INSTITUTO DE ECONOMÍA

DOCUMENTO de TRABAJO

403

2011

Gender Differences in Major Choice and College Entrance Probabilities in Brazil

Alejandra Traferri.

www.economia.puc.cl • ISSN (edición impresa) 0716-7334 • ISSN (edición electrónica) 0717-7593

Versión impresa ISSN: 0716-7334
Versión electrónica ISSN: 0717-7593

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA

Oficina de Publicaciones
Casilla 76, Correo 17, Santiago
www.economia.puc.cl

GENDER DIFFERENCES IN MAJOR CHOICE AND COLLEGE ENTRANCE PROBABILITIES IN BRAZIL

Alejandra Traferri*

Documento de Trabajo Nº 403

Santiago, Noviembre 2011

* atraferri@uc.cl

INDEX

GENDER DIFFERENCES IN MAJOR CHOICE AND COLLEGE ENTRANCE PROBABILITIES IN BRAZIL*

ALEJANDRA TRAFERRI†

Abstract. I study gender differences in major choice and college entrance probabilities in University of Campinas, a Brazilian public university dependent on the State of São Paulo. As with most Brazilian public universities, students select a major, and then compete for a place in that major by taking a major-specific entrance exam. This singular characteristic of the Brazilian case allows me to differentiate the effect of gender on major-specific entrance probabilities and preferences. I propose a model and econometric strategy which can account for two important issues, selectivity bias and the fact that expected utility depends on the probability of entering the different majors. I find evidence of gender differences in preferences and entrance probabilities. For most majors, gender differences in major choice are mostly explained by differences in preferences. However, for the most demanding majors (those that require higher grades from students), differences in major choice are explained in a large proportion by differences in entrance probabilities. Finally, I find that gender has important interactions with other variables. In particular, gender effects depend on education, socioeconomic characteristics and family background.

Keywords: Major Choice, Gender Differences, College Entrance Test, Vestibular, Brazilian Universities (JEL C35, I21, J24).

Date: This Draft: September 2011.

* I am grateful to Raquel Carrasco for her guidance and advice. I am also grateful to University of Campinas' Permanent Commission for Vestibulars (Comvest) for providing the data for this paper, especially to Mauricio Kleinke, Renato Pedrosa, and Leandro Tessler for granting access to the data and providing several important clarifications. I would like to thank Fernanda Estevan, Gaston Llanes, and Marian Vidal-Fernandez for useful comments at different stages of this research. Finally, I would also like to thank the Institute for Economic Development at Boston University, in which I conducted part of this research as a visiting scholar.

† Instituto de economía, Pontificia Universidad Católica de Chile, Avda. Libertador Bernardo OHiggins 340, Santiago, Chile. E-mail: atraferri@uc.cl.

1. Introduction

There are significant differences in the average major choices of men and women. This issue has been analyzed extensively for the American case (see, for example Freeman 1971, Turner and Bowen 1999, Zafar 2009), but the same pattern is present in Brazilian universities. In the case of the University of Campinas, for example, the proportion of men choosing engineering majors between 2006 and 2008 was 67 percentage points higher than the proportion of women choosing that major, and the proportion of women choosing Medicine was 28 percentage points higher than the proportion of men choosing that major (see Table 3 for more details).

The choice of college major depends not only on utility considerations (including expected earnings), but also on the relative advantages of each student. *Ceteris paribus*, students with a relative advantage in math will tend to do better in majors which make intensive use of mathematics, and therefore we expect them to choose engineering majors more often than students who have a relative advantage in verbal skills, for example.

Accordingly, men and women may choose different majors partly because they have different preferences or different relative advantages[1]. Moreover, gender differences may be affected by education, socioeconomic characteristics and family background. For example, gender differences in major choice may be smaller or larger for students who attended public schools, in comparison with students who attended private schools.

Therefore, it is important to know what is the relative importance of gender differences and other individual characteristics in explaining major choice and academic performance. Does gender have an effect on the probability of entering college once we take into account other individual characteristics? Do men and women choose different majors

[1]Several hypothesis have been proposed in the literature as to why these gender differences may arise. The fertility hypothesis says that women know that their work life will be interrupted when they have children, so the economic value of particular careers may be lower. The socialization hypothesis states that men and women are taught to be different since their infancy, and this has an effect on what they perceive their role in society should be. Finally, gender discrimination may also play a role, not only by generating a difference in expected earnings for particular careers, but also by pre-empting the entry of women (men) to careers traditionally dominated by men (women). See Turner and Bowen 1999 for a more detailed discussion. However, it is important to remark that, independently of the reasons why gender differences in preferences or academic performance, it is important to know whether these differences do, in fact, exist.

because they have different preferences or different relative advantages, or because they have received a different education and come from different socioeconomic backgrounds?

In past works, it has been difficult to find an answer to the previous questions because in most countries students are allowed to enter college (or not) before choosing major. Therefore, it is in general not possible to discern whether the major choice was motivated by the student preferring that choice more than others, or because the student believes she will do better in that major.

A good survey of previous works is Turner and Bowen (1999). Their discussion on the possible causes of gender differences makes clear that it is generally very difficult to determine the differential impact of each factor. They use data on SAT (Scholastic Aptitude Test) scores for American college students, and show that an important part of the gender gap is explained by differences in SAT scores. Other papers, like Altonji (1993) and Arcidiacono (2004), present dynamic models to study the choice of college and major. However, they are mainly concerned with the effects of differences in predicted earnings, and do not take into account differences in the probability of entering or finishing the different majors.

An important precedent for this paper is Montmarquette et al. (2002), who use an econometric approach which is closer to the one used in this paper. They present a model in which students take into account the probability of graduating when choosing major. They estimate this probability through a linear probability model, and then introduce the estimated probability in a multinomial logit of choice of major. However, they do not allow for correlation between the errors of the probability equation and the expected utility of each major, which implies that their analysis is prone to sample selection bias. As I will show, an alternative econometric strategy can account for correlation between both equations.

In this sense, the Brazilian case is particularly interesting. In most Brazilian public universities, the student chooses a major before taking a major-specific exam, which determines whether the student is allowed to enroll in the major or not. Therefore, when choosing among the different majors, the student takes into account not only the utility corresponding to each major, but also the associated probabilities of entry. This implies that we can perform a separate study of the factors affecting the probability of entry and the choice of major.

The determinants of major choice have been scarcely explored in the Brazilian case. A few papers analyze the determinants of performance in Entrance Test Exams (Guimarães and Sampaio 2007, 2008, Calvacanti et al. 2009), but the choice of major has not been analyzed in detail. Such analysis of gender differences in choice of major is important because of its possible relation with gender inequality. This is an important topic of research for the Brazilian Federal Government, which is currently designing public policies to reduce gender inequality. For example, the Federal Government has recently introduced over 400 projects directed at enhancing equal opportunities for men and women, which will be performed by 22 government institutions between 2008 and 2011 (Pinheiro et al. 2008).

In this paper, I estimate a model of major choice and college entry using data from entrance tests of the University of Campinas between 2006 and 2008. The basic model (Model I) consists of two estimations. First, I estimate a binary logit to study the determinants of the probability of entering the different majors. Then, I estimate a multinomial logit model of major choice. An important determinant of expected utility is the probability of entering the major. I calculate these probabilities from the estimations of the first step.

In the basic model, I assume the errors of the entry and expected utility equations are uncorrelated, so there is no selectivity bias by assumption. In other words, the fact that a student chooses a particular major does not mean that this student has a higher probability of entering that major, in comparison with a similar student who chose another major. In an extension (Model II), I allow for correlations between the errors of both equations, so that students with a larger entry shock for a particular major tend to have a larger preference shock for the same major. Model I is estimated through Maximum Likelihood in two steps. Model II is estimated through Maximum Simulated Likelihood in one unique step.

When estimating the second model, I find that the correlation between the errors of the two equations is positive and significant. Therefore, students who get a larger preference shock for some major tend to have a larger entry shock for that major too. Given that the coefficient is significant, the model without correlations will produce biased estimators. Therefore, it is important to consider correlated errors in the econometric design.

I find evidence of gender differences in the probability of entering the different majors. Controlling for other individual characteristics, men have on average a higher probability of entering some majors, and women have a higher probability of entering other majors. Interestingly, the effect of gender depends on past academic performance, given that for most majors, the interaction between gender and the ENEM grade is significant.[2]

I also find a significant effect of gender on major choice. The largest differences between men and women arise in the most demanding majors (i.e. those with highest minimum required grades). Nevertheless, there are significant differences between the choices of men and women in other majors as well. Overall, men have a higher probability of choosing mathematically-oriented majors (Technologies, Exact Sciences, and Engineering and Architecture), and women have a higher probability of choosing majors in Natural and Earth Sciences, Arts, Humanities, and Health and Biological Sciences. In addition, I find that the effect of gender on major choice depends on student characteristics. In particular, the size of gender differences depends on work status, type of secondary school and income, among other variables.

In order to determine if gender differences in major choice are caused by differences in preferences or probability of entry, I simulate women choices with male probabilities of entry, and men choices with female probabilities of entry. I find that preferences account for most of the difference in choices in majors with low or medium minimum required grades. In the most demanding majors, on the other hand, a large part of the difference in major choice is explained by differences in the probability of entry.

The rest of the paper proceeds as follows. In Section 2, I present the models to be estimated and the estimation strategy. In Section 3, I describe the process of major selection and the characteristics of the entrance exam for University of Campinas. In Section 4, I present an introduction to the data and show descriptive statistics of the sample and variables used in the estimations. In Section 5, I show and discuss the results of the estimations. Finally, Section 6 presents the conclusions of the paper.

