Choice of Academic Degree

Alaitz Artabe and Javier Gardeazabal*
University of The Basque Country (UPV/EHU)

December 5, 2012

Abstract

This paper investigates the extent to which individuals' choice of college degree depends on personal characteristics and degree attributes. We propose a method to estimate a discrete choice model that avoids the identification problem that is typically found when there is a large number of alternatives and utility depends on both individual characteristics and alternative attributes. Using pre-enrollment data from the University of the Basque Country, we estimate a discrete choice model. We find that the choice of college degree is influenced by individual characteristics, such as, gender, grades in Basque language and Mathematics, province of residence and distance to place of study, as well as, alternative attributes, such as, expected wage and the time taken to find a job.

*We would like to thank Amaia Zarraga for providing us with the pre-enrollment data and her help in making sense of them. Alaitz Artabe acknowledges financial support from the Basque Government through "Programas de Ayudas para la Formación y Perfeccionamiento del Personal Investigador", from the Spanish Ministry of Science and Innovation (ECO2009-11213) and from Fundación Banco Herrero. Javier Gardeazabal would like to thank the Department of Economics at the University of California Santa Cruz for its hospitality during part of this research and acknowledges financial support from the Spanish Ministry of Science and Innovation (ECO2009-09120), the Basque Government (IT-313-07) and the Spanish Ministry of Education (PR2011-0481).
Mailing address: Departamento de Fundamentos del Análisis Económico II, Universidad del País Vasco, Lehendakari Aguirre 83, 48015 Bilbao, Spain. E-mail: alaitz.artabe@ehu.es javier.gardeazabal@ehu.es

1 Introduction

After completing secondary education, students decide whether to continue studying for a higher degree or to look for a job. Those who decide to attend college have to decide which degree to pursue. These decisions are sequential, with the first being the decision as to whether or not continue studying and the second the choice of degree. This paper focuses on the choice of degree, that is conditional on having decided to continue studying. In particular, this paper establishes a degree choice problem and solves it from the point of view of a rational individual who selects the one that reports the maximum utility level from among the set of feasible degree choices.

In order to establish the degree choice problem, it is necessary to determine what the factors are affecting individual preferences regarding degrees. Some of these factors are the specific characteristics of individuals such as gender, age or knowledge. For instance, there is ample evidence that some degrees have a male or female majority. Although the pattern is changing, most engineers are male and most teachers are female. Thus, an individual characteristic, gender, is a factor determining their choice. Similarly, a more mathematically inclined person might choose a science degree as opposed to humanities or social sciences.

Degree attributes are potential factors determining degree choices. Some degree attributes might attract the interest of potential students. For instance, potential high earnings as a lawyer might induce more students to study law. On the contrary, other degree attributes might discourage students. For example, high unemployment rates among law graduates might induce some students to pursue other degrees.

Student characteristics are degree-invariant and degree attributes are student-invariant. However, some factors affecting degree choice vary with degrees and individuals. For instance, tuition might vary with degree and student if colleges price discriminate among students and degrees. Similarly, another potential factor affecting degree choice, the distance from the student's hometown to the place of study, varies with degrees and students.

Although it may be less obvious, combinations of personal characteristics and degree-specific attributes might also affect the decision to choose a particular degree. For instance, some women might be willing to pursue a degree known to lead to jobs with flexible working hours.

This paper analyzes how individual characteristics, degree attributes and combinations of both, might affect degree choices. For this purpose, we use a multinomial model of degree choice. According to this model, the utility individuals derive from a degree depends on their characteristics, degree attributes and variables that depend on both individuals and alternatives. Furthermore, the multinomial choice model assumes individuals choose the alternative that reports the highest possible utility among the set of feasible choices. Without further restrictions, this model would suffer from an identification problem. That is why the majority of the models proposed in the literature consider either individual characteristics or attributes, but not both simultaneously. In addition, our specific application of a multinomial choice model to a degree choice setting implies that the choice set includes a large number of alternatives, which complicates the analysis even further.

In order to solve the identification problem, we propose a two-step estimation procedure that allows utility to depend on individual characteristics and attributes and a large number of alternatives. Our procedure can also be used in other fields where individual choices depend on their characteristics and alternative-specific attributes. We illustrate our methodology using pre-enrollment data from the University of the Basque Country (UPV/EHU), a large public university that offers over a hundred degrees from different fields.

The rest of the paper is organized as follows. Section 2 reviews related literature on the demand for education and the methodologies used. Section 3 describes the degree choice model and the methodology used to estimate it. Section 4 describes the data. Section 5 illustrates the methodology using pre-enrollment data from the University of the Basque Country. Section 6 summarizes the results obtained.

2 Literature Review

The factors determining students' participation in higher education are extensively analyzed in the literature, e.g. Rice (1987), Albert (2000), Marcerano-Gutierrez (2007) and Cepar and Bjnec (2010). As the decision to go to college is inherently dicothomous, authors use binomial logit models to analyze how different factors affect individuals' choices, e.g. Gonzalez López-Valcárcel and Dávila Quintana (1998) and Rahona López (2006).

Previous studies on the degree choice problem invariably use discrete choice models, as the dependent variable is polycothomous. In general, the number of alternatives or number of feasible degree choices is very large, representing a problem for applied analysis. However, degrees can be aggregated into broad categories to reduce such a large number. For instance, SalasVelasco and Martín-Cobos Puebla (2006) focus on the choice between long degrees (4-year degrees) and short degrees (3-year degrees) using a binomial logit model. Likewise aggregating degrees into different fields, Ayalon (2003) and SalasVelasco and Martín-Cobos Puebla (2006) analyze the field-of-degree choice using a multinomial logit model.

Due to the aforementioned identification problem, most models used in the literature either consider individual characteristics or alternative attributes as factors that influence individuals' choices, but not both. In particular, authors use the multinomial logit to account for individual characteristics or the conditional logit model to capture the effect of alternative attributes.

