolating their impact on the location choice of foreign students. When dealing with this issue, we find that tuition fees have a more negative effect on the choice of a specific university, a result which is definitely new to the existing literature on foreign students. Our preferred estimates based on Poisson regression and IV estimation suggest that a 10% increase of tuition fees results in a decrease of inflows of about 8%. Compared with the results of Wales (2013), our estimate suggest that foreign students might exhibit a slightly higher response to fees than native ones. On top of that, we find support for the role of the quality of the university, a result already found in previous work (Beine et al., 2014; Van Bouwel and Veugelers, 2013). We also find that the host capacity of the university as well as the expected return on education in the city where the education is acquired are important, in line with the spirit of the migration model of foreign education (Rosenzweig, 2008). Our results are robust to a set of robustness checks that involve deviations from the exclusion restriction, alternative measures of tuition fee and university quality, the fact that some universities have multiple campus and the use of alternative econometric specification.

The article is structured as follows. Section 2 presents a theoretical model that is useful for deriving the estimable gravity equations. Section 3 is devoted to the exposition and clarification of the data that we use in the econometric estimations. Section 4 presents the estimable gravity equations and discusses the main econometric issues, including the treatment of the zeros for the dependent variable and the way we deal with endogeneity issues. Section 5 presents the results while Section 6 concludes.

2. Theoretical background

In order to analyze the role of tuition fees on the location choice of foreign students, we derive a tractable students' migration equation from a simple theoretical model based on the human capital literature and on the RUM approach (McFadden, 1984). Education is considered an investment in future earnings and employment (Becker, 1964) for rational students who seek to maximize their lifetime earnings. The quality of education may affect their expected returns to education (Card and Krueger, 1992). The prospective student migrant compares the present value of future earnings if he or she decides to study in a university at home with the present value of future earnings if education is obtained at a university abroad. If the increase in the present value of the future income is greater than the cost of migrating, plus other education costs, the student decides to move to the university yielding the highest net present value. Nevertheless, this is conditional because each university might face capacity constraints or impose quotas on foreign students.

2.1. Students' choice

The set of destination countries is $D = \{d_1, \ldots, d_{n_d}\}$ with n_d the number of destination countries (d is the index for destination country) and the set of origin countries is $O = \{o_1, \ldots, o_{n_o}\}$ with n_o the number of origin countries (o is the index for the origin country). Countries can be inside both D and O. The set of universities in country d is $U^d = \{u_1^d, u_2^d, \ldots, u_{n_u^d}^d\}$ with n_u^d the total number of universities in country d (u^d is the index for university). The set of students in each country o who aspire to undertake studies in higher education is $S^o = \{s_1^o, s_2^o, \ldots, s_{N_s^o}^o\}$, with N_s^o the total number of young people in country o who aspire to study. The index for student is s.

Utility derived from studying in university u^d located in country d of student s from country o (VS^s_{o,d,u^d}) is expressed as:

$$VS^s_{o,d,u^d} = VS_{o,d,u^d}(IW^s_{d,u^d}, CM_{o,d}, CS_{u^d}, CL_{u^d}, A_d) + \epsilon^s_{o,d,u^d} \tag{2.1}$$

where IW^s_{d,u^d} is the intertemporal expected value of labor income after graduating from university u^d, $CM_{o,d}$ a vector of country-pair migration costs; CS_{u^d} the cost of education (here the fees of university u^d); CL_{u^d} the cost of living in the city of university u^d and A_d some country-specific unpriced amenities. One part of utility is deterministic and varies by origin and university destination pair $VS_{o,d,u^d}(.)$. The other part is stochastic and captures unobserved components of the individual utility associated with each university choice (ϵ^s_{o,d,u^d}).

The expected intertemporal labor income of student s from country o studying in university u^d located in country d (IW^s_{o,d,u^d}) is defined by:

$$IW^s_{o,d,u^d} = \int_{\underline{T}^s}^{\overline{T}} e^{-\rho t} W^s_{o,d,u^d}(.)dt \tag{2.2}$$

with $\underline{T}^s$ as the age of student s upon graduating and $\overline{T}$ as a fixed retirement age. $e^{-\rho t}$ is a discount factor with ρ the rate of time preference. Individuals have the same rate of time preference and the same indirect utility functions.[4] $W^s_{o,d,u^d}(.)$ is the annual expected labor income. We suppose that students form myopic expectations about the expected wages by referring to the wages prevailing in the local labor market of the university:

$$IW^s_{o,d,u^d} = \frac{(e^{-\rho\overline{T}} - e^{-\rho\underline{T}})}{\rho} W^s_{o,d,u^d}(.) = B\left(\frac{Q_{u^d}}{\overline{Q}_d}\right)^{\beta_0} w_{u^d} \tag{2.3}$$

with w_{u^d} the value of average earnings in area u^d; Q_{u^d} the quality of education where the higher education has been attained; and $\overline{Q}_d$ the average quality of education in the country d. β_0 is a strictly positive parameter. A positive difference between the quality of education obtained (Q_{u^d}) and the average quality of education in country d ($\overline{Q}_d$) implies a skill premium (the effective earnings will be greater than the local average earnings). $B = \frac{(e^{-\rho\overline{T}} - e^{-\rho\underline{T}})}{\rho}$ is a constant, which implies that the expected intertemporal labor income is not specific to an individual ($IW^s_{o,d,u^d}(.) = IW_{o,d,u^d} = (.)$).

4 In the absence of individual information in our database, we assume thereafter $\forall s\ \underline{T}^s = \underline{T}$.

The deterministic and observable component of utility is assumed to be logarithmic:

$$VS_{o,d,u^d} = \ln\left(\frac{(IW_{o,d,u^d})^{\beta_1} A_d^{\gamma_1}}{\delta_{o,d,u^d}}\right) \tag{2.4}$$

with $\delta_{o,d,u^d} > 1$ an iceberg total cost factor ($\delta_{o,d,u^d} = \delta(CM_{o,d}, CS_{u^d}, CL_{u^d})$). Migration from country o to university u^d in country d involves country-pair specific costs and localization-specific costs that reduce utility in an iceberg-type way.

We assume that the migration costs depend only on the destination country and not on the specific location within the country. These migration costs, $CM_{o,d}$ are composed of two parts: fixed costs (C_o) and variable costs ($C_{o,d}$). The fixed part measures the costs of moving, independent of the destination country (home-specific costs) whereas the variable migration costs depend on dyadic factors such as physical distance $d_{o,d}$, the cultural and linguistic proximity of the origin and destination countries, such as the use of a common official language ($l_{o,d}$) or the existence of colonial links ($col_{o,d}$).

We assume a fairly simple specification of the total factor cost δ_{o,d,u^d}:

$$\begin{aligned} \ln(\delta_{o,d,u^d}) = \gamma_2 \ln(C_o) + \alpha_1 \ln(d_{o,d}) + \alpha_2 \ln(l_{o,d}) + \alpha_3 \ln(col_{o,d}) \\ + \beta_3 \ln(CS_{u^d}) + \beta_4 \ln(CL_{u^d}) - \beta_5 \ln(E_{o,u^d}) \end{aligned} \tag{2.5}$$

We then have:

$$\begin{aligned} VS_{o,d,u^d} = {} & \ln(B) + \beta_2 \ln(Q_{u^d}) - \beta_2 \ln(\overline{Q}_d) + \beta_1 \ln(w_{u^d}) + \gamma_1 \ln(A_d) - \gamma_2 \ln(C_o) \\ & -\alpha_1 \ln(d_{o,d}) - \alpha_2 \ln(l_{o,d}) - \alpha_3 \ln(col_{o,d}) - \beta_3 \ln(CS_{u^d}) - \beta_4 \ln(CL_{u^d}) \end{aligned} \tag{2.6}$$

with $\beta_2 = \beta_0\beta_1$.

Following the random utility approach to discrete choice problems (McFadden, 1984), the probability that student s from country o chooses university u^d in country d is defined by:

$$\begin{aligned} & P_{o,d,u^d} = \text{Prob}[VS^s_{o,d,u^d} > VS^s_{o,i,u^i}] = \text{Prob}[VS_{o,d,u^d} - VS_{o,j,u^i} > \epsilon^s_{o,i,u^i} - \epsilon^s_{o,d,u^d}], \\ & \forall u^i \neq u^d \text{ and } \forall \text{i} \in \text{D} \end{aligned} \tag{2.7}$$

with ϵ being an iid extreme-value distributed random term.

Following Train (2003), this probability can be decomposed in three logits (a decision tree with three levels). In the upper level, the student decides whether to study at home or abroad. This upper-level utility depends on factors that vary with the choice of migrating or staying, $Z_h = \{C_o, A_d\}$. If the decision is to move abroad, there is a subsubset of destination countries from which the student must choose his or her location (middle level of the tree). The determining variables for this middle-level utility depends on factors that vary across countries, $Y_{o,d} = \{\overline{Q}_d, d_{o,d}, l_{o,d}, col_{o,d}\}$. At the lower level, the student chooses the university where he or she would like to study. The lower-level utility depends on characteristics that vary across universities, $X_u = \{Q_{u^d}, w_{u^d}, CS_{u^d}, CL_{u^d}\}$.