2. Model and estimation strategy

In this section, I present a model to study the decisions of N individuals ($i = 1, 2, \ldots N$), choosing among J majors ($j = 1, 2, \ldots J$). Let $M_{ij} \in \{0, 1\}$ be a binary indicator, which is equal to 1 if individual i chooses major j and 0 otherwise, and $E_{ij} \in \{0, 1\}$ be another binary indicator, which is 1 if individual i enters major j and 0 otherwise.

[2]ENEM (Exame Nacional do Ensino Médio, High School National Exam) is a non-mandatory Brazilian national exam, which examines students' knowledge of concepts taught in secondary school.

Clearly, M_{ij} will be equal to 1 for exactly one major j, and we observe E_{ij} only for the major the individual has actually chosen.

Individuals choose a major in order to maximize their expected indirect utility, which depends on the utility of entering the major and the probability of entering the major, which in turn depend on individual characteristics. Specifically, let $U_{ij} = \alpha_j \, x_i$ be the utility of individual i of entering major j, and let $p_{ij} = Pr(E_{ij} = 1|x_i)$ be the probability of entering major j, where x_i is a k_x-vector of individual characteristics (including a 1 for the intercept), and the α_j are k_x-vectors of parameters.

The expected indirect utility of individual i from choosing major j is:

$$u^*_{ij} = p_{ij} \, U_{ij} + \varepsilon_{ij}, \tag{1}$$

where ε_{ij} is an individual-major taste shock, which is unobserved by the econometrician, but known to the individual when choosing among the different majors. Introducing the expression for utility and rearranging equation (1) we get:

$$u^*_{ij} = p_{ij} \, \alpha_j \, x_i + \varepsilon_{ij}. \tag{2}$$

According to equation 1, the utility assigned to a given major depends on the probability of entering that major. This probability depends on the characteristics of the individual, and is determined by a binary model with latent equation:

$$y^*_{ij} = \gamma_j \, z_i + \eta_{ij}, \tag{3}$$

where z_i is a k_z-vector of individual characteristics, possibly overlapping with x_i, γ_j is a vector of major-specific parameters, and η_{ij} is an error term which is unobserved (or partially observed) by the individual when choosing major, and unobserved by the econometrician before and after the individual chooses major. z_i includes a 1 for the intercept.

As usual, we cannot observe the latent variables u^*_{ij} and y^*_{ij}. The rules determining the observed variables are:

$$\begin{aligned} M_{ij} &= 1\left[u^*_{ij} > \max_{k=1,\dots J,\, k \neq j} u^*_{ik}\right], \\ E_{ij} &= 1\left[y^*_{ij} > 0\right], \end{aligned}$$

where $1[\,.\,]$ is an indicator function.

There are two difficulties in estimating the above model. First, we only observe E_{ij} when $M_{ij} = 1$. Therefore, there will be a selection process if the errors in equations (1) and (3) are correlated. Second, the first latent equation depends on the parameters of the second latent equation, through the probability of entering the major.

2.1. **Basic model.** To complete the description of the model, we need some assumption about the error terms. I will start by assuming independence of the error terms (Model I), and then relax this assumption by introducing correlations between the error terms (Model II). The benchmark model is characterized by the following assumption:

Assumption 1 (Model I)**.** *ε_{ij} are i.i.d. according to a double exponential distribution, and have zero mean and variance equal to $\pi^2/6$. η_{ij} are i.i.d. with a cumulative density function F, and have zero mean and unit variance. ε_{ij} and η_{ik} are independent for any j and k.*

The first part of Assumption 1 corresponds to what is known as the multinomial logit Model (MNL, McFadden 1974).[3] Alternatively, I could have assumed a multivariate normal distribution for ε_{ij}, which would have yielded a multinomial probit Model (MNP). The advantage of the MNL is that it provides closed form solutions for the probabilities, and is therefore more tractable (the MNP usually requires numerical integration for solving multiple integrals, which becomes unfeasible when the number of alternatives is large). As is well known, the main disadvantage of the MNL is the IIA (Independence of Irrelevant Alternatives) property.[4]

I will use the MNL for two reasons. First, it allows for comparison with previous studies of major choice. Turner and Bowen (1999), Montmarquette et al. (2002) and Arcidiacono (2004), for example, use the MNL as their discrete choice model. Second, in the following section, I will generalize the model to allow correlations between the errors of equations (1) and (3), which will eliminate the IIA property. Using the MNL for this baseline estimation allows me to compare its results with the alternative specification.

[3]Some authors also refer to this model as the Conditional logit Model, but it is more appropriate to use the term multinomial logit for the case in which the model is derived from utility maximization.

[4]The IIA property requires that the relative odds ratio between two alternatives does not change when a new alternative is added to the set of alternatives or when the characteristics of a third alternative change. In the case of the MNL, the ratio of probabilities of two events is

$$\frac{P_{ij}}{P_{ik}} = \frac{\exp(p_{ij}\,\alpha_j\,x_i)}{\exp(p_{ik}\,\alpha_k\,x_i)}.$$

It is easy to see that this ratio does not depend on the utility parameters of the other choices, which implies that the MNL has the IIA property. See Ben-Akiva and Lerman (1985) and Anderson, De Palma, and Thisse (1992) for more details.

With respect to the distribution of η_{ij}, I could use a normal (probit) or logistic (logit) distribution. The choice between the binary logit and probit models is largely one of convenience and convention, since the substantive results are generally indistinguishable. For the purpose of this paper, I will use the logistic distribution, because of its tractability.

Model I is the easiest model that can be estimated. The errors are independent, which means there is no selection problem. In other words,

$$Pr(E_{ij} = 1|x_i, M_{ij} = 1) = Pr(E_{ij} = 1|x_i),$$

and the same is true for the expectations. It is important to understand the meaning of this assumption: the fact that an individual chooses a given major does not give any information as to whether he has a higher probability of entering that major than other similar students who are not choosing that major.

Under Assumption 1, the probability of entering major j is

$$\begin{aligned} p_{ij} &= Pr(E_{ij} = 1|x_i) \\ &= Pr(\gamma_j\, z_i + \eta_{ij} > 0) \\ &= \frac{1}{1 + \exp(-\gamma_j\, z_i)}, \end{aligned} \tag{4}$$

and the probability of choosing major j is

$$\begin{aligned} P_{ij} &= Pr(M_{ij} = 1|x_i) \\ &= \frac{\exp(p_{ij}\, \alpha_j\, x_i)}{\sum_{j=1}^{J} \exp(p_{ij}\, \alpha_j\, x_i)}. \end{aligned} \tag{5}$$

The estimation approach is very simple. Given that the parameters of the probability of entering a major are needed to determine the probability of choosing that major, I first run a binary logit of E_{ij} on z_i, for each major j using only the observations of individuals who choose that specific major. Then, we use these estimations to run a multinomial logit of M_{ij} on $p_{ij}\, x_i$, using all observations.

It is important to remark that the proposed two-step procedure will give unbiased estimators only if Assumption 1 is correct. If the errors of the choice and entry equations are correlated, then there will be a selection process which the two-step procedure will not take into account, producing biased estimators. Notice, however, that this has been the approach taken by previous papers examining gender differences in major choice. In the next section, I present an alternative approach which will yield unbiased estimators for a specific correlation structure. Designing an econometric model and estimation strategy to account for more general correlation structures is an interesting topic for further research, but is beyond the scope of the present paper.

In terms of identification, the parameters of the model are already identified because of the non-linear functional form of probability, and because entrance probabilities enter utility in a multiplicative form. However, as I describe in Section 4 I will include exclusion variables in the choice and entry equations, which will also help in the identification of the parameters. In particular, some variables, like the one describing whether the student has taken a private preparation course for the entrance test, affect only the probability of entering a major. Other variables, like the variables describing the reasons why the student chose the particular major, affect only preferences.

2.2. **Model with correlations.** In this section we study what happens if ε_{ij} and η_{ij} are correlated. Specifically, let $\varepsilon_{ij} = \nu_{ij} + \sigma\,\mu_{ij}$ and $\eta_{ij} = e_{ij} + \sigma\,\mu_{ij}$, where μ_{ij} is the common factor affecting the taste and entry shocks. If σ is significant, then students with a higher entry shock for some major tend to have a higher taste shock for that major too.

It is important to determine what the individual and the econometrician know before choosing a major. ν_{ij} and μ_{ij} are known by the individual before choosing a major, but are unobserved by the econometrician. e_{ij} is not known by the individual (at least before taking the exam), nor by the econometrician (before and after the exam).[5]

Assumption 2 (Model II)**.** *ν_{ij} are i.i.d. according to a double exponential distribution, with zero mean and variance $\pi^2/6$. e_{ij} are i.i.d. with cumulative density function F, and have zero mean and unit variance. μ_{ij} are i.i.d. according to a cumulative density function G with mean 0, probability density function g, and unit variance. μ_{ij}, ν_{ik} and e_{ih} are independent for any j, k, and h.*

The econometric approach is inspired in the mixed logit, which allows the parameters to differ among individuals.[6] McFadden and Train (2000) show that the mixed logit can approximate any discrete choice process, by appropriately choosing the distribution G, and that the mixed logit eliminates the well known problem of the Independence of Irrelevant Alternatives (IIA) of the Conditional and multinomial logit

[5]An alternative way to introduce correlation would have been to assume that the ε_{ij}'s and η_{ij}'s have a joint normal distribution, and to allow the covariance matrix to be non-diagonal.

[6]Read chapter 6 of Train (2003) for a good description of the mixed logit.

models.[7] Usually, a normal distribution is used when the parameters can take positive or negative values, and a lognormal distribution is used when the parameters must have a specific sign. For the purposes of this paper, μ_{ij} will follow a normal distribution.

Estimation of Model II is not as straightforward as the previous case. The problem is that the correlation between the errors implies that $Pr(E_{ij} = 1|x_i, M_{ij} = 1) \neq Pr(E_{ij} = 1|x_i)$, so the estimation of the second equation using only the observations of individuals who chose a particular major gives biased estimates. For this reason, the two-step procedure cannot be used and we have to estimate the whole system in one step.