This identification problem has been dealt with in other fields. Two procedures have been suggested to overcome the identification problem. The first one is a modification of the conditional logit to allow for individual characteristics and alternative attributes. Greene (2003), Kim and Kwon (2003), Mazzanti (2003) and Iyengar and Haln (2009) create dummy variables for the alternatives and then interact them with individual characteristics. This procedure, however, is not feasible when the number of alternatives is large, as the number of regressors equals the number of alternatives times the number of individual characteristics.

The second procedure, e.g. Murdock (2006) and Timmins and Murdock (2007), requires a large number of alternatives and involves two steps. In the first step, these authors estimate a multinomial model that includes individual characteristics and alternative-specific constants that can be considered as alternative-specific unobserved effects. In the second step, the alternative-specific constants estimated in the first step are regressed on alternative attributes. In other words, alternative attributes enter the utility function through the alternative-specific unobserved effects.

Within the demand for education literature, a similar two-step procedure is used by Bratti (2002) to investigate students performance across UK universities. The only difference with the two-step procedure outlined above is that Bratti uses an ordered probit in the first step as performance is measured ordinally. Bratti introduces college-dummy variables in the first step and then regresses the coefficients on college dummies on college attributes in order to assess whether differences in students' performance can be explained by colleges attributes.

Our contribution to the literature is twofold. First, we propose a third estimation procedure that can be thought of as a combination of the two previously used procedures. As in the first procedure outlined above, we interact individual characteristics with alternative-specific dummies, but also with all attributes. As in the second procedure outlined above, we estimate the model in a two-step manner. In the first step, we estimate an alternative-specific conditional logit model where we include all individual characteristics and alternative specific constants, as is done in the literature. In the second step, we regress alternative specific constants and, this is the novelty, slopes on alternative attributes. Second, we provide an illustration of our procedure using real data from pre-enrollment records at the University of the Basque Country and are able to determine the role of personal characteristics and degree attributes in determining students' degree choices.

3 The Model

Assume that the utility individual i obtains from pursuing degree s is given by

$$U_{is} = U(Z_i, X_{is}, W_s, \varepsilon_{is}) \tag{1}$$

where $i = 1, ..., N$; $s = 1, ..., S$; $Z_i = (Z_{1i}, ..., Z_{Ki})$ are individual i's characteristics; $W_s = (W_{1s}, ..., W_{Ls})$ are degree s's attributes; $X_{is} = (X_{1is}, ..., X_{Jis})$ are covariates that vary with both individual and degree and ε_{is} is a zero-mean random disturbance. Assume individual i chooses alternative s if and only if $U_{is} > U_{ir}$ for all $r \neq s$. In our degree-choice setup, Z_i includes gender, age, and other personal characteristics, W_s, includes expected post-graduation earnings and other degree attributes, and X_{is} includes variables that depend on the degree and the individual such as the distance from the student's home to college.

Further assume that the utility function can be approximated by a linear function as follows

$$U(Z_i, X_{is}, W_s, \varepsilon_{is}) = \sum_{j=1}^{J} X_{jis}\beta_j + \sum_{k=0}^{K} Z_{ki}\delta_{ks} + \varepsilon_{is} \tag{2}$$

for $i = 1, ..., N$ and $s = 1, ..., S$ and

$$\delta_{ks} = \sum_{l=0}^{L} W_{ls}\beta_{kl} + u_{ks} \tag{3}$$

for $k = 0, ..., K$, $Z_{0i} = 1$, $W_{0s} = 1$, $E(\varepsilon_{is} \mid Z_i, X_{is}, W_s) = 0$, $E(u_{ks} \mid W_s) = 0$ and $E(u_{ks}^2 \mid W_s) = \sigma_u^2$.

Note that Equation 2 is the latent variable model underlying an alternative-specific conditional logit. The extension here corresponds to the assumption summarized in Equation 3 whereby the alternative specific coefficients δ_{ks} depend on the attributes linearly. The model given by 2 and 3 can be written as

$$U_{is}(Z_i, X_{is}, W_s, \varepsilon_{is}) = \sum_{j=1}^{J} X_{jis}\beta_j + \sum_{k=0}^{K}\sum_{l=0}^{L} Z_{ki}W_{ls}\beta_{kl} + \sum_{k=0}^{K} Z_{ki}u_{ks} + \varepsilon_{is}$$

where the second term in the right hand side contains all interactions between individual characteristics and attributes, including the alternative specific constants, and therefore allows for a fairly general specification.

We propose the following estimation procedure.

Step 1. Estimate the parameters in Equation 2 using McFadden's alternative specific conditional logit (McFadden 1974). Denote these estimates by $\widehat{\beta}_j$ and $\widehat{\delta}_{ks}$.

Step 2. Estimate the parameters of the system of Equation 3 by least squares using the estimated values $\widehat{\delta}_{ks}$ as dependent variables.

Two comments are in order. First, the second step involves estimation of 3 where the dependent variable is replaced by an estimate, that is

$$\widehat{\delta}_{ks} = \sum_{l=0}^{L} W_{ls}\beta_{kl} + u_{ks} - \widehat{v}_{ks} \tag{4}$$

where $\widehat{v}_{ks}$ is the estimation error from the first step,

$$\widehat{v}_{ks} = \delta_{ks} - \widehat{\delta}_{ks}.$$

Therefore, the second step can be viewed as an error of measurement in the dependent variable problem, which does not generate bias in the regression estimates.

Second, Equation 3 represents a system of K equations, one equation per personal characteristic used in the first step plus the constant term. Therefore, this system of equations is a Seemingly Unrelated Regression Equations (SURE).