Utility can be rewritten as:

$$VS^s_{o,d,u^d} = \ln(B) + VS_h(Z_h) + VS_{o,d}(Y_{o,d}) + VS_u(X_{u^d}) + \epsilon^s_{o,d,u^d} \tag{2.8}$$

with

$$\begin{aligned}
VS_u(X_{u^d}) &= \beta' \ln(X_{u^d}) = \beta_2 \ln(Q_{u^d}) + \beta_1 \ln(w_{u^d}) - \beta_3 \ln(CS_{u^d}) - \beta_4 \ln(CL_{u^d}) \\
VS_{o,d}(Y_{o,d}) &= \alpha' \ln Y_{o,d} = -\alpha_1 \ln(d_{o,d}) - \alpha_2 \ln(l_{o,d}) - \alpha_3 \ln(col_{o,d}) \\
VS_h(Z_h) &= \begin{cases} \gamma' \ln Z_d = \gamma_1 A_d - \gamma_2 \ln(C_o) & \text{if moving} \\ \gamma' \ln Z_s = \gamma_1 A_o & \text{if staying} \end{cases}
\end{aligned} \tag{2.9}$$

where β', α' and γ' denote parameter vectors.

With this decomposition of utility, the probability associated with Equation (2.7) can be written as the product of three standard logit probabilities:

$$P_{o,d,u^d} = P_{o,u^d|d,h} P_{o,d|h} P_{o,h} \tag{2.10}$$

where $P_{o,u^d|d,h}$ is the conditional probability of choosing a university u^d given that an alternative in subnest d is chosen; $P_{o,d|h}$ is the conditional probability of choosing a country d, given that an alternative between staying or moving is chosen; and $P_{o,h}$ is the unconditional (marginal) probability of choosing to study in a foreign country or in home country o. These probabilities can be expressed as:

$$P_{o,u|d,h} = \text{Prob}[VS_{o,d,u^d} - VS_{o,d,u^i} > \epsilon^s_{o,d,u^i} - \epsilon^s_{o,d,u^d}] = \frac{\exp(VS_u(X_{u^d}))}{\exp I^u(d,h)}, \quad \forall u^i \neq u^d \tag{2.11}$$

for the conditional probability $P_{o,u|d,h}$, and

$$\begin{aligned}
P_{o,d|h} &= \text{Prob}[VS_{o,d,u^d} - VS_{o,j,u^d} > \epsilon^s_{o,j,u^d} - \epsilon^s_{o,d,u^d}] \\
&= \frac{\exp(VS_{o,d}(Y_{o,d}) + (1-\lambda^u) I^u(d,h))}{\exp I^d(h)}, \quad \forall j \neq d
\end{aligned} \tag{2.12}$$

for the conditional probability $P_{o,d|h}$. This conditional probability for the degenerate branch (Stay branch), $P_{o,h|s}$, is trivially equal to 1 (a partially degenerate nested logit). And, for the unconditional (marginal) probability:

$$\begin{aligned}
P_{o,h} &= \text{Prob}[VS_{o,h,u} - VS_{o,k,u} > \epsilon^s_{o,k,u} - \epsilon^s_{o,h,u}] \qquad \text{with k} \neq \text{h} \\
&= \frac{\exp(VS_h(Z_h) + (1-\lambda^j) I^j(s))}{\exp(VS_h(Z_h) + (1-\lambda^j) I^j(s)) + \exp(VS_h(Z_k) + (1-\lambda^j I^j(m))}
\end{aligned} \tag{2.13}$$

The inclusive values I^u and I^j are defined by:

$$I^u(d,h) = \ln\left(\sum_{u=1}^{n_u^d} \exp(VS_u(X_u))\right) \tag{2.14}$$

$$I^j(h) = \ln\left(\sum_{j=1}^{n_d} \exp\left(VS_{o,j}\left(Y_{o,j}\right) + (1-\lambda^u) I^u(j,h)\right)\right) \tag{2.15}$$

The nested multinomial logit model defined by Equations (2.10)–(2.13) connects the levels of the decision tree with each other in the sense that the attributes of the lower-branch alternatives influence the choice among any choice set of upper branches. In a sequential choice model, the levels of the hierarchy would be unrelated.

The aggregate multi-country migration flow equation to university u^d is given by multiplying the number of young people in country o who want to study (N_s^o) with the probability of migration to university u^d of a randomly drawn student of country o (P_{o,d,u^d}):

$$M_{o,d,u^d} = P_{o,d,u^d} N_s^o = P_{o,u^d|d,m} P_{o,d|m} P_{o,m} N_s^o \tag{2.16}$$

with M_{o,d,u^d} as the number of young people from country o who want to study at university u^d located in country d. It follows that the total number of foreign young people who wish to study at university u^d located in country d is given by:

$$M_{d,u^d} = \sum_{o \neq d} P_{o,d,u^d} N_s^o = \sum_{o \neq d} P_{o,u^d|d,m} P_{o,d|m} P_{o,m} N_s^o \tag{2.17}$$

However, as already stated, this number (M_{d,u^d}) is not the number of foreign students who will be enrolled in university u^d, this is the number of foreign students who want to go on to study at university u^d. We call this the *ex ante* enrollment demand. It is not enough that students wish to go to this university, the university must also allow their registration. Universities have enrollment policies or constraint capacity that can lead to the number of foreign students enrolled being lower than M_{d,u^d}. To know the actual number of foreign students enrolled we need to model universities' enrollment behavior.

2.2. Universities' behavior

We assume that all universities have the same enrollment behavior and, in the short term, it is determined by three factors: the capacity for enrolling foreign students which is constrained [this capacity, $EC_{u^d}^{\beta_5}$, is a share (defined by β_5) of the total enrollment capacity EC_{u^d}], the university quality (Q_{u^d}) and the fees (CS_{u^d}).

Capacity and quality may change over the long term with investment in capital and staff but they are fixed in the short term. Fees in the long run can also be adjusted according to enrollment demand (when they are not regulated). However, these three factors are fixed in the short term. Therefore, the foreign student enrollment capacity could be constrained for university u^d, and the actual number of foreign students ($\tilde{M}_{d,u^d}$) should verify: $\tilde{M}_{d,u^d} = EC_{u^d}^{\beta_5}$, where $\tilde{M}_{d,u^d}$ is the observed allocation, which corresponds to the *ex post* enrollment.

For each university u^d, two configurations are therefore possible: $M_{d,u^d} \leq EC_{u^d}^{\beta_5}$, the *ex ante* enrollment demand for university u^d is lower than its enrollment capacity or $M_{d,u^d} > EC_{u^d}^{\beta_5}$, that implies $M_{d,u^d} > \tilde{M}_{d,u^d} = EC_{u^d}^{\beta_5}$, the *ex post* (observed) enrollment is lower than the *ex ante* demand. In this last case, the constraint is binding, and some students are forced to request enrollment in a university different from their first preference.

It is well known that many universities have turned away applications from foreign students due to capacity constraints, which supports the assumption that some universities are constrained. In that case, the total allocation is also constrained and the choices based only on preferences [defined by the system (2.10–2.13)] differ from the observed *(ex post)* allocation consistent with the preferences and with capacity constraints. We define how this *ex post* allocation can be done.

[9.8.2018–8:26pm] [1–31] Paper: OP-JEGJ180045
Copyedited by: MANUSCRIPT CATEGORY: ORIGINAL ARTICLE

2.3. Equilibrium allocation with enrollment capacity constraints

We follow the allocation solution developed by De Palma et al. (2007).[5] The set of constrained universities is $\mathfrak{C}$ and $\overline{\mathfrak{C}}$ is the set of unconstrained universities, with $\mathfrak{C} \cup \overline{\mathfrak{C}} = U^d$. The existence of a feasible allocation requires the total world enrollment capacity not be binding. It implies:

$$\sum_{o \neq d} \sum_{d} \sum_{u^d} M_{o,d,u^d} < \sum_{d} \sum_{u^d} EC_{u^d}^{\beta_5} \tag{2.18}$$

Any student who wants to study abroad could be enrolled in a university, but not necessarily in his or her preferred university. As we have assumed that at least one university has an enrollment constraint, the *ex post* total allocation is different from the total *ex ante* allocation. The (*ex post*) probability that student s coming from country o is enrolled in university u^d in country d is denoted by $\tilde{P}_{o,d,u^d}$. The *ex post* allocation[6] is given by:

$$\tilde{M}_{o,d,u^d} = \tilde{P}_{o,d,u^d} N_s^o = \tilde{P}_{o,u^d|d,m} \hat{P}_{o,d|m} \hat{P}_{o,m} N_s^o \tag{2.19}$$

De Palma et al. (2007) show that, under two simple assumptions (allocation rules), the allocation probabilities can still be written as a multinomial logit model but with an additional *correction factor* that expresses an individual allocation ratio. This allocation ratio is defined by π_{u^d}, with $\tilde{P}_{o,u^d|d,m} = \pi_{u^d} P_{o,u^d|d,m}$.