The first step is to construct the likelihood function. Let μ_i be a J-vector containing the common factors for all majors, and suppose μ_i is observed by the econometrician. This means that μ_i becomes a variable in the estimation, just as one of the x_i or z_i. Let $P_{ij}(\mu_i) = Pr(M_{ij} = 1|x_i, \mu_i)$ and $p_{ij}(\mu_{ij}) = Pr(E_{ij} = 1|x_i, \mu_{ij})$.[8] Under our assumptions, we have that:

$$Pr(E_{ij} = 1|x_i, \mu_{ij}) = Pr(E_{ij} = 1|x_i, M_{ij} = 1, \mu_{ij}). \tag{6}$$

According to equation 6, if the model is correctly specified (i.e. if the errors of the choice and entry equations are correlated according to Assumption 2) and we were able to observe μ_i, then there would not be a selection problem when estimating the probabilities of entry. In other words, if Assumption 2 is correct, once we control for μ_i, the fact that a student chooses a major does not mean that she has a different probability of entering that major than a student with identical x_i, z_i and μ_i who did not chose that major.

There are $2J$ possible events for which we have to find a probability: the probability of choosing major j and entering, and the probability of choosing major j and not entering, for each j. Given μ_i, the probability

[7]Under Assumption 2, the ratio of probabilities of two events is

$$\frac{\mathcal{P}_{ij}}{\mathcal{P}_{ik}} = \frac{\int_{-\infty}^{\infty} \frac{\exp(p_{ij}\,\alpha_j\,x_i)}{\sum_{j=1}^{J}\exp(p_{ij}\,\alpha_j\,x_i)}\, g(\mu_i)\,d\mu_i}{\int_{-\infty}^{\infty} \frac{\exp(p_{ik}\,\alpha_k\,x_i)}{\sum_{j=1}^{J}\exp(p_{ij}\,\alpha_j\,x_i)}\, g(\mu_i)\,d\mu_i},$$

where μ_i is a J-vector containing the $\mu_i j$ for $j = 1, \ldots J$. Clearly, this ratio depends on the utility parameters of the other choices, which implies that the IIA property does not hold under Assumption 2.

[8]Notice that P_{ij} depends on the common factors for all majors (μ_i), while p_{ij} depends only on the shock for major j (μ_{ij}).

of choosing major j and the probability of entering major j are

$$\begin{aligned} P_{ij}(\mu_i) &= \frac{\exp(p_{ij}\,\alpha_j\,x_i + \sigma\,\mu_{ij})}{\sum_{j=1}^{J}\exp(p_{ij}\,\alpha_j\,x_i + \sigma\,\mu_{ij})} \\ p_{ij}(\mu_{ij}) &= F(\gamma_j\,z_i + a_i + \sigma\,\mu_{ij}). \end{aligned}$$

Then, the probability of choosing major j and entering is $P_{ij}(\mu_{ij})\,p_{ij}(\mu_{ij})$, and the probability of choosing major j and not entering is $P_{ij}(\mu_{ij})\,(1-p_{ij}(\mu_{ij}))$.

Of course, we do not observe μ_i. Nevertheless, we know its distribution, so we can integrate out the μ_i, i.e. calculate the expected value of the above probabilities:

$$LE_{ij} = \int_{-\infty}^{\infty} P_{ij}(\mu_i)\,p_{ij}(\mu_{ij})\,g(\mu_i)\,d\mu_i \tag{7}$$

$$LN_{ij} = \int_{-\infty}^{\infty} P_{ij}(\mu_i)\,(1-p_{ij}(\mu_{ij}))\,g(\mu_i)\,d\mu_i. \tag{8}$$

With LE_{ij} and LN_{ij} we obtain the following log-likelihood function:

$$L = \sum_{i=1}^{N}\sum_{j=1}^{J} M_{ij}\,\left(E_{ij}\,\log(LE_{ij}) + (1-E_{ij})\,\log(LN_{ij})\right)$$

There is no closed-form solution for the above integrals, and numerical integration is unfeasible when the number of majors is large.[9] Therefore, I will approximate the above probability through simulations and maximize the simulated log-likelihood function.

The estimation process is as follows. For a given value of the parameters, a realization of μ_i is drawn from G for each individual. Using these draws, I calculate P_{ij} and p_{ij}. The process is repeated for R draws and yields the following approximate probabilities:

$$\begin{aligned} \breve{LE}_{ij} &= \frac{1}{R}\sum_{r=1}^{R}\breve{P}_{ij}^{r}\,\breve{p}_{ij}^{r} \\ \breve{LN}_{ij} &= \frac{1}{R}\sum_{r=1}^{R}\breve{P}_{ij}^{r}\,(1-\breve{p}_{ij}^{r}), \end{aligned}$$

where $\breve{P}_{ij}^{r}$ and $\breve{p}_{ij}^{r}$ are the simulated probabilities of choosing and entering major j corresponding to draw r.

[9] In this paper, for example, I will group the majors in 9 major concentrations, which would require solving 9 integrals for each of the above probabilities.

Using these simulated probabilities we obtain the following simulated log-likelihood function:

$$\breve{L} = \sum_{i=1}^{N} \sum_{j=1}^{J} M_{ij} \left(E_{ij} \log(\breve{LE}_{ij}) + (1 - E_{ij}) \log(\breve{LN}_{ij}) \right).$$

The Maximum Simulated Likelihood (MSL) estimator simply maximizes the above simulated log-likelihood, and is obtained through an iterative maximization algorithm as the usual ML estimator. The only difference is that in each step, we use a particular draw of the random term μ_{ij} to simulate the probabilities in order to construct the objective function. With respect to the choice of R, the estimators will be asymptotically consistent if R grows at a rate greater or equal than $\sqrt{N}$. Therefore, in applications R is usually chosen to be slightly larger than $\sqrt{N}$.

2.3. **Alternative models.** In this section, I discuss alternative models which could have been used to analyze major choice.

First, as I mentioned earlier, an alternative to the multinomial logit (MNL) is the multinomial probit (MNP). As is well known, the MNP does not have the problem of the IIA property. On the other hand, the MNL has the advantage of delivering closed form solutions for probabilities, which reduces the computational burden of the estimations. This property is very important for the present paper because the number of alternatives and observations is very large. Moreover, the model with correlations (Model II) eliminates the IIA property of the standard MNL, so the main argument to use the MNP instead of the MNL loses strength.

Second, another econometric approach which could have been used is the one proposed by Mallar (1977). Mallar studied a model with a set of interrelated dichotomous (binary) relationships, where the probability that one event happens affects the probability that other events occur. The approach is to transform the model, so that each probability is a non-linear function of a linear index. Then, assuming that each linear index depends on the other linear indexes (rather than on the other probabilities), it is possible to obtain a reduced form for the linear indexes, and estimate an independent equation for each probability, in which the dichotomous variable depends only on the exogenous variables (i.e. the model is transformed so that each probability does no longer depend on the other probabilities). Mallar shows that the structural parameters can be obtained from the parameters of the reduced form.

We could interpret a polychotomous (multinomial) model as a model in which the n choices are interrelated dichotomous events. Then, we could apply Mallar's approach to our problem, by adding the respective probabilities of entry. However, the above referenced transformation is a strong assumption for a multinomial model. Mallar's approach is specially suitable for the case where *truly* binary variables depend on one another, but is less satisfying for the analysis of polychotomous choices, for which the MNL and MNP have been specially designed.

In addition to the previous reasons, an important advantage of the MNL for the analysis of major choice is that this is the model which has been used the most in the literature, so it facilitates comparison with other papers. For these reasons, I will use the econometric model proposed in the previous sections to perform the analysis.

3. The university entrance process

Before describing the data, it is important to understand the entrance process for Brazilian universities (vestibular). In this section, I describe the process for University of Campinas, but the process is similar for most public universities. I will only present a brief description of the process of choosing a major and entering the university, considering the most relevant features for the present paper. The actual entrance process is much more complex.[10] Basically, university candidates must choose their preferred majors, and only the best students within each major are called to fill the seats. Majors with more candidates per offered seat are more competitive and thus imply a lower probability of entry.

Candidates must follow the following process:

(1) Each candidate chooses 3 majors in order of preference.
(2) Candidates take the first-stage exam. The exam is the same for all students and has 2 parts: multiple choice questions and essay. The multiple choice questions may belong to Mathematics, Physics, Chemistry, Biology, History or Geography. The essay evaluates knowledge of the Portuguese language, and is graded only for students who answered correctly at least 50% of the multiple choice questions. All students are ordered within their major of first-choice. Only the top students within each major

[10]For more details, read University of Campinas' general resolutions (resoluções gerales) number 31 of August 10, 2005; number 41 of August 21, 2006; and number 30 of August 8, 2007.

can participate in the second-stage of the exam.[11] It is important to remark that in this first-stage of the exam students compete only with other students *choosing the same major* for the possibility of taking the second-stage exam.

(3) Candidates take the second-stage exam. The exam is the same for all students, and has 8 parts: Literature and Portuguese Language, Biological Sciences, Geography, History, Mathematics, Chemistry, Physics, and English language.[12] The different parts of the exam are given different weights for each major (e.g. engineering majors put more weight on mathematics, literature puts more weight on Portuguese language).

(4) Students with the highest scores are called to fill a seat. If the student is not called for her first choice, she may be called for her second or third choice, if her score is high enough. As students may decide not to enroll in a major when they are called, there may be several calls until all seats at the different majors are filled.

(5) Candidates decide whether to enroll or not. Even when a candidate is called to fill a seat, she may decide not to enroll in the university. There are many reasons why this may happen:
   (a) The candidate may be one of the so called "treineiros," students who have not finished high school, and are only training for the exam in a subsequent year.
   (b) The candidate is called to fill a seat at another university, and chooses that university over this one.
   (c) The candidate decides not to enroll for other reasons (e.g. a change in her socioeconomic situation).