3.1 The marginal effects on utility

For continuous variables, the marginal effects of individual characteristics on expected utility are

$$\frac{\partial E\left(U_{is}|\, X_{is}, Z_i, W_s\right)}{\partial Z_{ki}} = \delta_{ks} \tag{5}$$

and the marginal effect of attributes on utility are

$$\frac{\partial E\left(U_{is}|\, X_{is}, Z_i, W_s\right)}{\partial W_{ls}} = \sum_{k=0}^{K} Z_{ki}\beta_{kl} \tag{6}$$

Notice that the marginal effects of attributes depend on the values of individual characteristics. Averaging over individuals gives us the marginal effect evaluated at the average value of the individual characteristics

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\partial E\left(U_{is}\middle| X_{is},Z_i,W_s\right)}{\partial W_{ls}} \quad = \quad \sum_{k=0}^{K}\overline{Z}_k\beta_{kl} \tag{7}$$

Statistical significance of the marginal effects of individual characteristics can be assessed simply by looking at the significance of the δ_{ks} parameters in the first step. Statistical significance of the marginal effects of attributes is more difficult. A way of doing so is to test the cross-equation restriction

$$\sum_{k=0}^{K}\overline{Z}_{ki}\beta_{kl}=0.$$

Note that testing this cross-equation restriction requires system estimation and can not be perform if the system 3 is estimated using OLS equation by equation.

4 The data set

We assembled data from three different sources. Student preferred degree and personal characteristics were obtained from 2009 pre-enrollment records from the University of the Basque Country. Data on degree attributes were gathered from the 2009 *Encuesta de Inserción Laboral Universitaria,* a survey conducted by the Basque Government employment office, *Lanbide,* among 2006 graduates, and from the 2009 statistical compendium of the University of the Basque Country *UPV/EHU en cifras.*

As will be seen, this data set has special features that make it particularly appropriate for our purpose. The University of the Basque Country is the only public university in the Spanish autonomous region of the Basque Country or *Euskadi.* It is a large university with over 40,000 students and several campuses spread over three provinces, *Araba*, *Bizkaia* and *Gipuzkoa*. A very large fraction of Basque students who enter college attend the University of the Basque Country.

Most of the pre-enrollment data corresponds to a cohort of students that finished high school in 2009 and pre-enrolled at the University of the Basque Country. As part of the pre-enrollment, students select up to five degrees that they are interested in and rank them according to their preferences. As students do not have incentives to behave strategically, their rankings should reflect their preferences, so the degree ranked in the first place should be the degree that reports them the highest utility. Note that we use data from the pre-enrollment and not the final enrollment because the final choice can be conditioned by *numerus clausus* applied to some degrees, so the final choice might not be their preferred degree. Students choose from among 87 undergraduate degrees.[1] Some of the pre-enrollment records are for students who did not finally register at the University of the Basque Country, either because they decided to interrupt their studies or opted to persue a degree at a university other than the University of the Basque Country. However, if those students pre-enrolled at the University of the Basque Country, the alternative ranked first should be their preferred alternative from among those offered at the University of the Basque Country. Summarizing, we believe this data set is particularly valid to provide information about students preferences regarding degrees.

[1] Some degrees are taught at different campuses. For our purposes, the same degree offered at different campuses is considered to be different degrees.

We obtained pre-enrollment records for 5,029 students and restricted the analysis to all degrees that were ranked in first place by at least ten students, thus reducing the analysis to 77 degrees. Apart from their preferred degree, several personal characteristics are reported in students' pre-enrollment records including, age, gender, place of residence, the grade obtained in the university admission exam, *selectividad*, as well as whether the admission exam was taken in Spanish or Basque language and students' grades in the Mathematics, Spanish, Basque, English, Philosophy and History parts of the university admission exam as well as a global grade which is a weighted average of the university admission exam (60%) and the average grade from high school (40%). A detailed definition of each variable is given in an Appendix.

Table 1 reports descriptive statistics for individual characteristics. Age ranges from 17 to 42 and although the mean is much closer to the minimum of the support, the existence of some older individuals allows us to investigate its effect on the choice of degree. The average gender is below 0.5 indicating that most enrolling students were females. Table 1 also reports descriptive statistics on the average grades in the university admission exam, *selectividad,* the global grade and grades in different subjects.

Our pre-enrollment data set is consistent with previous findings. First, the average grade of students who prefer long degrees (7.0554) is higher than the average grade of those who choose shorter degrees (6.4582), as documented by Jimenez and Salas-Velasco (2000) and Salas-Velasco and Martín-Cobos Puebla (2006) for other data sets. Second, Table 2 indicates that males tend to choose technical degrees more often than females as documented earlier by González López-Valcárcel and Dávila Quintana (1998) and Ayalon (2003). Third, grades in different subjects might be crucial in determining the decision to choose a degree as documented by Strenta, Elliott, Adair, Matier and Scott (1994). Table 3 reports average grades in different subjects grouped by chosen field-of-degree. Those who choose technical degrees, experimental sciences and health sciences degrees have higher average grades than those who choose a degree in Humanities or Social Sciences.

The distance from home to the school where the degree is taught is an important factor for the students to decide their preferred degree as previously suggested by González López-Valcárcel and Dávila Quintana (1998), Montgomery (2002), Long (2004) and Drewes and Michael (2006). The cost of attending college increases with distance as students have either to commute incurring monetary costs and time spent in commuting or move to another town incurring renting expenses. From the information on the place of residence we computed for each student the distance from the municipality of residence and the municipality where each school is located.[2] Distance to college changes with individuals and alternatives. On the other hand, Table 4 classifies students according to the province where college is located and the province of residence. Most students who live in *Bizkaia* want to study in *Bizkaia* and the same is true for *Gipuzkoa* but not for *Araba*. Therefore, in addition to distance from hometown to college, we include a dummy variable that indicates if the student's hometown is in the same province where the degree is taught. Should this dummy be a significant determinant of degree choice, the interpretation would be that, in addition to the negative effect of distance on utility, there is a provincial border effect similar to the one documented in international trade studies e.g. McMallum (1995).[3]

[2] Distances were computed using municipal geographic coordinates and then converted to kilometers using the spherical law of cosines.