Free allocation rule: For an unconstrained university $u^d \in \overline{\mathfrak{C}}$,

$$P(\{s \text{ allocated to } \mathrm{u^d} \ | \text{s prefers } \mathrm{u^d}\}) = 1 \qquad \forall \mathrm{s}, \quad \forall \mathrm{u^d} \in \overline{\mathfrak{C}}$$

No priority rule: The *no priority rule*, concerns the allocation in an *ex post* constrained university. With this rule, if a student s has a stronger preference (*ex ante*) for constrained university u^d than another student s', student s will also have a proportionally greater chance to be allocated *ex post* to this University.

For an *ex post* constrained university, the individual allocation ratio of university u^d, is the same for all students:

$$\frac{\tilde{P}^{s}_{o,u^d|d,m}}{P^{s}_{o,u^d|d,m}} = \frac{\tilde{P}^{s'}_{o,u^d|d,m}}{P^{s'}_{o,u^d|d,m}} = \Phi_{u^d} \qquad \forall s, s' = s_1^o, \cdot, s_{N_s^o}^o, \quad \forall u^d \in \mathfrak{C}$$

Under these two assumptions, De Palma et al. (2007) show that the allocation probabilities are given by the adjusted MNL formula:

$$\tilde{P}_{o,u^d|d,m} = \frac{\exp(VS_u(X_{u^d}) + \ln(\pi_{u^d}))}{\sum_{u=1}^{n_u^d} \left(\exp(VS_u(X_u)) + \ln(\pi_{u^d})\right)}, \quad \text{with} \tag{2.20}$$

5 De Palma et al. (2007) applied this constraint approach on the housing market in Paris. In their analysis, demand is constrained by the existing stock of housing.

6 Without constraints at the country level—for example, with quotas on student visas (that implies a $\tilde{P}_{o,d|m}$) or constraints on students emigration (that implies a $\tilde{P}_{o,m}$)—the formula of $P_{o,d|m}$ and $P_{o,m}$ are not modified by constraints at the university level. However, this does not mean that their values are not affected by capacity constraints at the university level. When they are taken into account, the calculus of the inclusive value $I^u_{d,h}$ is also modified, and, therefore, the values of $P_{o,d|m}$ and $P_{o,m}$. These new values are represented by $\hat{P}_{o,d|m}$ and $\hat{P}_{o,m}$.

$$\pi_{u^d} = \begin{cases} \dfrac{EC_{u^d}^{\beta_5}}{M_{o,d,u^d}} < 1 & \text{if } u^d \in \mathfrak{C} \\ \dfrac{1 - \sum_{u\in\mathfrak{C}} \dfrac{EC_{u^d}^{\beta_5}}{M_{o,d,u}} P_{o,u|d,m}}{\sum_{v\in\overline{\mathfrak{C}}} P_{o,v|d,m}} > 1 & \text{if } u^d \in \overline{\mathfrak{C}} \end{cases} \tag{2.21}$$

They propose a solution algorithm for the model. This algorithm could be used in our nested logit model to find the allocation solution and the estimated coefficients with enrollment capacity constraint. The algorithm iteratively estimates the constraints and the individual and aggregate allocation ratios until they converge. While we do not observe $\tilde{M}_{o,d,u^d}$ for each university in the data, we can use this theoretical model and the solution approach proposed by De Palma et al. (2007), for our database for Italy. We do this, both by adding the assumption that all the universities in Italy have their *ex ante* enrollment capacity constrained and by using a sequential estimation procedure.

2.4. Estimable equilibrium equation

The estimation of a nested multinomial logit model can be done by FIML (full information maximum likelihood) or through a sequential procedure.[7] Due to data constraints, the sequential procedure is often favored.[8] For estimating the (constrained) coefficient in the first step, we need to use the iterative procedure proposed by De Palma et al. (2007), which requires us to carry out all the steps. This is because the *ex post* allocation in an *ex ante* non-constrained university in country d can be modified by the reallocation implied by the constraints on universities in country d or other countries. However, this is not possible for our study due to data constraints (we should have all the data of the flows of international students for all the universities in the world). Nevertheless, this limitation can be overcome if we assume that each university in the country of interest (Italy in this study) faces a binding enrollment capacity constraint.[9]

Consequently, if we assume that in country d we have:

$$\sum_{o\neq d} P_{o,u^d|d,m} P_{o,d|m} P_{o,m} N_s^o = M_{d,u^d} > EC_{u^d}^{\beta_5} = \tilde{M}_{d,u^d} \qquad \forall u^d \in U^d \tag{2.22}$$

which implies that

$$\tilde{M}_{d,u^d} = EC_{u^d}^{\beta_5} = \sum_{o\neq d} \tilde{P}_{o,u^d|d,m} \hat{P}_{o,d|m} \hat{P}_{o,m} N_s^o = EC_{u^d}^{\beta_5} \qquad \forall u^d \in U^d$$

7 Even if the estimation procedure is sequential, we have shown that the nested logit model (the decision process) is not sequential but recursive.
8 Such sequential estimation leads to consistent estimates but less efficient that FIML estimation.
9 This hypothesis seems reasonable since universities exhibit a combination of various types of constraints. First, educational institutions face natural production constraints in the form of infrastructures, number of professors and teaching assistants. In Italy, non-EU students need to pass a test of linguistic proficiency, which induces some selection. Furthermore, some disciplines like medicine organize contests to select students.

and

$$\tilde{P}_{o,u^d|d,m} = \frac{\exp(VS_u(X_{u^d}) + \ln(\pi_{u^d}))}{\sum_{u=1}^{n_u^d}\left(\exp(VS_u(X_u)) + \ln(\pi_{u^d})\right)}, \quad \text{with} \tag{2.23}$$

$$\pi_{u^d} = \frac{EC_{u^d}^{\beta_5}}{M_{d,u^d}} \qquad \forall u^d \tag{2.24}$$

With this allocation rule, Equation (2.19), which determines the *ex post* number of students coming from country o and studying in university u^d in country d, is written as:

$$\tilde{M}_{o,d,u^d} = \tilde{P}_{o,u^d|d,m}\hat{P}_{o,d|m}\hat{P}_{o,m}N_s^o = \pi_{u^d}P_{o,u^d|d,m}\hat{M}_d^o = \frac{EC_{u^d}^{\beta_5}}{M_{d,u^d}}\frac{\exp(VS_u(X_{u^d}))}{\sum_{u=1}^{n_u^d}\exp(VS_u(X_u))}\hat{M}_d^o \tag{2.25}$$

with $\hat{M}_d^o$ being the number of students who would like to study in country d, taking into account the capacity constraints.

Taking logs of Equation (2.25) and substituting VS_u by Equation (2.9), we obtain the following structural gravity equation:

$$\begin{aligned} \ln(\tilde{M}_{o,d,u^d}) &= \beta_1\ln(w_{u^d}) + \beta_2\ln(Q_{u^d}) - \beta_3\ln(CS_{u^d}) - \beta_4\ln(CL_{u^d}) \\ &+\beta_5\ln(EC_{u^d}) - \ln(M_{d,u^d}) - \ln(\sum_{u=1}^{n_u^d}\exp(VS_u(X_u))) + \ln(\hat{M}_d^o) \end{aligned} \tag{2.26}$$

Before proceeding to the econometric specification corresponding to Equation (2.26), some comments are in order. First, β_5 is the average propensity of all universities to apply the capacity constraint to foreign students. Theoretically, this average propensity should be between 0 and 1. Second, the term $\ln(\sum_{u=1}^{n_u^d}\exp(VS_u(X_u)))$ does not vary across universities and will be captured by the constant. Third, $\hat{M}_d^o$ is specific to the origin country and could be included in a fixed effect controlling for all factors that are specific to the foreign student's country of origin. Finally, $\ln(M_{d,u^d})$, the *ex ante* demand from foreign students to each university of country d is not observed by the econometrician. We will therefore discuss the implications of its omission in the context of IV estimation.

3. Data and descriptive statistics

This section presents the key data used to estimate Equation (2.26) capturing the relationship between foreign student inflows and tuition fees in Italy. Section 3.1 presents the data on international students flows. Section 3.2 describes our tuition fees measure. Section 3.3 reports information on the other covariates. For further information and comprehensive descriptive statistics we invite the interested reader to refer to the Web Appendix.

3.1. International students flows

To measure $\tilde{M}_{o,d,u^d}$ in Equation (2.26), we take advantage of the data on bilateral flows of international students from all countries of the world to Italy for the academic year 2011–2012. Following Beine et al. (2014), the international students we consider are the ones who migrated exclusively for the sake of education. Those who spent either one or more semesters abroad in institutional programs, such as the ERASMUS students, do not comply with our definition of international students and are therefore excluded from the data. We omit these students from the analysis for two reasons. First, bilateral agreements constrain the student's choice in terms of location. Second, in some curricula, attending a period of study abroad can be compulsory.

The statistical office of the Italian Ministry of Education (MIUR) produces data on foreign students for 88 universities.[10] When conducting our empirical estimation we do not include telematic universities (10) and institutions devoted to domestic education only (7). Therefore our final sample consists of 71 universities. For further details, see Section A.1 of the Web Appendix.

Table 1 reports some descriptive statistics on the number of foreign students. Italy is not a major destination for international students who represent on average 3.79% of the total student population.[11] These students originate from 173 different countries.[12] Since the data are dyadic in nature, it is not easy to describe their structure concisely. Nevertheless, Table 8, in Section C of the Web Appendix, reports for each university the number of foreign students along with the names of the countries sending more students to each university.