In stage (4), students may be offered a place at their second or third choices in an early call, but regardless of whether they accept or reject this offer, they may still be offered a superior choice in later calls. However, students who reject an offer to fill a seat in one of their chosen majors in any given call, will not be offered lower choices in subsequent calls. For example, if in the first call a student is offered a seat at her second choice major, she may reject it and still be offered a seat at her first choice in later calls, but she will no longer be offered a seat at her third choice.

[11] The number of students going to the second stage is determined following a series of complex rules. The basic intention is to have a maximum of eight candidates per seat in the second-stage exam.

[12] Some majors require an additional aptitude test.

With respect to the enrollment decision, in the available database it is possible to identify treineiros, but it is not possible to identify students deciding not to enroll for other reasons. Treineiros will be excluded from the analysis, because they may have different motives for choosing a major than non-treineiros.[13]

As explained above, students submit an ordered list of 3 majors with their university applications. For the purposes of this paper, I will focus on the study of the determinants of the first choice. There are several reasons for this decision. First, the available database does not have information on students' second and third choices. Second, the proportion of students enrolling in their second or third choice is much smaller than the proportion of students enrolling in their first choice.[14] Table 1 shows the number of candidates, offers to fill seats (calls), and enrollments, depending on the order of the major in the list of preferred majors. The table shows that 90.95% of enrolled students enroll in the major which they selected as their first choice, and only 9.05% of enrolled students enroll in a major which was not their first choice. Moreover, the share of calls for second and third choices is 14.05%, which means that students reject second and third choices in a higher proportion than first choices. Third, the reasons for choosing the second and third alternatives may be very different from the reasons for choosing the first alternative. For example, the second choice may be a 'safe bet,' that is, a major for which the probability of entry is much higher than the first choice, and which the student selects in order to maximize the probability of entering in some major.

Finally, students may be applying to other universities in addition to University of Campinas. Then, it could be the case that these students select a major at University of Campinas as a 'safe bet,' and choose a different major in the other university. The information on applications

[13]Treineiros want to take the exam in order to gain experience in the vestibular process. Given that the exam is basically the same for all majors (all that changes from major to major is the weight given to each part of the exam), treineiros care less about which major they choose as their first choice. Therefore, many treineiros choose easier majors in order to have a higher probability of getting to the second stage of the exam. In the database, for example, it is possible to see that treineiros choose Technology majors (the least demanding group of majors) in a much higher proportion than non-treineiros.

[14]Even though the database has no information on the majors that students select as second or third alternatives, we know in which majors students are offered a seat, and also in which majors students decide to enroll (if any). Therefore, we can determine if the student receives an offer or enrolls in a major which was her first choice or not. Notice however, that we cannot determine the second or third choice of students who are not offered any seat.

to other universities is not available in the database. However, it is unlikely that students will choose a major at University of Campinas as a safe bet, given that University of Campinas is one of the most prestigious universities in Brazil, and it is generally considered to be very difficult to enter this university. For example, over 60% of the students in the sample stated that they chose this university for its reputation, or because this university is the best for the major they want to study.

[TABLE 1 ABOUT HERE.]

4. Data and descriptive statistics

The dataset is composed of major choices, entrance test outcomes and individual characteristics of applicants to the University of Campinas between 2006 and 2008. The database has 134,563 observations (not including treineiros), each corresponding to one candidate. After eliminating observations with missing values, we are left with 120,058 observations.

Table 2 shows the number of candidates for a seat at the University, the number of students called to fill a seat in their first choice major, and the number of students enrolled in their first choice major; separated by gender, where M stands for male and F for female.[15] Interestingly, we can see that the difference in the probability of being offered a seat between men and women is between 2 and 4 percentage points, depending on the year. Considering the whole sample, male students' average probability of being called to fill a seat is 12.34%, while female student's probability is 9.57%.

[TABLE 2 ABOUT HERE.]

Majors have been grouped in 9 areas. Groups have been constructed taking into account similarity of the fields of study and degree of difficulty. The degree of difficulty is determined by the test score of the last person called to fill a seat: more difficult majors will require higher grades from students to be called to fill a seat. The major composition of the different groups is the following:

[15]All tables in the paper are based on the sample used in the estimations, i.e. they do not include treineiros and students with missing information for some variable used in the estimations. Tables including treineiros or students with missing information may be obtained from the author upon request.

Technologies: Construction Technology, Environmental Sanitation Technology, Telecommunications Technology, Information Technology.

Exact Sciences: Statistics, Mathematics (teaching certificate), Computer Science, Physics (teaching certificate), Physics-Mathematics-Applied Mathematics, Computation.

Engineering and Architecture: Electrical Engineering, Civil Engineering, Chemical Engineering, Mechanical Engineering, Electrical Engineering, Computer Engineering, Control and Automation Engineering, Architecture and Urban Planning.

Natural and Earth Sciences: Geography, Geology-Geography, Chemical Technology, Chemistry, Food Engineering, Agricultural Engineering.

Arts: Music Conducting, Music Composition, Music (teaching certificate), Music Instruments, Popular Music, Dance, Visual Arts, Scenic Arts.

Social Sciences: Social Sciences, Literature, History, Economics, Social Communication (Media Studies).

Humanities: Pedagogy (teaching certificate), Chemistry-Physics (teaching certificate), Linguistics, Language Studies, Language Studies (teaching certificate), Philosophy.

Health and Biological Sciences: Pharmacy, Medicine, Biological Sciences.

Other Health and Biological Sciences: Nursing, Physical Education, Phonology, Dentistry, Biological Sciences (teaching certificate).

University of Campinas organizes majors in 4 areas, according to similarity of fields of study: (i) Exact, Technological and Earth Sciences; (ii) Humanities; (iii) Arts; and (iv) Biological and Health Sciences. The approach taken to construct major concentrations was to divide these areas in groups according to the grade of the last person called to fill a seat. Humanities, and Biological and Health Sciences were divided in two groups each. Arts was kept as one group because there was not much heterogeneity between majors in terms of minimum grades. Exact, Technological and Earth Sciences was divided in 4 groups because this was the area with the largest number of majors, and with more heterogeneity in fields of study and minimum grades. Architecture was placed in the same category as most Engineering majors because it belongs to the same faculty as Civil Engineering, and has a similar minimum grade. Food Engineering and Agricultural Engineering were placed separately from the other Engineering majors because they have much lower minimum grades. Instead, these majors were placed in Natural and Earth Sciences, which is composed of majors with similar degree of difficulty and field of study.

The major offering of University of Campinas remained unchanged from 2006 to 2008. University of Campinas gives two kinds of academic degrees. A Bachelor's degree corresponds to a BSc or BA degree in American Universities. A Teaching Certificate is an inferior degree, which usually requires a lower grade to enter and is intended for graduates who want to teach at the secondary level of education (high school). I have indicated which majors correspond to Teaching Certificates in the list of majors. All other majors correspond to Bachelor's degrees.

One difference between this paper and previous works is that I will have to include more groups in the analysis.[16] This difference arises because previous works are not concerned with the difference in the degree of difficulty, but only with the similarity of the fields of study. For example, if we only considered major similarity, Medicine and Nursing would be grouped together. However, Medicine requires a much higher grade for students to be called to fill a seat. Therefore, there may be students who choose Nursing over Medicine because they are more likely to enter this major.

Table 3 shows descriptive statistics of the different major groups. The first row shows the average minimum grade required to be called to fill a seat in a major of that group. A higher minimum grade indicates that it is in general more difficult to enter a major of that group. We can see that Health and Biological Sciences, and Engineering and Architecture are the two most difficult groups. Technologies and Exact Sciences, on the other hand, are the least difficult groups. The second row (candidates) shows the number and percentage of students who selected a major within that group as their first choice. The third row (called) shows the number and percentage of candidates who are called to fill a seat in their first choice major. Finally, the fourth row (enrolled) shows the number and percentage of students that decide to enroll in their first choice major when offered the chance.

[TABLE 3 ABOUT HERE.]

Table 3 shows there are differences in the proportion of men and women choosing the different majors, but also in the proportion of

[16] For example, Montmarquette et al. (2002) and Arcidiacono (2004) consider only 4 major groups.

men and women who are called to enter a major. For example, women represent 22.37% of candidates for a seat in Engineering and Architecture, but represent only 18.94% of students called to fill a seat. This difference in shares means that men are offered seats in a higher proportion than women.

The dependent variables used in the estimations are the choice of major (first equation), and the outcome of the entry process (second equation). As explained in the previous section, the entry variable measures whether the candidate was effectively called to fill a seat in her first choice major. Table 4 shows the list of independent variables, and indicates the group of variables that will appear in each equation. Most variables will appear in both equations, but some of them will appear only in the equation determining the choice of major or in the equation determining the probability of entry. The table also shows which categories will be the reference categories in the estimations.

[TABLE 4 ABOUT HERE.]

The main explanatory variable is gender. The estimations also include interactions between gender and several variables. Other individual characteristics are represented by race, age and work status.

Education variables are also very important. There is currently an intense debate in Brazil on whether students coming out of public schools have a lower chance of entering public universities because of the low quality of primary and secondary public education. Also, students of technical schools are considered to be better in mathematics and related subjects, which may affect their probability of choosing and entering the different majors. Finally, the estimations also include a variable which indicates if the student is already enrolled in another major in University of Campinas or another university.

The most important socioeconomic variable is income. When students are surveyed, they are only asked in which income segment their family income lies. Therefore, we do not have actual income, but a categorical variable indicating the income of the family relative to the minimum wage. Allegedly, higher income should imply a higher probability of entering any major, but it is more difficult to conjecture what should be the effect on major choice. Also, poorer students are exempt from paying the registration fee, so this can also be used as an indicative of family wealth. Other important variables are the ones describing the occupation and education of father and mother.