[3] McMallum showed that Canadian inter-provincial trade was 22 times greater than trade between Canadian provinces

Degree-specific attributes also influence individual preferences. A particularly important degree attribute is the minimum grade required to enter a degree. We consulted the administrative records at the University of the Basque Country to obtain the ex-post minimum entering grade in each degree, that is, the value of the aforementioned global grade corresponding to the individual who entered a degree with the lowest grade. We will refer to this grade as the *Threshold Grade*. In highly demanded degrees, where the *numerus clausus* is binding, this threshold grade is high. Whereas in those degrees where the *numerus clausus* is not binding, the threshold grade corresponds to the minimum grade required to be accepted to the University.

In addition to the data obtained from pre-enrollment records, we also gathered data from *Encuesta de Inserción Laboral Universitaria,* an independent survey carried out by the Basque Government employment office *Lanbide.* This survey is conducted on a population of college graduates three years after completion of their degrees. Respondents are asked questions about their employment situation, earnings and more. These items are used to construct degree attributes that measure labor market conditions for graduates. We use two such indicators, the average time needed to find a job after completing their degree and the average wage earned by graduates with the same degree. Other studies, e.g. Jimenez and Salas-Velasco (2000), Long (2004) and Maringe (2006), Salas-Velasco and Martín-Cobos Puebla (2006), suggest that this choice of labor market indicators is a reasonable one.

Four more attributes were collected from the statistical compendium *UPV/EHU en cifras.* We first constructed a set of provincial dummies indicating the province where each degree is taught. A second attribute compiled from the statistical compendium is the *Student/Teacher Ratio*, which measures the ratio of the number of students to the number of teachers in each school. A low student-teacher ratio signals more personalized teaching and, *ceteris paribus*, better quality of instruction.

A third attribute from the statistical compendium has to do with the language of instruction. The Basque Country is a bilingual community. Basque and Spanish are the two official languages in the Basque Country. Therefore language of instruction can be an important factor influencing the decision to study a particular degree. Almost all subjects of all the degrees are taught in Spanish and a large fraction of them in Basque. We constructed the variable *Basque Ratio,* which measures the percentage of subjects that are offered in Basque.

The fourth attribute computed from the statistical compendium is the number of credits required in each degree. The number of credits can be considered a proxy of the difficulty of the degree. Descriptive statistics of attributes are reported in Table 1.

5 Empirical results

The random utility model is estimated using a two-step procedure. In the first step, we estimate Equation 2 using McFadden's alternative-specific conditional logit. As there are 77 degrees and 10 individual characteristics plus a constant term per degree, this results in 847 parameter estimates in the first step. Reporting such a large number of parameters is difficult. We opted for presenting those parameter estimates in table 5. For instance, a +/0/- in table 5 indicates that the variable in the corresponding column has a positive/insignificant/negative effect on utility. In all cases the Degree in

and US states of similar size and proximity and attributed the observed difference to the so called border effect.

Economics is taken as the reference.

As regards the demographic effects, the results are as follows. As shown in column (1) of table 5, age is a significant factor for eight degrees. Interestingly enough, although age is only a significant factor for a few degrees, its effect is negative, meaning that younger students tend to select those degrees. Column (2) shows that *gender* significantly affects the degree choice. Women are more likely to choose degrees related with education (*Child Education* and *Social Education*), health (*Nursing, Pharmacy, Dentistry* and *Psychology*) and languages (*Translation and Interpretation*) while men are more likely to choose *Physical Education*, *Physical Activity and Science* and *Engineering*.

Column (3) shows that using Basque as the vehicular language of instruction is not a significant factor for most degrees. However, Basque as a vehicular language increases the probability of choosing *Child Education*.

We turn now to the analysis of the effects of the grades in the admission exam and parts of it. The grades obtained in *Basque* are significant as shown in column (4). Obtaining high marks in *Basque* increases the probability of choosing degrees such as *Basque Philology*, *Child Education* and *Primary Education*. Columns (5) to (8) indicate that the marks obtained in *Philosophy*, *History*, *English* and *Spanish* are not significant for most degrees. However, *Mathematics* grades are a significant factor for most degrees as shown in column (9). The higher the marks obtained in *Mathematics*, the higher the probability of choosing degrees such as *Mathematics*, *Chemistry*, *Biology* and *Industrial Engineering*. Columns (10) and (11) show that the *global grade* and the mark obtained in the admission exam are not significant for most degrees.

In addition to personal characteristics, the set of explanatory variables includes two variables that vary with students and degrees: *Province* and *Distance*. The coefficients associated with these variables are alternative invariant, accordingly the results for these variables are reported in Table 5. *Distance* has a negative and significant impact on utility and living in the same *province* where the degree is taught has a positive and significant effect on utility. Hence, there seems to be a border-effect in the election of the degree, so students tend to choose degrees in their province of residence.

The second step of our procedure focuses on estimating the system of equations 3, which accounts for the influence of attributes on utility. Each column in Table 6 reports SURE estimates of one of the equations of system 3. For instance, column (1) reports the regression of the constant terms $\widehat{\delta}_{0s}$ (the estimates of the constant terms from the alternative-specific conditional logit) on all attributes. Similarly, column (2) reports the estimates of the regression of $\widehat{\delta}_{1s}$ (the estimates of the coefficients on *Gender* from the alternative-specific conditional logit) on all attributes. Likewise, the other columns report estimates of the other regressions in system 3. As attributes enter utility interacted with personal characteristics, the parameter estimates reported in Table 6 are not the marginal effects of attributes and therefore difficult to interpret. Several coefficients estimates in this table are not significant. SURE estimates of Equation 3 excluding insignificant regressors remain similar to those obtained in Table 6.