Section A of the Web Appendix provides comprehensive descriptive statistics on the inflows of international students to Italian universities.

3.2. Tuition fees

The cost of education CS_{u^d} in Equation (2.26) is captured by the level of tuition fees. The Italian system of higher education has distinctive features in terms of tuition fees compared with other countries. First, Italy is one of the few European continental countries in which tuition fees vary across institutions. Second, while the tuition fees within one university can vary across fields of study, this variation is very limited. Third and interestingly, there is no price discrimination between native students and foreign students.[13]

Italian universities are classified either as private or public institutions. In contrast to most continental European countries, tuition fees charged by Italian public universities are not uniformly determined by the central government. According to the Italian law (Decree of the President of the Republic of 25 July 1997, No. 306), the total amount of fees collected by a public university cannot exceed 20% of the funding received by this

10 Data source: MIUR, Indagine sull'istruzione universitaria.
11 By comparison, for the same year, the percentage of international students in French universities was 13.3%.
12 In the empirical part, we pay attention to not loosing the information relative to the empty corridors, that is, origin-destination pairs with zero migration flow. The total number of observations is then equal to the number of universities multiplied by the number of origin countries.
13 Namely, Italian institutions do not charge higher tuition fees for non-European students. Only other five European countries treat equally non-European students: the Czech Republic, Hungary, Iceland, Liechtenstein and Norway (European Commission, 2012).

Table 1. Descriptive statistics of foreign student flows (2011)

Number of universities (a)[a]	71
Origin countries (b)	173
Number of observations ($a \times b$)	12,283
Percentage of zeros[b]	65.2
Total number of students (host capacity)[c] (c)	1,681,742
Number of foreign students[c] (d)	63,669
Percentage of foreign students as share of total students[c] (d/c)	3.79

aTelematic universities and universities hosting <15 students not included.
bFlow of students from country i studying in univ. u is zero.
cNumbers are computed aggregating all origin countries.

university from the MIUR. Conversely, for Italian private institutions, this 20% limit does not apply, and they do charge higher fees. Tuition fees in Italian public universities depend on many determinants, in particular, on the student's family income and on the year of enrollment.

Our primary source of data on (average) tuition fees is based on a survey conducted by the economic newspaper '*Il Sole 24 Ore*'.[14] Section B.3 of the Web Appendix shows the distribution of tuition fees for Italian universities. Only private institutions charged average fees above the level of €2000.

3.3. Other covariates

The estimation of Equation (26) requires data on the specific destination (the city where the university is located) such as the cost of living and the expected income along with variables capturing universities' characteristics such as their quality. Table 2 provides a quick summary of the data used to estimate the benchmark regression. Sections B.1 and B.2 of the Web Appendix present the data used to proxy the cost- of-living and the expected-income at destination, respectively.

Section B.4 of the Web Appendix provides information on the construction of our benchmark quality indicators, which relies on the Shangai Top-Ranking. Section 5.2.3 shows that our results are robust to a different quality indicator.

The theoretical model highlights the relevance of capacity constraint. Section (B.5) of the Web Appendix details information on the size of Italian universities. Further, when conducting the empirical estimation we control for the availability of English-teaching programs (see Section B.6 of the Web Appendix). Finally, in Section C of the Web Appendix, Table 7 reports, for each institution, the value of the main variables involved in our empirical analysis.

14 See *Il Sole 24 Ore*, 18/07/2011-N.194. The data were obtained from the MIUR. To check the robustness of our econometric results we also use alternative measures of fees, coming from the ETER database. See Section 5.2.2.

Table 2. Definitions and sources of main variables

Variable	Term in Equation (26)	Definition	Source
International students	$\tilde{M}_{o,d,u^d}$	Number of foreign students coming from country *i* and enrolled in university *u*	MIUR
Fees	CS_{u^d}	Average fees charged by university *u*	Newspaper *Il Sole24 ore*
Quality	Q_{u^d} (ranking)	Quality of university *u* based on Top 500 ranking	Top 500 Shanghai Ranking ARWU
Host capacity	EC_{u^d}	Total number of students enrolled at university *u*	MIUR
Cost of living	CL_{u^d}	Cost of living in city/district *j*, where institution *u* is located	Numbeo dataset
Expected return	w_{u^d}	GDP per capita in the district where university *u* is located	GDP at NUTS 3 level, Eurostat
English programs	Eng_{u^d}	Existence of of teaching programs in English at university *u*	Conference of Italian University Rectors (Fondazione Crui)

AQ10

4. Econometric specification

4.1. From theory to econometric specification

Our econometric specification is based on Equation (2.26) that provides the determinants of choosing a specific university, conditionally upon studying abroad in a specific destination country (i.e., Italy). The benchmark estimated equation takes the following form:

$$\ln(\tilde{M}_{o,d,u^d}) = \alpha + \alpha_d + \beta_1 \ln(\exp return_{u^d}) + \beta_2 \ln(quality_{u^d}) + \beta_3 \ln(fees_{u^d}) + \beta_4 \ln(living\cos t_{u^d}) + \beta_5 \ln(hostcapacity_{u^d}) + \beta_6 English_{u^d} + \epsilon_{d,u^d} \tag{4.1}$$

where $\tilde{M}_{o,d,u^d}$ denotes the observed number of students coming from country o and studying in university u^d in country d. As noted above, this is applied to one specific academic year, 2011–2012. The data are, therefore, dyadic and time-invariant in nature.

$fees_{u^d}$, $living\cos t_{u^d}$, $quality_{u^d}$, $hostcapacity_{u^d}$ and $\exp return_{u^d}$ stand, respectively, for CS_{u^d}, CS_{u^d}, Q_{u^d}, EC_{u^d} and w_{u^d} in Equation (2.26). α_d is a set of fixed effects controlling for all factors specific to the country of origin of the foreign students. It includes $\ln(\hat{M}^o_d)$ in Equation (2.26). Given that we focus on a specific country, i.e., Italy, α_d also controls for bilateral factors between the origin country and the university. α is a constant term that includes the theoretical term $\ln(\sum_{u=1}^{n_u^d} \exp(VS_u(X_u)))$ from Equation (2.26) that does not vary across institutions. ϵ_{d,u^d} is an error term that is assumed to be independently and identically distributed.[15]

15 The estimation is not subject to important issues of multicollinearity. The respective correlations between the (log of) fees and (the log of) the other covariates amount to −0.044 with the score, −0.042 with the ranking, 0.274 with the cost of living, 0.544 with the expected income, −0.381 with the capacity and −0.018 with the existence of English programs.

One concern related to the previous specification is that it neglects the existence of teaching programs provided in English at the destination university. Given the importance of English as an international language, the existence of such programs can be a determinant for foreign students in their location and enrollment choice.[16] Furthermore, it is possible that universities with English teaching programs can display characteristics different from other universities, prompting some correlation with other covariates such as the quality or the fees. If it is the case, the specification derived from the theory might be problematic.[17] To address such a concern, we extend specification (4.1) by including a dummy variable, $English_{u^d}$, capturing the availability of English teaching programs at the university level.

Before we proceed to the estimation, a couple of comments are in order. First, we make clear that Equation (4.1) corresponds to the last stage of the migration process of foreign students. Previous stages concern (i) the decision to study abroad or domestically and (ii) the choice of the country of destination. This article focuses only on the last stage.[18] Beyond the limitations in data availability that prevent the estimation of the full decision tree, this is not desirable for several reasons. The main objection is that pooling universities of different countries would lead to a clear rejection of the IIA hypothesis implicit in the estimation of Equation (4.1). The rejection of the IIA hypothesis would occur because the choice structure involves two countries that might be considered as nests in the decision process. Given that it is very likely that the degree of substitution between two universities varies with respect to the country of destination, we prefer in the end to estimate the model separately for each country of destination. This issue is also related to the well-known problem of multilateral resistance of migration (Bertoli and Fernández-Huertas Moraga, 2013; Beine et al., 2015). In other words, pooling several countries and integrating the choice of the destination country would entail the estimation of a nested logit model with two potential nests. This is obviously beyond the scope of this article and is left for future investigation.

Second, Equation (4.1) omits the term $\ln(M_{d,u^d})$ in Equation (2.26) which is unobservable. This term indeed captures the total demand to university u^d coming from all origin countries before the impact of the constraints associated with the educational capacities. While in theory this is observable for each university, it is not available to the econometrician and will be included in the error term. This in turn might lead to estimation biases that we will discuss in the identification strategy, especially in the IV procedure. See Section 4.3.