It is always desirable to have exclusion variables to help with identification. Some variables only affect the probability of entry. For example, students with a high ENEM grade may add points to the vestibular score.[17] Also, if the student took a pre-vestibular course, she is likely to perform better in the vestibular exam. Finally, the entry equation also includes an interaction between ENEM and gender, to test whether higher ENEM grades have a differential impact on men and women. All these variables will influence the probability of entering the major the student chooses, but will not affect preferences.

Likewise, variables describing the reasons why the student chose the major and the University of Campinas will affect the equation determining major choice, but will not affect the probability of entering a given major. The choice equation also includes interactions between gender and work, secondary school variables, other major, registration fee and the variables showing the reasons for choosing the major and the university.[18]

Table 5 shows summary statistics for the independent variables. Columns 1 and 2 show the average value of each variable for men and women, and column 3 shows the sample average. For categorical variables, the average is the proportion of individuals for whom the variable is equal to 1. Interestingly, we see important differences between men and women with respect to their individual characteristics, socioeconomic variables and education variables. This shows why it is important to control for all these characteristics in the regressions, when trying to elucidate the effect of gender on preferences and probabilities of entry.

[TABLE 5 ABOUT HERE.]

[17] ENEM is a voluntary exam which students can take after finishing secondary school. Students who have taken the ENEM exam increase their final vestibular exam score, only if the ENEM grade is higher than the unadjusted vestibular score. If the student did not take the ENEM test, or if the ENEM grade is below the unadjusted vestibular score, then the vestibular score is not changed. Therefore, taking the ENEM exam may be beneficial (if the student gets a high score), but is never harmful for students.

[18] I have also estimated a model with more interactions, and found interactions to be generally non significant in the entry equation (besides the interaction between gender and ENEM). In contrast, many interactions were significant in the choice equation. The unrestricted model had too many coefficients (each interaction increases the number of coefficients to be estimated by 9), which reduced the global significance of the model. For the purposes of this paper, therefore, I will only include in the estimations the interactions which were significant in the unrestricted model.

5. Results of the estimations

In this section, I discuss the estimation results for the models presented in Section 2. I start with the benchmark model (Model I), and then proceed to analyze the model with correlations (Model II).

5.1. **Benchmark model.** As explained in Section 2, the benchmark model is estimated in two stages. First, I estimate the parameters of the entry equation, and with the resulting parameters, I calculate the probabilities of entering the different majors for each individual. Second, I estimate a multinomial logit model for the choice of major, using entrance probabilities estimated in the first stage.

Table 6 shows the estimated coefficients for the entry equation (equation 3). Table A1 shows the corresponding average marginal effects (in percentage terms), which are calculated as the average of the marginal effects of all individuals. For Gender and ENEM, the average marginal effect includes the effect of the interaction, so the sign of the marginal effects may differ from the sign of the coefficients. For all other variables, the sign of the coefficients will always coincide with the sign of the marginal effects.

[TABLE 6 ABOUT HERE.]

[TABLE A1 ABOUT HERE.]

With respect to the effect of gender on the probabilities of entry, a positive (negative) sign would indicate that males (females) have a greater probability of entering a given major group. The coefficient of gender is positive and significant for 6 groups, and it is non significant for 3 groups. Given that the estimations include an interaction between Gender and ENEM, this result only means that men have a higher probability for entering 6 major groups considering students with ENEM equal to 0. To perform a complete analysis of the effects of Gender on entrance probabilities, we have to study the sign and significance of the coefficients of ENEM and the interactions between ENEM and Gender.

The coefficient of ENEM is positive and significant for all groups. The coefficient for the interaction is negative and significant for 7 major groups, but is always smaller in absolute value than the ENEM coefficient. Therefore, a higher ENEM grade implies a higher average probability of entering all majors for both men and women, but for

7 groups of majors, a higher ENEM grade has a greater impact on women's probability of entry.

The first row of Table A1 shows the average marginal effects of Gender considering all students. We can see that the marginal effect is negative for 6 major groups, and is non significant for 3 groups. For example, holding other personal characteristics constant, women have on average a 7.80 percentage points higher probability of entering a major in Other Health and Biological Sciences, and a 4.78 percentage points higher probability of entering a Technologies major.

However, given that the interaction between ENEM and Gender is significant, the average marginal effect of gender on entrance probabilities depends on the ENEM grade. Table 8 shows the marginal effect of gender on entrance probabilities for groups of students with different ENEM. Given that for most majors ENEM coefficients for females are larger than for males, for a large enough ENEM this will affect the sign of the average marginal effect of gender conditional on ENEM. In the case of Engineering and Architecture, for example, the average marginal effect is positive for low ENEM grades, is non significant for intermediate ENEM grades, and is negative for high ENEM grades. In the case of Health and Biological Sciences, on the other hand, the effect is non significant for most ENEM groups, and is negative and significant only for the group of students with highest ENEM grade.

[TABLE 8 ABOUT HERE.]

Going back to Table A1, White is significant only for 2 major groups, Arts and Health and Biological Sciences, and in both cases, the coefficient and marginal effect is positive. The fact that White is non significant for many groups may in part be due to the PAAIS affirmative action program, which gives additional points to black and aboriginal students, and may be thus counteracting any advantage White students may have. Work is significant and negative for 4 major groups (working implies a lower probability of entering university), and significant and positive for 1 major group.

The coefficients of the age variables are significant and negative for 7 major groups, and are positive for only one group, Health and Biological Sciences. For most major groups, the marginal effects decrease in absolute value as age increases. This result is surprising, because it means that for most majors, getting older has a positive effect on the probability of entry, which may be due to two effects. First, many students try to enter the university several years before succeeding. These students may have a higher chance of entering as time goes by because they become more experienced. Second, there may be a selection process, through which older students who are still trying to enter the university are the most constant and hard-working students. In the case of Health and Biological Sciences, on the other hand, the coefficients of the age categories are positive and marginal effects decrease with age, which means that younger students have an advantage on average to enter this major group.

The coefficient of Primary School Private is significant and positive for 5 major groups. The largest marginal effects are those corresponding to Humanities (4.90 percentage points) and Exact Sciences (4.72 percentage points). The coefficient of Secondary School Private is significant and negative for 4 major groups, and significant and positive for 1 major group. The coefficient of Secondary School Mixed (students who attended both public and private schools) is significant and negative for 4 major groups. For Engineering and Architecture; Natural and Earth Sciences; and Social Sciences, coming from a private secondary school implies a decrease of 4 to 5 percentage points in the probability of entry. This surprising finding may in part be due to the PAAIS affirmative action program. Students who only attended public schools for their secondary education receive extra points in the vestibular exam, which may more than compensate the positive effects that attending private secondary schools could have on the probability of entering university.

The effect of attending a technical school is positive for 5 major groups, and the largest effect is on Humanities (7.8 percentage points). Surprisingly, the effect on Engineering and Architecture is negative. Nevertheless, it is important to remark that the sign and significance of the effects changes in the estimations corresponding to Model II. For example, in Model II, the effect on Engineering and Architecture will become non significant.

Students who are already enrolled in another major, at University of Campinas or another university, have in general a higher probability of accessing another major. For example, being enrolled in another major increases the probability of entering Humanities in 16.40 percentage points, and increases the probability of entering Exact Sciences in 15.26 percentage points. This result shows that having some experience in higher education has a positive impact on the possibilities of entering another major.

Being exempt of the registration fee indicates that the student comes from a poorer socioeconomic background. For almost all major groups, Registration Fee has a negative effect on entrance probabilities. Therefore students coming from a poorer economic background have in general a lower probability of entering university. The exception is Health and Biological Sciences, for which being exempt of the registration fee has no impact on entrance probabilities.

Finally, preparing for the vestibular exam in a private academy increases the probability of entering 7 majors. Interestingly, the preparation course has a greater effect on majors with a lower minimum grade, like Technologies, Exact Sciences, and Other Health and Biological Sciences. For example, preparing for the vestibular exam increases the probability of entering a major in Other Health and Biological Sciences in 4.28 percentage points, and increases the probability of entering Technologies in 3.95 percentage points.

Next, I discuss the estimation results for the parameters of the choice equations (equation 1). Table 9 shows the estimated coefficients and Table A2 shows the corresponding average marginal effects. Marginal effects are calculated as the average of the individual marginal effects, and are shown in percentage terms. As with any polychotomous choice model, the sign of the coefficients may not coincide with the sign of the marginal effects because we have to consider the effect of a variable on the utility of one alternative, in comparison with the effect on the utility of the other alternatives. Moreover, the variables which are included in both equations have a double effect on the probability of choosing a major: on one hand, they affect the utility of entering the different majors, but at the same time they affect the probability of entering the different majors, which also affects expected utility. For these reasons, it is more useful to perform the analysis in terms of marginal effects.

[TABLE 9 ABOUT HERE.]

[TABLE A2 ABOUT HERE.]

After controlling for other individual characteristics, and taking into account the effect of the interactions, males have on average a greater probability of choosing mathematically-oriented majors, like Technologies, Exact Sciences, and Engineering and Architecture. Men also have a greater probability of choosing Social Sciences. Women, on the other hand, have a greater probability of choosing Health and Biological Sciences; Other Health and Biological Sciences; Natural and Earth Sciences; and Arts. These findings are consistent with those of the previous literature, and show that the Brazilian case exhibits similar patterns to those found in other countries.

With respect to the magnitude of the effects, the largest effects of gender are on Engineering and Architecture, and Health and Biological Sciences. In particular, men have a 13.86 percentage points higher probability of choosing Engineering and Architecture, and women have a 10.38 percentage points higher probability of choosing Health and Biological Sciences.