Equation 7 states that the marginal effect of attributes on utility is a weighted average of the coefficient estimates in a row of Table 6 evaluated at the average value of individual characteristics. Table 7 reports two sets of marginal effects: a set of estimates using the estimated coefficients from Table 6 which include all attributes in each equation and another set of estimates using the estimated coefficients of a model that includes only those attributes previously found to be statistically significant.

It turns out the estimates of marginal effects change dramatically from one set of estimates to the other. The reason for this change is that some of the insignificant coefficients in Table 6 are nevertheless large in absolute value, thus contributing to change the magnitude of the marginal effect. Therefore, the following discussion focuses on the latter set of estimates. The *Basque Ratio,* which indicates the fraction of credits taught in Basque is significant with higher fractions of subjects taught in Basque increasing utility. Higher average *wages* among graduates from a particular degree, exert a positive and significant effect on utility. The degree of difficulty, as measured by the number of *credits* necessary to complete a particular degree also has a significant and positive effect on utility. However, the period of time necessary to *find a job* after graduation significantly and negatively affects utility. Whether the school is located in *Bizkaia* or *Guipuzkoa* (*Araba* is the reference), the *threshold grade* and the *Student/Teacher ratio* do not have significant marginal effects on utility.

Conclusions

This paper investigates the extend to which individual decisions on college degree choice depend on their personal characteristics and degree attributes. We propose a two-step method to estimate a discrete choice model that avoids the identification problem typically found when the number of alternatives is large and utility depends on both individual characteristics and alternative attributes.

Our two-step procedure is as follows. First, we estimate a random utility model by alternative-specific conditional logit using personal characteristics and alternative specific constants. By doing so, we are able to identify the effect of those characteristics on individuals' utility and analyze how they influence their demand for a College degree. As attributes and their interactions with individual-specific characteristics are not introduced in the first step, we avoid the identification problem mentioned above. In the second step, we regress the alternative-specific coefficient estimates obtained in the first step on alternative attributes allowing us to compute the marginal effects of attributes on expected utility.

We illustrate our methodology modeling the demand for College degree of students at the University of the Basque Country (UPV/EHU). The main results obtained are as follows. *Gender* significantly influences degree choice, women obtain higher utility from degrees related with education, health and languages, while men tend to choose degrees related to sports and engineering. Other important factors that have an effect on students' utility are the grades obtained in *Basque* and *Mathematics*. Students whose grades in mathematics are high, prefer to choose degrees that involve skills in *mathematics*, such as, *Mathematics*, *Chemistry*, *Biology* and *Industrial Engineering*. Similarly, obtaining high marks in *Basque* increases the probability of choosing degrees such as *Basque Philology*, *Child Education* and *Primary Education*.

D*istance* from town of residence to school is a significant factor in students' choices, with students preferring degrees taught closer to their place of residence. Besides the distance effect, there also exists a border effect inducing students to choose degrees taught in the same *province* where they live.

Besides the effect of personal characteristics on utility, our two-step procedure also allows us to measure the influence of degree attributes. In particular, students tend to choose degrees where a high percentage of subjects are taught in Basque. On the contrary, the student/teacher ratio does not appear to influence degree choice significantly. The difficulty of a particular degree, as measured by the number of *credits* necessary for completion, also has a positive and significant effect on utility. Finally, labor market conditions do influence degree choice: students tend to choose degrees with high expected *wages* and shorted spells of unemployment after graduation.

The effect of personal characteristics on College degree choice reported in this paper is in line with previous findings and simple intuition, thus providing further external validity to earlier evidence. In addition, this paper develops a procedure to account for the effect of alternative attributes on degree choices. The evidence reported in this regard is also in line with intuition. Our reading of these findings is that the procedure designed in this paper seems to be able to identify the effect of degree attributes on the demand for higher education.

Data Appendix

Personal characteristics

- *Gender* takes the value 0 if the individual is a woman and 1 if the individual is a man.
- *Age* measured in years.
- *Vehicular Language* takes the value 1 if the students took the admission exam in Basque and 0 otherwise.
- *Basque, Philosophy, History*, *Spanish*, *English*, and *Mathematics* equal the mark obtained in the corresponding part of the admission exam and zero if the student did not take that part of the exam.
- *Overall mark* obtained in the admission exam.
- *Global Grade* is the average between the grade obtained in the admission exam and the average grade obtained in High School.

Degree attributes.

- *Bizkaia* and *Gipuzkoa* are dummies that take the value 1 if the degree is taught in the province of Bizkaia / Gipuzkoa and 0 otherwise.
- *Threshold Grade* is the minimum grade needed to access to a degree.
- *Basque Ratio* is the percentage of credits that are offered in Basque over the total number of compulsory credits.
- *Student/Teacher Ratio* is the ratio of the number of students to the number of instructors in each School.
- *Wage* is the average wage of employed graduates three years after graduation.
- *Credits* records the number of credits necessary to complete a degree.
- *Find Job* is the average number of months graduates need to find a job from the moment they start looking for a job.

Variables that vary with students and degrees

- *Province* is a dummy that takes the value 1 if the student's residence is in the province where degree is taught.
- *Distance* between the town of residence of a student and the town where a degree is taught.

References

[1] Albert, C., 2000. Higher Education Demand in Spain: The Influence of Labour Market Signals and Family Background. *Higher Education,* 40: 147-162.

[2] Ayalon, H., 2003. Women and Men go to University: Mathematical Background and Gender Differences in Choice of Field in Higher Education. *Sex Roles,*48: 277-190.

[3] Bratti, M., 2002. Does the choice of university matter? A study of the differences across UK universities in life sciences students' degree performance. *Economics of Education Review*, 21: 431-443.

[4] Cepac, Z., Bojnec, S., 2010. Higher Education Demand Factors and the Demand for Tourism Education in Slovenia. *Organizacija,* 43: 257- 266.