4.2. Econometric method

Another issue is the prevalence of a high percentage of zero values for the bilateral migration flows. In our sample, for the year 2011 under investigation, we have 65.2% of zero values for the bilateral flow of foreign students to Italy. The presence of a high

16 Interestingly, Kahanec and Králiková (2011) find that the availability of English teaching programs acts as a pull factor.
17 Note that, if the correlation between availability of English teaching courses is positively correlated with either education quality or tuition fees, this would lead to upward biased coefficients.
18 In fact, our estimation can be seen as the first step in estimating the full decision tree in a sequential way. The inclusive values of this first stage could be recovered and use in the estimation relative to the two other levels.

proportion of zero values is well-known to generate biases in the key estimates using traditional panel fixed-effect estimates (Silva and Tenreyro, 2006). The use of $\log(1+\tilde{M}_{o,d,u^d})$ as the dependent (the so-called scaled OLS) allows us to solve the selection problem due to the drop of the zero observations. Nevertheless, the scaled OLS estimation technique would give inconsistent estimates in the presence of hetero-skedasticity. Silva and Tenreyro (2006) show that Poisson regressions are robust to different patterns of heteroskedasticity. We follow this route in the subsequent estimation and use the Poisson estimates as the benchmark. However, our tables will report the scaled OLS estimates of Equation (4.1) for robustness checks.[19]

4.3. Dealing with endogeneity concerns

In the model of Section 2, tuition fees are exogenous and decided by university authorities independent of numbers of students or other characteristics. In reality, the exogenous nature of fees in specification Equation (4.1) is questionable on several grounds. First, fees might depend on the attractiveness of the university: successful universities attracting a large number of (foreign) students can easily raise the tuition fees compared with other universities. This leads to a reverse causality issue between student flows and fees. While the bilateral nature of $\tilde{M}_{o,d,u^d}$ mitigates this aspect, it is important to deal with the potential endogeneity of fees.[20]

On top of that, fees might be correlated with some unobserved characteristics of the university such as the quality of amenities on campus or in the hosting city. Another possibility is that universities set quotas for foreign students that are unknown to the econometrician. This can in turn lead to a quantity-price trade-off and induce a positive correlation between fees and quotas. The source endogeneity of tuition fees also calls for a specific treatment.

We deal with the endogeneity of fees by using a traditional IV approach. Basically, we use the public vs private status of the university as an instrument of tuition fees, following a similar solution adopted in Beine et al. (2014) at the country level. In particular, we create and use a dummy variable that captures the status (private vs public) of the university. The underlying assumption is that private universities have a higher control over tuition fees. They tend to increase fees not only because of the costs but also because they receive fewer subsidies. Furthermore, they are not constrained by the regulation in terms of the cap that applies to public universities. We should expect a positive correlation between the private status and the level of tuition fees. In terms of exclusion restriction, the underlying assumption is that foreign students should not have particular preferences for private or public universities on top of the quality of education, host capacity, cost of living and income at the destination area. This seems

19 Alternative estimation methods like the double hurdle model can also handle the high proportion of zeroes in the dependent variable, but are less directly related to the theoretical mechanisms presented in Section 2. A second important advantage of the Poisson method is that it can handle endogeneity issues through Poisson IV estimation.

20 Another way of looking at this endogeneity problem is contained in Equation (2.26). In fact, the fee level (CS_{u^d}) in each university is likely to be positively correlated with the *ex-ante* total foreign demand M_{d,u^d}, which is omitted from Equation (4.1).

to be a reasonable assumption and is confirmed by examples of many successful public universities in the US such as Berkeley or Michigan State University.[21]

5. Results

We first present the benchmark results. We then consider various robustness checks.

5.1. Benchmark regressions

The inclusion of origin-country fixed effects allows us to control for the role of the usual push factors (for instance, GDP at origin) as well as the influence of bilateral determinants (colonial links, proximity, languages). The estimates reported in Table 3 are in line with a traditional view of the role of fees and of quality. In particular, both types of estimation techniques deliver a negative and significant role for fees in the choice of a university, in line with the view that fees are part of the cost function of foreign education. Estimates vary little with respect to the two quality indexes.

Nevertheless, in spite of these promising results, a couple of comments are in order. First, while fees appear to have a negative role, failure to account for their possible endogeneity leads us to take these results with caution. Second, while the benchmark results suggest significant and intuitive roles for fees, the quality of the university, host capacity and the expected income in the area, we fail to find some robust evidence of a role for the cost of living. While the Poisson estimations deliver the expected negative effect of cost of living, but with a low level of significance, the Scaled OLS estimates are insignificant. Since all estimates are potentially biased by the presence of endogenous fees, it is also important to check whether this result survives after an explicit treatment of endogeneity through IV estimates. These are reported in Table 4.

The estimates of Table 4 provide interesting insights. First, the use of IV estimation leads to a significant correction in the estimate of the influence of tuition fees. Endogeneity of fees might be due either to reverse causality (i.e., attractive universities are more likely to charge higher fees) or to some positive correlation of fees with unobserved factors of attractiveness (e.g., universities with better amenities tend to charge higher fees). In both cases, this results in a positive correlation between fees and the error term of model Equation (4.1), resulting in an upward biased estimate of the impact of tuition fees. A comparison of Tables 3 and 4 shows that the use of instrumentation corrects the bias in the expected direction, with a more negative impact of fees on the university choice. This holds for both estimation techniques.

Second, the IV results lead to a significant change in all the estimates of the determinants of the choice of university except for quality. Correcting the impact of fees could suggest that the non-IV Poisson estimate tends to overestimate the true impact or, in other words, underestimate the impact in absolute terms. Such a bias is consistent with, for instance, a positive correlation between fees and unobserved amenities. It is also consistent with a phenomenon of reverse causality (attractive universities are more expensive). The IV estimates of Equation (4.1) support the role of all possible determinants of the model, suggesting that the choice of a particular university results

21 In a robustness check, we look at the impact of reasonable deviations from the exclusion restriction on the estimation of the effect of the fees. See Section 5.2.1.

Table 3. Determinants of location of foreign students: benchmark estimation

	(1) Scaled OLS	(2) Poisson	(3) Scaled OLS	(4) Poisson
Fees	−0.059***	−0.184**	−0.062***	−0.187***
	(0.01)	(0.06)	(0.01)	(0.06)
Cost of living	−0.017	−0.772*	−0.078	−0.898**
	(0.06)	(0.41)	(0.06)	(0.40)
Quality (ranking)	0.077***	0.127***		
	(0.01)	(0.02)	–	–
Quality (score)	–	–	0.111***	0.225***
			(0.01)	(0.04)
Host capacity	0.152***	0.513***	0.157***	0.499***
	(0.01)	(0.06)	(0.01)	(0.06)
Income	0.634***	1.584***	0.665***	1.606***
	(0.03)	(0.16)	(0.03)	(0.15)
English	0.045***	0.346***	0.053***	0.384***
	(0.01)	(0.09)	(0.01)	(0.09)
Origin FE	Yes	Yes	Yes	Yes
R^2	0.565	–	0.564	–
Pseudo R^2	–	0.742	–	0.743
Number of Obs	12,283	12,283	12,283	12,283

Standard errors in parentheses. For scaled OLS, robust standard errors.
$^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.001$.

Table 4. Determinants of location of foreign students: IV estimation

Variables	(1) Scaled IV	(2) Poisson IV	(3) Scaled IV	(4) Poisson IV
Fees	−0.246***	−0.853***	−0.245***	−0.791***
	(0.02)	(0.18)	(0.02)	(0.17)
Cost of living	−0.238***	−2.470***	−0.300***	−2.314***
	(0.06)	(0.71)	(0.06)	(0.64)
Quality (ranking)	0.081***	0.139***	–	–
	(0.01)	(0.03)		
Quality (score)	–	–	0.120***	0.248***
			(0.01)	(0.04)
Host capacity	0.104***	0.398***	0.108***	0.386***
	(0.01)	(0.06)	(0.01)	(0.06)
Income	0.914***	2.787***	0.939***	2.637***
	(0.04)	(0.41)	(0.04)	(0.37)
English	0.071***	0.383***	0.079***	0.432***
	(0.01)	(0.10)	(0.01)	(0.09)
Origin FE	Yes	Yes	Yes	Yes
R^2	0.553	–	0.553	–
F First stage	4697.89	–	4677.66	–
Number of Obs	12,283	12,283	12,283	12,283

Standard errors in parentheses. For scaled IV, robust standard errors.
$^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.001$.

from a complex assessment of benefits and costs as outlined in the theoretical RUM framework of Section 2. Interestingly, the estimates suggest that foreign students explicitly take into account the cost of the living and the expected income for the city of destination. The estimated elasticity suggests that a 10% increase in the tuition fee tends to decrease the average bilateral flow to that university by about 8%.

5.2. Robustness checks

While the results suggest that fees exert a negative impact on the attractiveness of a university for foreign students, we should keep in mind that these findings rely on a set of assumptions or at least methodological choices. In this section, we assess the robustness of the results with respect to a set of important choices we have made in the benchmark analysis. These concern the validity of the instrument in the IV estimations, measures of fees and quality of universities, the existence of multiple campuses in some universities and the specific econometric specification underlying the econometric estimation.

5.2.1. Deviations from the exclusion restriction

The exclusion restriction of our IV might be subject to discussion. While we control for a set of determinants such as host capacity and quality, it could be that some foreign students take into account the status of the university when choosing a location. For instance, it could be that foreign students believe that private universities are better organized and provide better services for students in terms of advice, personal tutoring and other aids. It could also be that students believe private universities are more accountable to students for the quality of teaching. The greater attractiveness of private institutions seems to be the prevailing dominant view. Nevertheless, this view is not the only one. For instance, it might be expected that there is a higher recognition of degrees conferred by public universities, suggesting that the private status of some institutions might deter more than attract some students. In that case, there might be a positive or negative correlation of our status variable and the error term of Equation (4.1), invalidating the exclusion restriction of the IV procedure.