The marginal effects of White and Work are generally significant. The largest effect of White is on the probability of choosing a major in Engineering and Architecture: white students have a 1.42 percentage points lower probability of choosing a major in this group. With respect to Work, students who work have a 11.39 percentage points lower probability of choosing Health and Biological Sciences, and a 16.13 percentage points higher probability of choosing Engineering and Architecture.

Students who went to a technical secondary school have a 21.12 percentage points higher probability of choosing Engineering and Architecture, and have a lower probability of choosing all other major groups. Being enrolled in another major is also generally significant. Students who are enrolled in another major, in the same or another university, have a 14.59 percentage points higher probability of choosing Engineering and Architecture, and a 5.41 percentage points lower probability of choosing Health and Biological Sciences.

Students who are exempt from paying the registration fee are more likely to choose Engineering and Architecture, Humanities, and Other Health and Biological Sciences, and are less likely to choose Technologies, Natural and Earth Sciences, Arts, Social Sciences and Health and Biological Sciences.

The coefficients of the variables indicating the reasons for choosing the major and the university are also significant. It is interesting to comment the results for some major groups. For example, choosing a major for job market reasons implies a decrease of 3.32 percentage points in the probability of choosing an Arts major and a decrease of 4.02 percentage points in the probability of choosing Health and Biological Sciences. Likewise, choosing a major for its social contribution implies an increase of 13.84 percentage points in the probability of choosing Health and Biological Sciences, and a decrease of 12.07 percentage points in the probability of choosing Engineering and Architecture.

5.2. **Model with correlations.** As explained in Section 2.2, Maximum Simulated Likelihood Estimators (MSLE) will be consistent if R grows at a rate larger or equal to $\sqrt{N}$, where N is the number of individuals. As a consequence, the number of computations increases at a rate of $N^{3/2}$, which makes it difficult to use a large sample. Therefore, for the model with correlations I will use the sample corresponding to the year 2008.[19] This sample has 39,494 observations. I will use $R = 200$ draws for each individual, which is larger than $\sqrt{N}$.[20]

Table 11 shows the coefficients of the entry equations for Model II, and Table 12 shows the corresponding average simulated marginal effects. Reported marginal effects are the average of the marginal effects calculated for each individual and each draw, and are shown in percentage terms.

[TABLE 11 ABOUT HERE.]

[TABLE 12 ABOUT HERE.]

The coefficient of Gender is positive for 4 major groups, negative for 4 major groups, and non significant for 1 major group. The coefficient of the interaction between gender and ENEM is positive for 3 major groups, negative for 4 groups and non significant for 2 groups. As in Model I, whenever the coefficient of the interaction is negative, it is smaller in absolute value than the coefficient of ENEM, which means that a higher ENEM grade implies a higher average probability of entry for both men and women. Unlike Model I, however, the coefficient of the interaction is positive for some major groups, which means that

[19]In order to determine the effects of using a smaller sample, in Appendix A, I include the estimation results for Model I using the sample corresponding to year 2008. Comparing the estimations of Model I with the full sample and the reduced sample, we can see that the differences in the effects of gender on choices are minimal. In the entry equation, there is a loss of significance of the gender effect for two major groups, but the sign and significance of the other groups remains unchanged.

[20]It may be argued that part of the differences in the results of Model I and II is caused by the difference in the estimation procedures. Specifically, Model II is estimated by MSL which involves simulating the expected value of the probabilities in equations 7 and 8. However, if the correlation between the estimations turns out to be non significant, then the expected value of these probabilities will be equal to the probabilities given by equations 4 and 5. Given the large number of draws used in the simulations, the simulated expected probabilities will be a good approximation of the true expected values. Therefore, the differences in the estimations will not be caused by the difference between ML and MSL.

for some majors, a higher ENEM grade has a larger impact on men in comparison with women.

As in the previous model, the presence of an interaction between Gender and ENEM implies that the average marginal effect of gender will vary depending on the ENEM grade of the group of students considered. Table 13 shows the marginal effect of gender for groups of students with different ENEM. Consider first the two major groups with highest average minimum grades. For Engineering and Architecture, the average effect considering all students is positive, and is also positive for most ENEM groups. However, the average marginal effect becomes negative for the top ENEM category, which is due to the fact that the coefficient of the interaction between ENEM and Gender is negative for this group. Health and Biological Sciences shows a very different pattern. The average effect considering all students is negative, and is also negative for most ENEM groups (all groups but the top group, where it becomes non significant).

[TABLE 13 ABOUT HERE.]

If we now consider the major group with the lowest average minimum grade (Technologies) we can see that the marginal effect is negative for low ENEM grades and is non significant for higher ENEM grades. On the other hand, for Exact Sciences the average effect considering all students is non significant, and the effect is negative for students with ENEM equal to zero, and positive for the top ENEM groups. Similar analyses can be performed for other major groups, which lead to the conclusion that the relation between the average marginal effect of gender and the ENEM grade depends on the particular major group under analysis.

Going back to Table 12, White is significant for 4 major groups. The largest effect is on the probability of entering Exact Sciences (2.27 percentage points), Other Health and Biological Sciences (2.20 percentage points), and Arts (2.18 percentage points). Interestingly, Work is significant only for two major groups, Engineering and Architecture, and Health and Biological Sciences, and is negative in both cases. Therefore, students who work have a lower chance of entering the most demanding majors.

With respect to the coefficients of the age variables, there are 4 major groups which exhibit a similar pattern. For Natural and Earth Sciences, Social Sciences, Humanities, and Other Health and Biological Sciences, the marginal effects are negative and decreasing in absolute value as age increases. This means that older students have an advantage to enter a major in these groups. The possible reasons behind this result have already been discussed in the previous section.

Primary education variables lose significance in comparison with Model I. Primary School Private only has a positive effect for Humanities, and Primary School Mixed has a positive effect on Engineering and Architecture; and Health and Biological Sciences.

Secondary School Private has a negative effect on the probability of entering the two most difficult majors (Engineering and Architecture; and Health and Biological Sciences). For these majors, the positive effect that private secondary schools may have is completely overcome by the extra points students get with the PAAIS program. Finally, the effect of Secondary School Mixed is negative for 2 major groups, Social Sciences and Health and Biological Sciences.

The effect of Secondary School Technical is positive and significant for Natural and Earth Sciences, and is negative for Arts and Health and Biological Sciences. This result is surprising because it means that attending a technical secondary school does not imply a higher probability of entry into technical and math-oriented majors.

The effect of being enrolled in another major is positive for all major groups except Arts, which means that having some experience in University increases the chances of entering a second career for most majors. Interestingly, the effect is larger for easier majors, and is smaller for the two most difficult majors (Engineering and Architecture, and Health and Biological Sciences).

The effect of Registration Fee is negative for all major groups except Arts. This result means that poorer students have on average a lower probability of entering most majors. As in Model I, Preparation Course is positive for 7 major groups. Finally, ENEM maintains sign and significance, which means that having a higher ENEM grade increases the probability of entering all majors.

Next, I analyze the estimation results for the parameters of the choice equations. Table 14 shows the estimated coefficients and Table 15 shows the corresponding simulated average marginal effects. In addition to the variables shown in the previous section, Table 14 shows the estimated coefficient and standard deviation of σ. The marginal effects are the average of the simulated effects corresponding to 200 draws for each individual, and show the average effects on the probability of choosing a major group in percentage terms.

[TABLE 14 ABOUT HERE.]

[TABLE 15 ABOUT HERE.]

Before analyzing the sign and significance of the effects, it is important to comment on the coefficient σ, which is significant. This means that the errors of the choice and entry equations are correlated: students who get a larger preference shock for some major tend to have a higher entry shock for that major as well. Econometric models that do not take this correlation into account will produce biased estimators, and therefore it is important to consider correlated errors in the econometric design.

In comparison with Model I, the sign of the average marginal effects of Gender on choice probabilities remains unchanged for all groups except for Social Sciences, for which the marginal effect is now non significant. The magnitudes of the effects are also similar to the previous case, except in the cases of Engineering and Health and Biological Sciences, for which they increase in absolute value. According to Model II, men have on average a 24.14 percentage points higher probability of choosing Engineering, controlling for other individual characteristics. Likewise, women have on average a 16.95 percentage points higher probability of choosing Health and Biological Sciences.

White has an effect on 4 major groups, and the largest effect is on Engineering and Architecture. Working decreases the probability of choosing Health and Biological Sciences in 12.05 percentage points, and increases the probability of choosing all other majors except Natural and Earth Sciences and Other Health and Biological Sciences, for which the effect is non significant.

Attending a technical secondary school implies a higher probability of choosing math-related majors (Technologies, Exact Sciences, and Engineering and Architecture), and Natural and Earth Sciences, and implies a lower probability of choosing Arts, Social Sciences, Health and Biological Sciences and Other Health and Biological Sciences. The largest effects are on Engineering and Architecture (7.78 percentage points), and Health and Biological Sciences (-7.77 percentage points). According to Model II, then, attending a technical secondary school does not affect the probability of entering mathematically oriented majors, but it does affect the probability of choosing these majors.

Being enrolled in another major also affects major choice. The two largest effects are on Health and Biological Sciences, and Social Sciences. Being enrolled in another major decreases the probability of choosing Health and Biological Sciences in 7.81 percentage points, and increases the probability of choosing Social Sciences in 4.26 percentage points.

As in the previous model, students who are exempt from paying the registration fee are less likely to choose Engineering and Architecture, and Health and Biological Sciences, and more likely to choose Exact Sciences and Humanities. Therefore, the conclusions of the analysis of Model I still hold: poorer students tend to avoid choosing harder majors. In particular, being exempt from the registration fee implies a decrease of 15.54 percentage points in the probability of choosing Health and Biological Sciences, and a decrease of 10.98 percentage points in the probability of choosing Engineering and Architecture, which are the two most demanding major groups, according to Table 3.