[5] Drewes, T., Michael, C., 2006. How Do Students Choose a University?: An Analysis of Applications to University in Ontario, Canada. *Research in Higher Education,* 47(7): 781-800.

[6] Jimenez, J. D., Salas-Velasco, M., 2000. Modeling Educational Choices. A Binomial Logit Model Applied to the Demand for Higher Education. *Higher Education,* 40: 293-311.

[7] Gonzalez-Espitia, C.G., 2009. Desarrollos Recientes Sobre Demanda de Educación y sus Aplicaciones Empíricas Internacionales. *Borradores de Economía y Finanzas, Nº 29, Universidad ICESI.*

[8] Gonzalez López-Valcárcel, B., Dávila Quintana, D., 1998. Economic and Cultural Impediments to University Education in Spain. *Economics of Education Review,* 17: 93-103

[9] Greene, W.H., 2003. *Econometric Analysis (Fifth Edition),* Prentice Hall, New Jersey

[10] Iyengar, S. and Hahn, K.S., 2009. Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use. *Journal of Communication,* 59: 19-39

[11] Kim, H., Kwon N., 2003. The advantage of Network Size in Acquiring New Subscribers: a Conditional Logit Analysis of the Korean Mobile Telephony Market. *Information Economics and Policies*, 15: 17-33

[12] Long, B.T., 2004. How have College Decisions Changed Over Time? An Application of the Conditional Logistic Choice Model. *Journal of Econometrics,* 121: 271-296

[13] McFadden, D. L. 1974. Conditional logit analysis of qualitative choice behavior. In Frontiers in Econometrics, ed. P. Zarembka, 105–142. New York: Academic Press.

[14] Mazzanti, M., 2003. Discrete Choice Models and Valuation Experiments. *Journal of Economic Studies,* 30: 584-604

[15] Marcenaro Gutiérrez, O.D., Navarro Gómez, M.L., 2001. Un análisis microeconómico de la demanda de educación superior en España. *Estudios de Economía Aplicada,* 19: 69-86

[16] Marcerano-Gutierrez, O., Galindo-Rueda, F., Vignoles, A., 2007. Who Actually Goes to University?. *Empirical Economics,* 32: 333-357

[17] Maringe, F., 2006. University and Course Choice: Implications for Positioning, Recruitment and Marketing. *International Journal of Educational Management,*20(6): 466-479

[18] McCallum, J., 1995. National Borders Matter: Canadian-U.S. Regional Trade Patterns. *American Economic Review,* 85(3): 615-623.

[19] Montgomery, M., 2002. A Nested Logit Model of the Choice of a Graduate Business School. *Economics of Education Review,* 21: 471-480

[20] Murdock, J., 2006. Handling unobserved site characteristics in random utility models of recreation demand. *Journal of Environmental Economics and Management*, 51: 1-25

[21] Rahona López, M., 2006. La influencia del entorno socioeconómico en la realización de estudios universitarios: una aproximación al caso español en la década de los noventa. *Hacienda Pública Expañola/ REvista de Economía Pública*, 178: 55-80

[22] Rice, P.G., 1987. The Demand for Post-compulsory Education in the UK and the Effects of Educational Maintenance Allowances. *Economica,* 54: 467-475

[23] Salas-Velasco, M., Martín-Cobos Puebla, M., 2006. La Demanda de Educación Superior: Un Análisis Microeconómico con Datos de Corte Transversal. *Revista de Educación,* 339: 637-660

[24] Strenta, A.C., Elliott, R., Adair, R., Matier, M., Scott, J., 1994. Choosing and Leaving Science in Highly Selective Institutions. *Research in Higher Education,* 35: 513- 547

[25] Timmins, C., Murdock, J., 2007. A Revealed Approach to the Measurement of Congestion in Travel Cost Models. *Journal of Environmental Economics and Management,* 53: 230-249

Table 1: Descriptive statistics of individual characteristics

Individual characteristics	Observations	Mean	Standard deviation	Minimum	Maximum
Age	5029	18.1873	0.7406	17	42
Gender	5029	0.4277	0.4948	0	1
Vehicular Language	5029	0.5864	0.4948	0	1
Mathematics	5029	4.9312	3.0649	0	10
History	5029	2.7796	3.4948	0	10
Spanish	5029	5.8042	1.4219	1.2	10
English	5029	6.6213	1.8684	0	10
Basque	5029	7.0135	2.3997	0	10
Philosophy	5029	3.7831	3.3835	0	10
Selectividad	5029	6.7822	1.0519	5	9.71
Global Grade	5029	6.7820	1.0517	5	9.711

Alternative attributes	Observations	Mean	Standard deviation	Minimum	Maximum
Bizkaia	77	0.4675	0.5022	0	1
Gipuzkoa	77	0.2857	0.4547	0	1
Threshold Grade	77	5.4657	0.9989	4.6	8.27
Basque Ratio	77	0.8805	0.2337	0	1
Student/Teacher Ratio	76	17.1316	19.7716	1	124
Wage	75	1537.775	198.7159	1140	1993
Credits	77	241.0487	72.3719	182	702
Find Job	75	3.9333	1.9406	1	8

Table 2: Gender by fields of knowledge

	Women	Men	Women	Men
	Number		Fraction	
Humanities	168	112	0.0583	0.0521
Experimental Sciences	171	134	0.0594	0.0623
Social Sciences and Law	1,485	845	0.5159	0.3984
Technical Studies	348	848	0.1209	0.3923
Health Sciences	706	212	0.2431	0.0986
TOTAL	2,878	2,151	1	1