To cope with such a concern, we conduct a new econometric procedure introduced by Conley et al. (2012) that accounts for possible deviations from the exclusion restriction. The idea is to consider the parameter capturing that restriction (the coefficient of the IV in the structural equation) as a random parameter drawn for a given distribution. The procedure allows for possible means different from zero, i.e. for asymmetric deviations from the exclusion restrictions (See Conley et al., 2012, for details).[22] We consider two alternative procedures. The first one, named 'union of confidence interval' (UCI), provides an alternative IV estimation assuming only a support for the exclusion parameter. The other one, called 'local to Zero estimation' assumes a normal distribution with a given mean and standard deviation. Table 5 reports the results of the UCI procedure.[23] An additional advantage of this procedure is that it allows to

22 Note that this procedure is particularly appealing in our context since it applies to situations in which the instrument is strong.

23 The results of the Local to Zero estimation yields similar conclusions and can be obtained upon request.

[9.8.2018–8:26pm] [1–31] Paper: OP-JEGJ180045
Copyedited by: MANUSCRIPT CATEGORY: ORIGINAL ARTICLE

AQ9

Table 5. Estimated impact of tuition fees with plausibly exogenous instrument

(1) Min deviation	(2) Max deviation	(3) Estimate	(4) Std. deviation	(5) *t*-ratio
Symmetric intervals				
–0.1	0.1	−0.246***	0.066	−3.70
–0.2	0.2	−0.246**	0.113	−2.18
–0.3	0.3	−0.246*	0.140	−1.75
Asymmetric intervals				
–0.3	0	−0.110	0.086	−1.28
–0.2	0	−0.155**	0.066	−2.33
–0.1	0	−0.200***	0.043	−4.60
0	0.1	−0.291***	0.043	−6.71
0	0.2	−0.337***	0.066	−5.04
0	0.3	−0.382***	0.086	−4.39

Estimated equation: Equation (4.1).
Estimation method: UCIs (Conley et al., 2012).
Columns (1) and (2) provide the minimum and maximum values of the parameter.
Capturing the deviation from the exclusion restriction.
Column (3) provides the mean estimate of the fee elasticity.
Column (4) provides the standard deviation of the estimate.

make an assessment of the validity of the instrument when the degree of over-identification is not positive.

Table 5 focuses on the estimation of the elasticity of foreign students to tuition fees for different values in the range of possible values taken by the key parameters capturing the deviation from the exclusion restriction.[24] The higher the range of admissible values, the less precise the estimated coefficient. Symmetric ranges around zero correspond to an agnostic view of the possible deviation of the exclusion restriction of the status of the university as an instrument. A range of positive (negative) values corresponds to the view that foreign students value private Italian universities more (less).

Results of Table 5 suggest that the negative and significant elasticity of tuition fees in the traditional IV estimation is robust to reasonable deviations from the exclusion restriction. The significance level drops below the 5% level only for values of the parameter over 0.2 in absolute terms. This means that, even if the private status of the university deters or attracts (on average) less than 0.2% of foreign students coming from each origin country, our IV estimates support a negative effect of tuition fees. Above that value, our estimates become less significant, albeit still negative at a 10% significance level. We should nevertheless emphasize that this procedure relies on Scaled IV estimation. As Tables 3 and 4 make clear, Scaled OLS and IV estimates are positively biased since they do not account for the issues related to the presence of many zeroes. The Poisson estimates account for this issue and deliver more negative values.

24 The other estimates of Equation (4.1) are not reported here due to space restrictions but are available upon requests. In general, they are unaffected by the alternative procedure compared with the benchmark estimations.

Therefore, the starting estimates used in the current procedure are over-conservative in terms of the negative impact of fees.

The bottom panel of Table 5 also reports results obtained with asymmetric intervals of values of the deviation parameter. By restricting the range of possible deviations, the estimation of the effect becomes slightly more precise. Also, accounting for asymmetry allows us to issue a different point estimate of the impact of tuition fees. The results support the negative impact of tuition fees. Interestingly, our estimations show that if foreign students are more attracted by private Italian universities (which seems the prevailing view), the impact of tuition fees becomes even more negative.

5.2.2. Alternative measure of fees

The benchmark results presented in Tables 3 and 4 rely on a specific measure of tuition fees. This measure was retrieved from a report by the Sole24Ore newspaper, which was based on data transmitted by the MIUR. It might be desirable to use alternative measures of fees to check the robustness of the results.

One alternative source of data relies on the European Tertiary education register (ETER) dataset.[25] The data on fees coming from the ETER dataset exhibit two main differences. First, the fees are computed from balance sheet data provided by the Ministry for each university. Therefore, instead of directly observed fees paid by the students, these fees are based on the part of the budget of the universities which is based on the total amount paid by the students to the university. Second, these amounts include indirect costs of education paid by the students (such as the additional fees related to the registration at the exams and so on). They are therefore not directly comparable to the fees used in the benchmark regressions. Nevertheless, even if they are conceptually different, the fees based on the ETER data are highly correlated with the original ones (correlation amounts to 0.935).

Table 6 presents the results using the fees computed from the ETER database. For the sake of brevity, we just report the results based on the instrumentation of fees (using the same instrument), both for the scaled and the Poisson models. The results show that our main findings are highly robust. As a matter of fact, the elasticity of inflows to the tuition fee level is very in line with the one estimated in the benchmark regressions. It is slightly higher in absolute terms in the scaled IV regressions (around -0.3 instead of -0.25) and slightly lower in the Poisson IV estimations (about -0.75 instead of -0.8 before). These estimates show the robustness of the results with respect to alternative measures of tuition fees.

5.2.3. Alternative measure of quality: QS rankings

Another desirable robustness test concerns the way quality of the university was assessed. Quality of the university is a very robust determinant of the location of the foreign students, not only in our previous regressions but in most papers representative of the literature. It is therefore important to assess the sensitivity of the results to alternative measures of quality. In the previous regressions, we use two alternative measures of quality based on the Top 500 Shanghai ranking. Quite recently, other

25 Data have been provided by the ETER, funded by the European Commission under the contracts EAC-2013-0308, EAC-2015-02080. See https://www.eter-project.com.

Table 6. Estimates of determinants: ETER measures of fees

Variables	(1) Scaled IV	(2) Poisson IV	(3) Scaled IV	(4) Poisson IV
Fees	−0.294***	−0.781***	−0.294***	−0.738***
	(0.02)	(0.14)	(0.02)	(0.14)
Cost of living	−0.083	−1.542***	−0.144**	−1.526***
	(0.06)	(0.46)	(0.06)	(0.43)
Quality (ranking)	0.082***	0.119***	–	–
	(0.00)	(0.02)		
Quality (score)	–	–	0.118***	0.212***
			(0.01)	(0.04)
Host capacity	0.076***	0.405***	0.082***	0.399***
	(0.01)	(0.06)	(0.01)	(0.06)
Income	0.882***	2.374***	0.912***	2.305***
	(0.04)	(0.26)	(0.04)	(0.26)
English	0.079***	0.435***	0.086***	0.470***
	(0.02)	(0.09)	(0.01)	(0.09)
Origin FE	Yes	Yes	Yes	Yes
R^2	0.547	–	0.546	–
F First stage	4150.23	–	3549.47	–
Number of Obs	12,283	12,283	12,283	12,283

Standard errors in parentheses. For scaled IV, robust standard errors.
Instrument: private vs public status of university.
$^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.001$.

institutions have started to issue alternative rankings based on other aspects of quality of the universities. One well-known alternative ranking is the QS World University ranking which ranks universities in each major country (including Italy) based on a set of six criteria (academic reputation, employer reputation, faculty/student ratio, citations, international openness and international student ratio). The QS ranking is increasingly used by foreign students to make decisions with respect to their choice of education at home and abroad. We replicate the benchmark regressions of Tables 3 and 4 using the QS rankings as the basis to compute indicators of quality. These estimates are reported in Table 7.

The results show that the main previous findings are confirmed with an alternative measure of quality. In particular, the results in column (4) relying on Poisson estimation and the instrumentation of fees, i.e. our preferred econometric specification, generate key estimates of the impact of fees and quality that are very in line with the previous results.