Finally, the analysis of the reasons for choosing major and university lead to similar conclusions as before. Choosing a major for job market reasons implies a decrease of 3.54 percentage points in the probability of choosing Arts, a decrease of 7.39 percentage points in the probability of choosing Health and Biological Sciences, and an increase of 4.11 percentage points in the probability of choosing Engineering and Architecture. Likewise, choosing a major for its social contribution implies an increase of 21.26 percentage points in the probability of choosing Health and Biological Sciences, and a decrease of 18.24 percentage points in the probability of choosing an Engineering major.

5.3. **Gender differences in entrance probabilities and preferences.** According to the model presented in Section 2, students choose majors by comparing expected utilities ($p_{ij}\, u_{ij}$). As a consequence, gender affects major choice in two ways: (i) through its effect on entrance probabilities (p_{ij}), and (ii) through its effect on preferences (u_{ij}).

The marginal effects presented in the previous section were constructed taking into account both effects. In this section, I try to separate the two effects, to see what part of the difference in gender choices are generated by differences in the probability of entering the different majors, and what part is generated by differences in preferences.

Specifically, I perform two simulations. First, I simulate women's choices using male entrance probabilities (i.e. setting Gender equal to 1 in the entry equation, and equal to 0 in the choice equation), and compare them with men's choices (setting Gender equal to 1 in both equations). Then, I simulate men's choices using female entrance probabilities, and compare them with women's choices. Table 16 presents the results of the simulations for Model II, as well as choice probabilities calculated with own-gender entrance probabilities.

[TABLE 16 ABOUT HERE.]

As expected, changing the entrance probabilities used to calculate expected utility has an effect on the probabilities of choosing the different majors. For example, using female entrance probabilities, women's average probability of choosing Engineering and Architecture is 12.72%, but when we simulate women's choices using male entrance probabilities, this probability increases to 21.11%. Likewise, using male entrance probabilities, men's average probability of choosing Health and Biological Sciences is 22.41%, but when we simulate men's choices using female entrance probabilities, this probability increases to 30.07%. Therefore, it is clear that gender differences in entrance probabilities affect major choice. In particular, men have on average a higher probability of entering Engineering and Architecture, and a lower probability of entering Health and Biological Sciences (see Table 12). Therefore, men will choose Engineering and Architecture majors in a higher proportion than they would choose them if this difference in entrance probabilities did not exist. Likewise, women will choose Health and Biological Sciences majors in a higher proportion than they would choose them if this difference in entrance probabilities did not exist.

Table 16 also shows that the sign of the gender differences in choice probabilities is the same for both simulations, except in the case of Social Sciences. Moreover, the magnitudes of the differences are similar. For example, in the case of Engineering and Architecture, simulated gender differences in choice probabilities are 16.24 or 17.79 percentage points, depending on the simulation.

For most majors, simulated gender differences in choice probabilities are very similar to gender differences calculated using own-gender probabilities. Therefore, for these majors, differences in preferences explain most of the gender difference in major choice. Nevertheless, there are two important exceptions, which are precisely the two most difficult majors. Using own-gender entrance probabilities, men have on average a 26.28 percentage points higher probability of choosing Engineering and Architecture, but when we simulate choices controlling for gender differences in entrance probabilities, men have a 16.24 or 17.79 percentage points higher probability. Therefore, for Engineering and Architecture, there is a substantial part of the difference in choices which is explained by gender differences in the probability of entry. The same can be said about Health and Biological Sciences.

5.4. **Interaction between gender and other explanatory variables.** Given the presence of interactions between gender and other variables, the marginal effect of gender may differ for groups of students with different characteristics. For example, gender differences in choice probabilities may be smaller or greater for students who attended public secondary schools, in comparison with students who attended private secondary schools. Table 17 shows simulated average marginal effects for different groups of students for Model II.

[TABLE 17 ABOUT HERE.]

Table 17 shows there are significant differences in the effect of gender depending on the group of students under analysis. For example, comparing the effect of gender for working and non-working students, we can see that in the cases of Engineering and Architecture, Natural and Earth Sciences, Humanities, and Health and Biological Sciences, the difference in choice probabilities between men and women is larger (in absolute value) for students who work than for students who do not work. Nevertheless, this pattern is not uniform across majors. In the cases of Technologies, Exact Sciences, Arts, Social Sciences, and Other Health and Biological Sciences, the gender difference in choice probabilities is smaller for students who work, in comparison with students who do not work.

Similar analyses can be performed for other variables. In particular, it is interesting to examine the effects of secondary education on gender differences. For example, men have a 25.56 percentage points higher probability of choosing Engineering and Architecture if we consider the group of students who attended private secondary schools, but the difference reduces to 21.34 percentage points for students who attended public secondary schools. Likewise, men have a 19.82 percentage points lower probability of choosing Health and Biological Sciences in the group of students who attended private secondary schools, but the difference reduces to 11.13 percentage points for students who attended public secondary schools. Therefore, private secondary education leads to higher differences in the choices of men and women in the two most demanding major groups. Nevertheless, there are also major groups for which gender differences are larger in the group of students who attended public secondary schools (e.g. Other Health and Biological Sciences).

6. Conclusion

Gender differences in major choice have triggered an extensive literature trying to decipher the reasons for the existence of such differences. The Brazilian case is interesting, because in most public universities students choose a major before taking a major-specific exam which determines whether they can enter the major of their choice. This contrasts with the college entrance process in most other countries (including the US), where students are first allowed entry into university and then have to choose their preferred major. The singular characteristics of the Brazilian case allow us to test whether differences in choices are due to differences in the probabilities of entry or differences in the utility associated with the different majors.

I have presented two econometric models, and estimated them using data from the University of Campinas, a prestigious public university dependent of the State of São Paulo. The first model imposes independence between preference and entry shocks, but can be estimated with standard econometric software. The second model relaxes the independence assumption, but becomes harder to estimate, and I have to resort to a Maximum Simulated Likelihood approach.

After estimating the second model, I find that the correlation between the errors of the two equations is positive and significant: students who get a larger preference shock for some major tend to have a higher entry shock for that major as well. The significance of this coefficient means that the model without correlations will produce biased estimators. Therefore, it is important to consider correlated errors in the econometric design.

With respect to the effect of gender on entrance probabilities, there are several interesting findings. First, the average gender effect on entrance probabilities is positive for some majors and negative for other majors. Second, the effect of gender on entrance probabilities depends on the ENEM grade. Nevertheless, it is difficult to generalize on the nature of the relation between gender effects and ENEM, as it will generally depend on the specific major group under consideration.

In addition to gender, entrance probabilities are affected by other variables. ENEM has a positive effect on entrance probabilities for both men and women. Students who are already enrolled in another major have a higher probability of entering all major groups except Arts, for which the effect is non significant. Students who are exempted of the registration fee (which indicates that the student comes from a poorer family) have a lower probability of entering all majors except Arts, for which the effect is non significant.

An important issue being discussed in Brazil is what is the effect of private vs. public education on the possibilities of accessing higher education. It is generally argued that students of private schools receive a better education, which gives them an advantage for entering college. I find that the effects of attending a primary private school are generally non significant. I also find that students who attended private secondary schools have a lower probability of entering the most demanding majors (Engineering and Architecture, and Health and Biological Sciences), and have a higher probability of entering Exact Sciences and Other Health and Biological Sciences. In the case of Engineering and Architecture, and Health and Biological Sciences, the negative sign of the coefficient may be partly due to the PAAIS affirmative action program, which gives additional points to students who attended only public secondary schools.

With respect to the effects of gender on major choice, I find that men have on average a higher probability of choosing mathematically oriented majors (Technologies, Exact Sciences and Engineering and Architecture), and women have on average a higher probability of choosing Natural and Earth Sciences, Arts, Humanities, Health and Biological Sciences and Other Health and Biological Sciences. The average effect of gender on the probability of choosing Social Sciences is non significant.

In order to determine if gender differences in major choice are caused by differences in preferences or probability of entry, I simulate women choices with male probabilities of entry, and men choices with female probabilities of entry. I find that preferences account for most of the difference in choices in majors with low or medium minimum required grades. In the most demanding majors (Engineering and Architecture, and Health and Biological Sciences), on the other hand, a large part of the difference in major choice is explained by differences in the probability of entry.

Finally, I find that the effect of gender on major choice depends on education, socioeconomic variables and family background. For example, for Engineering and Architecture, and Health and Biological Sciences, the difference between men and women is larger among students who attended private schools, in comparison with students who attended public schools. Therefore, for these two major groups, private secondary education leads to larger differences between men and women. Nevertheless, there are also major groups for which gender differences are larger in the group of students who attended public secondary schools (e.g. Other Health and Biological Sciences).

References

Altonji, J. (1993): "The Demand for and Return to Education When Education Outcomes are Uncertain," *Journal of Labor Economics*, 11(1), 48.

Anderson, S., A. De Palma, and J. Thisse (1992): *Discrete Choice Theory of Product Differentiation.* MIT Press.

Arcidiacono, P. (2004): "Ability sorting and the returns to college major," *Journal of Econometrics*, 121(1-2), 343–375.

Ben-Akiva, M., and S. Lerman (1985): *Discrete choice analysis: theory and application to travel demand*. MIT press.

Cavalcanti, T., J. Guimarães, and B. Sampaio (2009): "Barriers to Skill Acquisition in Brazil: Evidences from a University Entrance Exam," Unpublished Manuscript.

Freeman, R. (1971): *The market for college-trained manpower: A study in the economics of career choice*. Harvard University Press.

Guimarães, J., and B. Sampaio (2007): "The Influence Of Family Background And Individual Characteristics On Entrance Tests Scores Of Brazilian University Students," Anais do XXXV Encontro Nacional de Economia 092, ANPEC - Associação Nacional dos Centros de Pósgraduação em Economia.

——— (2008): "Mind the Gap: Evidences from Gender Differences in Scores in Brazil," Anais do XXXVI Encontro Nacional de Economia 200807211527140, ANPEC - Associação Nacional dos Centros de Pósgraduação em Economia.