Table 3: Average marks by field of knowledge

	Humanities	Experimental Sciences	Social Sciences and Law	Technical Studies	Health Sciences
Mathematics	4.8921	6.4020	4.8074	6.8596	6.5414
English	6.7856	6.9221	6.2192	6.9577	7.3214
Spanish	5.8306	6.0179	5.6046	5.9675	6.2163
Basque	7.5461	7.6693	7.2118	7.1852	7.4850
Philosophy	6.3849	6.7382	6.2051	6.5122	7.0110
History	6.8694	6.8167	6.3752	6.8341	7.2684
Overall	6.7089	6.9810	6.3781	7.0027	7.4779

Table 4: Province of residence and Campus choice

	Campus choice		
Province of residence	Bizkaia	Gipuzkoa	Araba
Bizkaia	2,170	168	281
Gipuzkoa	410	942	270
Araba	347	105	292
La Rioja	4	1	2
Navarra	2	2	0
Other	30	2	1

Number of students

Table 7: Marginal effects of attributes on utility

	SURE (unrestricted)		SURE (restricted)	
	estimate	p-value	estimate	p-value
Bizkaia	-0.1066	(0.8102)	0.2622	(0.8312)
Gipuzkoa	0.1555	(0.4737)	0.8795	(0.4737)
Threshold Grade	1.1768	(0.0180)	-0.5023	(0.3518)
Basque Ratio	3.8264	(0.0251)	4.5138	(0.0168)
Student/Teacher Ratio	0.0023	(0.7484)	-0.0680	(0.5441)
Wage	-0.0047	(0.0707)	0.0014	(0.0000)
Credits	0.0025	(0.0874)	0.0211	(0.0000)
Find Job	-0.4345	(0.0598)	-0.3559	(0.0128)

Table 5: First-step estimates

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
					Grades in admission exam								
Alternative	Degree	Age	Gender	Vehicular Language	Basque	Philosophy	History	English	Spanish	Mathematics	Selectividad	Global Grade	Constant
2	Architect	0	0	0	0	0	0	0	0	+	0	0	0
3	Telecommunications Engineering	0	+	0	0	0	0	0	0	+	0	0	0
4	Industrial Engineering	0	+	0	0	0	0	+	0	+	0	0	0
5	Computer Engineering	0	+	0	0	0	-	0	0	+	0	0	0
6	Chemical Engineering	-	0	0	0	0	0	0	0	+	0	0	+
7	Chemical Engineering	0	0	0	+	0	0	0	0	+	0	0	0
8	Pharmacy	0	-	0	0	0	0	0	0	0	0	0	0
9	Fine Arts	0	-	0	0	-	-	0	0	-	0	0	0
10	Psychology	0	-	0	0	0	0	0	0	-	0	0	0
11	Law	0	-	0	0	0	0	0	0	-	0	0	0
12	Law	0	-	0	0	0	0	0	0	-	0	0	0
13	Dentistry	0	-	0	0	0	0	0	0	0	0	0	0
14	Medicine	0	-	0	0	0	0	0	0	-	0	+	-
15	Physics	0	0	0	0	-	-	0	0	+	0	+	0
16	Mathematics	0	0	0	0	0	0	0	0	+	0	0	-
17	Biology	0	0	0	0	0	0	0	0	-	0	0	0
18	Business	0	-	0	0	0	0	0	0	0	0	0	0
19	Political Science	0	0	0	0	0	+	0	+	0	0	0	0
20	Sociology	0	0	0	0	0	0	0	0	-	0	0	0
21	Marketing and Public Relations	0	-	0	0	0	0	0	0	-	0	0	0
22	Translation and Interpretation	0	-	0	0	0	0	0	0	-	0	+	0
23	Philosophy	0	0	0	0	0	-	0	+	0	0	0	0
24	German/Classic/French Philology	-	0	0	0	0	-	-	+	-	0	0	+
25	Hispanic Philology	0	0	0	0	0	0	0	0	-	0	+	-
26	English Philology	0	0	0	0	0	0	+	0	-	0	0	0
27	Basque Philology	0	0	0	+	0	0	0	0	-	0	0	0
28	History	0	0	0	+	0	0	0	0	-	0	0	0
29	History of Art	0	0	0	0	0	0	0	0	-	0	0	0
30	Audiovisual Communication	0	-	0	0	0	0	0	0	-	0	0	0
31	Journalism	0	-	0	+	0	0	0	+	-	0	0	0
32	Pedagogy	0	-	-	+	0	0	0	0	0	0	0	0
33	Chemistry	-	0	0	0	0	0	0	0	+	0	0	+
34	Chemistry	0	0	0	0	0	-	0	0	0	0	0	0
35	Environmental Science	0	0	0	0	0	0	0	0	-	0	0	0
36	Physical Activity and Sport Science	0	+	0	+	0	0	-	0	0	0	0	0
37	Nursing	0	-	0	0	0	0	0	0	-	0	0	0
38	Nursing	0	-	0	0	0	0	0	0	-	0	0	0
39	Social Work	0	-	0	+	0	0	0	+	0	0	0	0
40	Business†	0	0	0	0	0	0	0	0	0	0	0	0
41	Business†	0	0	0	+	0	0	-	0	0	0	0	0
42	Business†	0	0	0	0	0	-	0	0	0	0	0	0
43	Labor Relations†	0	-	0	0	0	0	0	0	0	0	0	0