5.2.4. Universities with multiple campuses

On important issue that we have so far disregarded is the fact that some universities have different campuses. Since we are looking at location choice of foreign students, it is desirable to see to what extend this might affect our key findings. The multi-campus situation is not obvious to account for, both from a theoretical and empirical point of view. From a theoretical point of view, it is unclear whether students choose specifically a university or a specific campus. Furthermore, it might be the case that universities

Table 7. Estimates of determinants: QS rankings as measures of quality

Variables	(1) Scaled OLS	(2) Poisson	(3) Scaled IV	(4) Poisson IV
Fees	−0.051***	−0.184***	−0.222***	−0.913***
	(0.01)	(0.06)	(0.02)	(0.17)
Cost of living	−0.077	−0.743***	−0.123**	−2.795***
	(0.06)	(0.42)	(0.06)	(0.72)
Quality (ranking)	0.098***	0.092***	0.098***	0.093***
	(0.01)	(0.02)	(0.01)	(0.02)
Host capacity	0.165***	0.606***	0.124***	0.506***
	(0.01)	(0.06)	(0.01)	(0.06)
Income	0.581***	1.607***	0.844***	3.036***
	(0.03)	(0.16)	(0.04)	(0.41)
English	0.012***	0.314***	0.036***	0.349***
	(0.01)	(0.09)	(0.01)	(0.10)
Origin FE	Yes	Yes	Yes	Yes
(Pseudo) R^2	0.569	0.739	0.559	–
F First stage	–	–	4851.41	–
Number of Obs	12,283	12,283	12,283	12,283

Standard errors in parentheses. For scaled IV, robust standard errors.
Instrument in columns. (3) and (4): private vs public status of university.
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.001$.

organize teaching programs that are spread over the various campuses, which make the choice cumbersome for students. From an empirical point of view, data constraints on the specific location of each teaching program might not be very precise and assumptions about a couple of key variables are obviously needed to tackle this issue.

Despite the difficulties in accounting for the issue of multiple campuses, we carry out a robustness check to assess the importance of that issue for the key findings. Our key approach is to carry out an analysis with the unit of observation being the number of foreign students from a given origin in a given campus rather than a specific university.

We proceed to several steps. First, we identify the universities with multiple campuses using the ETER dataset. The ETER data identify 18 universities of that type, with a total of 28 external campuses. Second, like in the benchmark estimations, we get rid of the small universities hosting less than 14 foreign students. Third, we use data from ETER giving the number of students per campuses for the period of investigation. These data allow also to give the number of native students per campus, allowing to retrieve the host capacity per campus. Then we split the data with respect to the various campuses instead of universities. To that aim, we make some assumptions. We assume that different campuses of the same university face the same tuition fee level and exhibit the same reported quality. In contrast, they are subject to different capacity constraints. They are also subject to different expected incomes and costs of living to the extend these campuses are located in different regions. As a matter of fact, multiple campuses are all located in different Italian districts. The expected income and the costs of living are given by the corresponding values of the district and the city in which the campus is located. If this is unavailable at the city level, we take the available value of the closest city to that campus.

Table 8. Estimates of determinants: accounting for multiple campuses

Variables	(1) Scaled IV	(2) Poisson IV	(3) Scaled IV	(4) Poisson IV
Fees	−0.179***	−0.771***	−0.184***	−0.712***
	(0.02)	(0.15)	(0.02)	(0.14)
Cost of living	−0.022	−1.817***	−0.083*	−1.662***
	(0.05)	(0.62)	(0.05)	(0.56)
Quality (ranking)	0.047***	0.112***	–	–
	(0.00)	(0.02)		
Quality (score)	–	–	0.073***	0.200***
	(0.01)		(0.00)	(0.03)
Host capacity	0.164***	0.496***	0.163***	0.498***
	(0.00)	(0.04)	(0.00)	(0.04)
Income	0.818***	2.531***	0.842***	2.393***
	(0.03)	(0.33)	(0.03)	(0.31)
English	0.043***	0.365***	0.052***	0.397***
	(0.01)	(0.09)	(0.01)	(0.08)
Origin FE	Yes	Yes	Yes	Yes
R^2	0.531	–	0.530	–
F First stage	5869.72	–	5829.43	–
Number of Obs	16,089	16,089	16,089	16,089

Standard errors in parentheses. For scaled IV, robust standard errors.
Instrument: private vs public status of university.
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.001$.

The resulting data of this procedure span 106 initial campuses, from which we retrieved 16 campuses as they host less than 14 foreign students as a whole. We end up with 93 locations instead of 71 universities considered in the benchmark analysis, yielding a total number of observations over 16,000 points. Table 8 gives the results of the estimations corresponding to Table 4, i.e. those with the fees instrumented by the status of the university. The results with respect to the impact of fees on the foreign student inflows are very similar to those in the benchmark analysis. In particular, the estimated elasticities of the Poisson IV regressions are very close to the previous ones. While the estimates of the impact of cost of living are slightly different in the scaled IV estimations, the ones estimated by the Poisson IV procedure are also very in line with those of the benchmark regressions. This is also the case for all the other determinants. All in all, these results show that our main findings are still valid when accounting for the existence of multiple campuses in Italian Universities.

5.2.5. Scaled regressions

Another concern related to model Equation (4.1) is that the model does not perfectly match the idea of the multinomial logit defined in the theoretical model (see Section 2). In particular, in a multinomial logit set-up, one increase in the attractiveness of a given university proportionally decreases the attractiveness of the other ones. If, e.g. the ranking of La sapienza tends to increase, this should lead both to a larger inflow of foreign students to La Sapienza and to, say, a decrease in the foreign students intake in

Tor Vergata (another university in Roma of about the same quality). The same also holds for the other covariates, including tuition fees.

To deal with this, we change the estimated specification Equation (4.1) by scaling all variables by a reference level. The reference level is chosen at the dyadic level, i.e. it varies across each pair and is specific to each origin country. We scale all variables in the specification by the level prevailing at the university at the destination that hosts the greatest number of students from origin country o. In practice, for each origin country o, we determine the university that hosted the largest number of international students during the academic year 2011–2012. This variable is labeled by $(\tilde{M}_{o,d,u^d})^*$.[26]

The extended model that we consider takes the following form:

$$\ln\left(\frac{\tilde{M}_{o,d,u^d}}{\tilde{M}^*_{o,d,u^d}}\right) = \alpha + \alpha_d + \beta_1 * \ln\left(\frac{fees_{u_d}}{fees_{(u_d)^*}}\right) + \beta_2 * \ln\left(\frac{living\,cost_{u_d}}{living\,cost_{(u_d)^*}}\right) + \beta_3 * \ln\left(\frac{quality_{u_d}}{quality_{(u_d)^*}}\right)$$
$$+\beta_4 * \ln\left(\frac{hostcapacity_{u_d}}{hostcapacity_{(u_d)^*}}\right) + \beta_5 * \ln\left(\frac{expreturn_{u_d}}{expreturn_{(u_d)^*}}\right) + \beta_5 * Eng_{d,u_d} + \epsilon_{(u_d)} \quad (28)$$

Table 9 presents the results. It is directly comparable in terms of structure with the ones reporting the benchmark regressions, i.e. Tables 3 and 4. Nevertheless, given the new definitions of the variables (both the dependent and the covariates), the estimates of the coefficients are not comparable. The new regressions are better used to check the robustness of the results in terms of statistical significance.

Columns (1) and (2) of Table 9 report the non-IV estimates of model Equation (5.1). Columns (3) and (4) contain the results obtained applying the IV strategy. In all estimations, we use the Score as indicator of quality.[27] Table 9 provides additional evidence of the negative impact of fees on international student inflows. Also, the estimates of the other covariates almost perfectly mirror the results obtained when considering the baseline model Equation (4.1).

6. Conclusions

This article revisits the issue of the determinants of student migration. In contrast to the existing literature that has focused up to now on country-specific factors, we look at the determinants at the university level. This allows us to address specifically the role of important factors such as tuition fees or the quality of the university. The impact of those factors is difficult to grasp in country-level studies due to the high heterogeneity among institutions in many countries. While the analysis considers a set of university-specific factors, we pay special attention to the role of tuition fees in the inclination of foreign students to choose a specific university. So far, the existing literature has obtained mixed results concerning the impact of tuition fees.

We build our empirical investigation on a nested logit model capturing the decision to choose a specific university abroad. We focus on the last decision nest, i.e. the choice of

26 When the largest flow from a given country of origin is shared among several universities, we scale each covariate by the average values among these universities.

27 Estimations with the '*Ranking*' as a proxy of university quality are available upon request and give similar results.

Table 9. Estimations of determinants: scaled regressions

Variables	(1) Scaled OLS	(2) Poisson	(3) Scaled IV	(4) Poisson IV
Fees	−0.009***	−0.144***	−0.021***	−0.376**
	(0.00)	(0.04)	(0.00)	(0.08)
Cost of living	−0.046***	−0.079	−0.061***	−0.557**
	(0.01)	(0.20)	(0.01)	(0.26)
Quality (score)	0.020***	0.136***	0.020***	0.135***
	(0.00)	(0.02)	(0.00)	(0.02)
Host capacity	0.017***	0.446***	0.014***	0.395***
	(0.00)	(0.03)	(0.00)	(0.04)
Income	0.102***	1.499***	0.120***	1.903***
	(0.00)	(0.09)	(0.01)	(0.16)
English	0.010***	0.299***	0.012***	0.328***
	(0.00)	(0.05)	(0.00)	(0.05)
Origin FE	Yes	Yes	Yes	Yes
(Pseudo) R^2	0.215	0.159	0.212	–
F First stage	–	–	4677.66	–
Number of Obs	12,283	12,283	12,283	12,283

Standard errors in parentheses. For scaled IV, robust standard errors.
In columns (3) and (4), Instrument: private vs public status of university.
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.001$.

a specific university for a student, conditional on going abroad and conditional on choosing a specific destination country. This choice is constrained by binding capacity constraints on the side of hosting universities. Our model allows the identification of the main factors such as tuition fees, quality of the university, host capacity, expected return on education at destination, cost of living and the existence of programs taught in English. We estimate the role of those factors, using data at the university level for foreign students in Italy. One of the important issues at the econometric level is the endogeneity of fees. We use a classical IV approach based on the status (private vs public) of the universities.

Our analysis generates interesting and new findings. First, we find evidence of the negative role of a university's tuition fees on the flow of students choosing to study in that university. The typical estimate implies that an increase in tuition fees of 10% would reduce the bilateral flow by about 8%, suggesting a non-negligible effect in terms of magnitude. Surprisingly, this negative and significant role is new in the literature. We stress the importance of dealing with the endogeneity of tuition fees. Failure to account for endogeneity results in significantly different results and conclusions. While for instance a positive impact is not to be ruled out at a theoretical level, it is nevertheless difficult to rationalize in practice. The negative impact of fees that we document to be robust to a set of robustness checks: deviations from the exclusion restriction in the IV procedure, use of alternative measures of tuition fees and quality, the issue of multiple campuses in some universities and a change of the specification of the underlying model more consistent with the multinomial logit model. While tuition fees are found to have some influence on the location of foreign students, our analysis also emphasizes and confirms the role of other important factors. We find support in favor of the role of the university's quality. Also, the expected return to education after graduation is found to be important. This last result is in line with the implications of the migration model of foreign education.

The evidence of a negative elasticity between foreign student inflows and tuition fees suggests that universities need to pay attention to their pricing policies. While foreign students clearly pay attention to the quality of education provided by the universities, they also take the cost of education into account. While some part of this cost, namely the cost of living, is not directly controlled by universities, the level of the fees charged to foreign students should receive a close scrutiny from the university authorities in order to maintain some external attractiveness. In other terms, significant increases in tuition fees like those that have been observed in some countries over the recent period need to be matched by significant increases in the quality of education that is provided by these institutions. Failure to disregard the quality-price trade-off in foreign education could result in a loss of attractiveness of the university and in a decrease in the number of incoming students from abroad.

While this article provides a complete level of analysis of the role of fees on foreign education, it has of course some limits and calls for future research on that topic. While the number of foreign students in Italy is on the rise, Italy is not the major destination over the world. A similar analysis applied to one of the major destinations such as the US, Australia or the UK would be desirable. Another aspect is that fees in Italy do not vary much by field of study. An extension of this analysis including foreign enrollment by field could be useful. Finally, it would be desirable to study contexts in which tuition fees have been subject to significant variation over time. Investigation in such a context would allow to overcome the limitation of cross-sectional analyses.

Supplementary material

Supplementary data for this paper are available at *Journal of Economic Geography* online.

Acknowledgements

An earlier version of this article has been presented in several workshops and academic conferences, including those taking place in Geneva, Neuchatel, Luxembourg (University of Luxembourg), Perth (UWA), Sydney (Macquarie), Paris (OECD). We would like to thank among many others A. Ariu, S. Becker, M. Cervelatti, S. Chang, S. Coulombe, J. de Melo, F. Docquier, A. Dupuy, E. Carroni, J. Fenske, C. Heaton, H. Jayet, M. Jetter, M. Joxhe, B. Lanz, J. Machado, A. Moro, M. Muller, G. Orefice, C.Parsons, M. Pecoraro, P. Picard, G. Tripathi and S. Zanaj as well as the editor and two anonymous referees for helpful comments and suggestions. We are indebted to P. Buchanan for excellent editing work on a previous version of this article. All errors remain our own.

References

AQ4 AQ5

Abbott, A., Silles, M. (2015) Determinants of international student migration. *World Economy*, 39: 621–635

~~Agasisti, T., Dal Bianco, A. (2007) Determinants of college student migration in Italy: empirical evidence from a gravity approach. Technical Report, Paper Presented at the Congress of the European Regional Science Association. Paris, September. Available at SSRN 1063481.~~

Alecke, B., Burgard, C., Mitze, T. (2013) The effect of tuition fees on student enrollment and location choice interregional migration, border effects and gender differences. Technical Report, Ruhr Economic Papers 404.

Becker, G. S. (1964) *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Chicago Press.

Beine, M., Bertoli, S., Fernández-Huertas Moraga, J. (2015) A practioners guide to gravity models of international migration. *World Economy*, 39: 496–512

Beine, M., Delogu, M., Ragot, L. (2016) The Role of Fees in Foreign Education: Evidence from Italy and the UK. CREA Discussion Paper, 2016-06.

Beine, M., Delogu, M., Ragot, L. (2017) Understanding the Impact of Tuition Fees for Foreign Students: The Case of the UK. CREA Discussion Paper, 2017-07.

Beine, M., Noël, R., Ragot, L. (2014) Determinants of the international mobility of students. *Economics of Education Review*, 41: 40–54

Bertoli, S., Fernández-Huertas Moraga, J. (2013) Multilateral resistance to migration. *Journal of Development Economics*, 102(C): 79–100

Bessey, D. (2012) International student migration to Germany. *Empirical Economics*, 42: 345–361

Bruckmeier, K., Fischer, G.-B., Wigger, B. U. (2013) Does Distance Matter? Tuition Fees and Enrollment of First-Year Students at German Public Universities. Technical Report, CES IFO Working Paper 4258.

Bruckmeier, K. G.-B. F., Wigger, B. U. (2015) Tuition fees and the time to graduation: evidence from a natural experiment. *Journal of Higher Education Policy and Management*, 37: 459–471

Card, D., Krueger, A. B. (1992) Does school quality matter? Returns to education and the characteristics of public schools in the United States. *The Journal of Political Economy*, 100: 1–40

Carneiro, P., Heckman, J. J. (2002) The evidence on credit constraints in post-secondary schooling. *Economic Journal*, 112: 705–734

Chapman, B., Ryan, C. (2005) The access implications of income-contingent charges for higher education: lessons from Australia. *Economics of Education Review*, 24: 491–512

Christofides, L., Hoy, N., Yang, L. (2010) Participation in canadian universities: the gender imbalances. *Economics of Education Review*, 29: 400–410

Conley, T. G., Hansen, C. B., Rossi, P. E. (2012) Plausibly exogenous. *The Review of Economics and Statistics*, 94: 260–272

De Palma, A., Picard, N., Waddell, P. (2007) Discrete choice models with capacity constraints: an empirical analysis of the housing market of the greater Paris region. *Journal of Urban Economics*, 62: 204–230

Dearden, L., Fitzsimons, E., Wyness, G. (2011) The Impact of Tuition Fees and Support on University Participation in the UK. Technical Report, IFS Workimg Paper W11/17.

Dreher, A., Poutvaara, P. (2008) Student flows and migration: an empirical analysis. *World Development*, 39: 1294–1307

European Commission. (2012) *National Student Fee and Support Systems 2011/2012*. European Commission.

Flannery, D., O'Donoghue, C. (2009) The determinants of higher education participation in Ireland: a micro analysis. *The Economic and Social Review*, 40: 73–107

González, C. R., Mesanza, R. B., Mariel, P. (2011) The determinants of international student mobility flows: an empirical study on the Erasmus programme. *Higher Education*, 62: 413–430

Jena, F., Reilly, B. (2013) The determinants of United Kingdom student visa demand from developing countries. *IZA Journal of Labor & Development*, 2: 1–22

Kahanec, M., Králiková, R. (2011) Pulls of International Student Mobility. IZA Discussion Papers 6233, Institute for the Study of Labor (IZA).

Leslie, L., Brinkman, P. (1987) Student price response in higher education: the student demand studies. *Journal of Higher Education*, 58: 181–204

McFadden, D. L. (1984) Econometric analysis of qualitative response models. *Handbook of Econometrics*, 2: 1395–1457

OECD. (2015) *Education at a Glance 2015: OECD Indicators*. Paris: OECD Publishing.

Perkins, R., Neumayer, E. (2014) Geographies of educational mobilities: exploring the uneven flows of international students. *The Geographical Journal*, 180: 246–259

Rosenzweig, M. (2008) Higher education and international migration in Asia: brain circulation. In *Annual World Bank Conference on Development Economics*, pp. 59–100.

AQ8 Rosenzweig, M. R., Irwin, D. A., Williamson, J. G. (2006) Global wage differences and international student flows [with comments and discussion]. In *Brookings Trade Forum*, pp. 57–96. ~~JSTOR~~.

AQ8 Silva, J. S., Tenreyro, S. (2006) The log of gravity. *The Review of Economics and statistics*, 88: 641–658

Soo, K., Elliott, C. (2010) Does price matter? Overseas students in UK higher education. *Economics of Education Review*, 29: 553–565

AQ7 Train, L. (2003) *Discrete Choice Methods with Simulation*. Cambridge University Press.

Van Bouwel, L., Veugelers, R. (2013) The determinants of student mobility in Europe: the quality dimension. *European Journal of Higher Education*, 3: 172–190

Wales, P. (2013) Access all areas? The impact of fees and background on student demand for postgraduate higher education in the UK. Technical Report, SERC Discussion Paper 128.

AQ6