Table 5 (Continued): First-step estimates

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
					Grades in admission exam								
Alternative	Degree	Age	Gender	Vehicular Language	Basque	Philosophy	History	English	Spanish	Mathematics	Overall Mark	Global Grade	Constant
44	Social Education†	0	-	0	0	0	0	0	0	-	0	0	0
45	Social Education†	0	-	0	+	0	0	0	0	-	0	0	0
46	Maritime Navigation†	0	0	0	0	0	0	0	0	0	0	0	0
47	Nutrition and Dietetics†	0	-	0	+	0	0	0	+	0	0	0	0
48	Child Education†	+	-	0	+	0	0	0	0	-	0	0	0
49	Child Education†	0	-	+	+	0	0	-	0	0	0	0	0
50	Child Education†	0	-	0	+	0	0	-	0	-	0	0	0
51	Primary Education†	0	-	0	+	0	0	0	+	0	0	0	0
52	Primary Education†	0	-	0	0	0	0	0	0	-	0	0	0
53	Primary Education†	0	-	0	+	0	0	-	0	-	0	0	0
54	Foreign Language Education†	-	-	0	0	-	-	0	0	-	0	0	+
55	Foreign Language Education†	0	-	0	0	0	0	+	0	-	0	0	0
56	Physical Education†	0	+	+	+	0	0	-	0	-	0	0	0
57	Musical Education†	0	0	0	+	0	0	0	0	-	0	0	0
58	Special Education†	0	-	0	0	0	0	0	0	0	0	0	0
59	Engineer in Topography†	0	0	0	0	0	0	0	0	0	0	0	0
60	Computer Science†	0	0	0	+	0	0	0	0	0	0	0	0
61	Computer Science†	0	+	0	0	0	0	0	0	0	0	0	0
62	Technical Architect†	0	0	0	+	0	0	0	0	+	0	0	0
63	Engineering, Civil Construction†	0	+	0	0	0	0	0	0	+	0	0	0
64	Engineering, Urban Trans. and Ser.†	0	0	0	0	0	0	0	0	+	0	0	0
65	Engineering, Mechanics†	0	+	0	0	0	0	0	0	+	0	0	0
66	Engineering, Mechanics†	0	+	0	0	0	0	0	0	+	0	0	0
67	Engineering, Mechanics†	0	+	0	0	-	-	0	0	+	0	0	0
68	Engineering, Chemistry†	-	0	0	0	0	0	0	0	+	+	-	+
69	Engineering, Electronics†	0	0	0	0	0	0	0	0	+	0	0	0
70	Engineering, Electronics†	0	+	0	0	-	-	0	0	+	0	0	0
71	Engineering, Electronics†	0	0	0	0	0	0	0	0	0	0	0	0
72	Engineering, Electronics†	0	+	0	0	0	0	0	0	+	0	0	0
73	Engineering, Electricity†	-	0	0	+	0	0	0	0	0	+	-	+
74	Engineering, Electricity†	0	0	0	0	0	0	0	0	0	0	0	0
75	Engineering, Telecomm./Telematics†	0	+	-	0	0	0	0	0	+	0	0	0
76	Engineering, Telecomm./Systems†	-	0	0	0	0	-	0	0	+	+	-	+
77	Mining Engineering†	0	0	0	0	0	0	0	0	0	0	0	0

Covariates that vary with individual and alternative	
Province	0.5579***
	(0.0723)
Distance	-0.0160***
	(0.0012)

† Three year degree. Same degree with different alternative number indicates different campus. Alternative 1, “Economics” is used as reference. One, two and three asterisks correspond to 10, 5 and 1 per cent significance levels.

Table 6: Second step estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
					Grades in admission exam							
	Constant	Gender	Age	Vehicular Language	Overall Mark	Mathematics	History	Spanish	Basque	English	Philosophy	Global Grade
Bizkaia	7.989*	0.485	-0.459*	-0.440**	0.476	0.264***	-0.004	-0.072	-0.093**	0.062	-0.024	-0.511
	(4.602)	(0.344)	(0.238)	(0.202)	(0.496)	(0.077)	(0.108)	(0.047)	(0.045)	(0.050)	(0.042)	(0.543)
Gipuzkoa	13.795**	0.238	-0.775**	-0.398	0.599	0.279***	-0.048	-0.008	-0.048	0.018	0.036	-0.758
	(6.415)	(0.392)	(0.367)	(0.246)	(0.619)	(0.082)	(0.174)	(0.048)	(0.063)	(0.047)	(0.048)	(0.652)
Threshold Grade	-5.246*	-0.733***	0.317**	-0.052	-0.265	-0.034	0.135	-0.021	0.044**	0.005	0.003	0.348
	(2.863)	(0.151)	(0.155)	(0.090)	(0.214)	(0.043)	(0.097)	(0.020)	(0.022)	(0.025)	(0.019)	(0.236)
Basque Ratio	-35.269*	0.135	2.106**	-0.673	0.700	-0.079	0.157	-0.184*	0.116	-0.075	-0.182	-0.328
	(18.295)	(0.567)	(1.027)	(0.547)	(0.968)	(0.212)	(0.394)	(0.092)	(0.103)	(0.129)	(0.131)	(1.091)
Student/Teacher	0.111*	-0.005	-0.005	0.001	-0.015*	0.001	0.000	0.001	-0.001	0.000	0.000	0.010
	(0.063)	(0.005)	(0.004)	(0.004)	(0.008)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.008)
Wage	0.024	0.002***	-0.002	0.001***	0.001	-0.000	-0.001	-0.000	-0.000	0.000	0.000	-0.001
	(0.019)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Credits	-0.026	-0.001	0.000	-0.002**	-0.003	-0.000	0.000	0.000	-0.001**	0.000	0.000	0.007***
	(0.022)	(0.002)	(0.001)	(0.001)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Find Job	1.062	-0.225***	-0.032	0.204***	-0.146	-0.065**	-0.045	0.018*	-0.005	-0.031***	0.023**	0.081
	(1.671)	(0.079)	(0.086)	(0.044)	(0.110)	(0.027)	(0.057)	(0.010)	(0.013)	(0.009)	(0.011)	(0.110)
Constant	22.896	1.679	-1.065	-1.038	0.409	0.636*	0.339	0.476**	0.223	-0.091	-0.161	-2.399
	(30.026)	(1.359)	(1.684)	(0.923)	(2.457)	(0.321)	(0.551)	(0.194)	(0.191)	(0.210)	(0.215)	(2.684)

Robust standard errors in parentheses. One, two and three asterisks correspond to 10, 5 and 1 per cent significance levels. Number of observations = 73.