Mallar, C. (1977): "The estimation of simultaneous probability models," *Econometrica*, 45(7), 1717–1722.

McFadden, D. (1974): "Conditional logit analysis of qualitative choice behavior," in *Frontiers in econometrics*, ed. by P. Zarembka, pp. 105–142. Academic Press, New York.

McFadden, D., and K. Train (2000): "Mixed MNL models for discrete response," *Journal of Applied Econometrics*, 5, 447–470.

Montmarquette, C., K. Cannings, and S. Mahseredjian (2002): "How do young people choose college majors?," *Economics of Education Review*, 21(6), 543–556.

Pinheiro, L., N. de Oliveira Fontoura, A. C. Querino, A. Bonetti, and W. Rosa (2008): *Brasil–Retrato das Desigualdades: Gênero e Raça*. IPEA, SPM, UNIFEM, Brasilia, 3rd edn.

Train, K. (2003): *Discrete choice methods with simulation*. Cambridge University Press.

Turner, S., and W. Bowen (1999): "Choice of major: The changing (unchanging) gender gap," *Industrial and Labor Relations Review*, 52(2), 289–313.

Zafar, B. (2009): "College Major Choice and the Gender Gap," Federal Reserve Bank of New York, Staff Report No. 364.

TABLE 1. Candidates for entry between 2006 and 2008

	Candidates	Called			Enrolled		
		1st Choice	Other Choice	Total	1st Choice	Other Choice	Total
2006	40,162	4,180	720	4,900	2,389	317	2,706
%		85.31	14.69	100	88.29	11.71	100
2007	40,402	4,632	883	5,515	2,508	243	2,751
%		83.99	16.01	100	91.17	8.83	100
2008	39,494	4,354	549	4,903	2,518	178	2,696
%		88.80	11.20	100	93.40	6.60	100
Total	120,058	13,166	2,152	15,318	7,415	738	8,153
%		85.95	14.05	100	90.95	9.05	100

Table 2. Gender composition of candidates for entry between 2006 and 2008

	Candidates			Called			Enrolled		
	M	F	T	M	F	T	M	F	T
2006	20,232	19,930	40,162	2,417	1,763	4,180	1,362	1,027	2,389
%				11.95	8.85	10.41	6.73	5.15	5.95
2007	20,387	20,015	40,402	2,650	1,982	4,632	1,419	1,089	2,508
%				13.00	9.90	11.46	6.96	5.44	6.21
2008	19,652	19,842	39,494	2,373	1,981	4,354	1,381	1,137	2,518
%				12.08	9.98	11.02	7.03	5.73	6.38
Total	60,271	59,787	120,058	7,440	5,726	13,166	4,162	3,253	7,415
%				12.34	9.58	10.97	6.91	5.44	6.18

Rows marked with the percentage sign indicate the percentage of students called or enrolled in their first choice major over the total number of candidates.

Table 3. Descriptive statistics of major groups.

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.	Total
Average Min. Grade	360.87	441.90	552.80	482.51	448.19	521.65	447.76	603.08	443.00	
Candidates	2,597	7,113	28,827	9,632	3,914	14,624	5,622	37,632	10,097	120,058
Male	1,639	5,148	22,378	4,118	1,679	7,494	1,455	13,593	2,767	60,271
%	63.11	72.37	77.63	42.75	42.9	51.24	25.88	36.12	27.40	50.20
Female	958	1,965	6,449	5,514	2,235	7,130	4,167	24,039	7,330	59,787
%	36.89	27.63	22.37	57.25	57.1	48.76	74.12	63.88	72.60	49.80
Called 1st ch	908	1,388	3,206	1,422	460	1,649	925	1,634	1,574	13,166
Male	578	1,059	2,573	711	201	903	307	682	426	7,440
%	63.66	76.3	80.26	50.00	43.70	54.76	33.19	41.74	27.06	56.51
Female	330	329	633	711	259	746	618	952	1,148	5,726
%	36.34	23.7	19.74	50.00	56.30	45.24	66.81	58.26	72.94	43.49
Enrolled 1st ch	638	820	1,642	935	380	815	603	702	880	7,415
Male	417	641	1,331	431	169	441	196	267	269	4,162
%	65.36	78.17	81.06	46.1	44.47	54.11	32.5	38.03	30.57	56.13
Female	221	179	311	504	211	374	407	435	611	3,253
%	34.64	21.83	18.94	53.9	55.53	45.89	67.5	61.97	69.43	43.87

TABLE 4. Variable definitions

Variable	Definition
Personal characteristics	
Gender	1 if male, 0 if female
White	1 if white, 0 otherwise
Work	1 if currently working, 0 if not working
Age1	1 if age is 17 or less
Age2	1 if age is between 18 and 19
Age3	1 if age is between 20 and 23
Age4 (*)	1 if age is 24 or more
Education variables	
Prim Sch Private	1 if attended only private primary schools
Prim Sch Public (*)	1 if attended only public primary schools
Prim Sch Mixed	1 if attended both private and public pr. sch.
Sec Sch Private	1 if attended only private secondary schools
Sec Sch Public (*)	1 if attended only public secondary schools
Sec Sch Mixed	1 if attended both private and public sec. sch.
Sec Sch Technical	1 if attended technical secondary school
Other Major	1 if already coursing another major
Socioeconomic factors	
Reg Fee	1 if exempt from paying registration fee
Income Low (*)	1 if family income is up to 5 minimum wages
Income Medium	1 if family inc. is between 5 and 15 min. wages
Income High	1 if family income is above 15 minimum wages
Educ Father None	1 if father has some primary school or none
Educ Father Prim	1 if father finished primary school
Educ Father Low Sec	1 if father finished low secondary school
Educ Father High Sec	1 if father finished high secondary school
Educ Father Uni (*)	1 if father finished university
Educ Mother None	1 if mother has some primary school or none
Educ Mother Prim	1 if mother finished primary school
Educ Mother Low Sec	1 if mother finished low secondary school
Educ Mother High Sec	1 if mother finished high secondary school
Educ Mother Uni (*)	1 if mother finished university

TABLE 4. Variable definitions (cont.)

Variable	Definition
Socioeconomic factors (cont.)	
Prof Father Professional (*)	1 if father is professional
Prof Father Non-manual	1 if father has job with non-manual tasks
Prof Father Manual	1 if father has job with manual tasks
Prof Father Other	1 if father has another kind of job
Prof Mother Professional (*)	1 if mother is professional
Prof Mother Non-manual	1 if mother has job with non-manual tasks
Prof Mother Manual	1 if mother has job with manual tasks
Prof Mother Housewife	1 if mother is a housewife
Prof Mother Other	1 if mother has another kind of job
Others	
Year	Vestibular exam year
Only in major choice equation	
Rsn Major Ability (*)	1 if chose major because of personal ability
Rsn Major Job Market	1 if chose major because of job market prospects
Rsn Major Soc Contrib	1 if chose major to contribute to society
Rsn Major Pers Realization	1 if chose major for personal realization
Rsn Major Other	1 if chose major for other reasons
Rsn Univ Best for course (*)	1 if chose Unicamp because is best for course
Rsn Univ Free	1 if chose Unicamp because it is free
Rsn Univ Reputation	1 if chose Unicamp for its reputation
Rsn Univ Other	1 if chose Unicamp for other reasons
Only in probability of entry equation	
Prep Course	1 if took a preparation course for vestibular exam
ENEM	ENEM test grade
ENEM * Gender	ENEM interacted with Gender

(*) Reference category in the estimations.

Table 5. Summary statistics

Variable	Male	Female	Total
Personal characteristics			
White	0.751	0.769	0.760
Work	0.203	0.147	0.175
Age1	0.165	0.172	0.168
Age2	0.534	0.554	0.544
Age3	0.222	0.218	0.220
Age4	0.079	0.056	0.067
Education variables			
Prim Sch Private	0.540	0.511	0.526
Prim Sch Public	0.297	0.338	0.317
Prim Sch Mixed	0.164	0.151	0.157
Sec Sch Private	0.650	0.640	0.645
Sec Sch Public	0.294	0.309	0.301
Sec Sch Mixed	0.056	0.051	0.053
Sec Sch Technical	0.112	0.065	0.089
Other Major	0.100	0.069	0.084
Socioeconomic factors			
Reg Fee	0.068	0.127	0.098
Income Low	0.264	0.323	0.293
Income Medium	0.470	0.446	0.458
Income High	0.266	0.231	0.249
Educ Father None	0.059	0.067	0.063
Educ Father Prim	0.072	0.087	0.079
Educ Father Low Sec	0.075	0.082	0.079
Educ Father High Sec	0.285	0.300	0.293
Educ Father Uni	0.508	0.463	0.486
Educ Mother None	0.048	0.054	0.051
Educ Mother Prim	0.074	0.086	0.080
Educ Mother Low Sec	0.081	0.090	0.086
Educ Mother High Sec	0.312	0.320	0.316
Educ Mother Uni	0.484	0.450	0.467

TABLE 5. Summary statistics (cont.)

Variable	Male	Female	Total
Socioeconomic factors (cont.)			
Prof Father Professional	0.487	0.453	0.470
Prof Father Non-manual	0.274	0.281	0.278
Prof Father Manual	0.105	0.118	0.111
Prof Father Other	0.134	0.148	0.141
Prof Mother Professional	0.320	0.296	0.308
Prof Mother Non-manual	0.261	0.269	0.265
Prof Mother Manual	0.043	0.043	0.043
Prof Mother Housewife	0.272	0.290	0.281
Prof Mother Other	0.105	0.102	0.104
Only in major choice equation			
Rsn Major Ability	0.541	0.485	0.513
Rsn Major Job Market	0.090	0.060	0.075
Rsn Major Soc Contrib	0.087	0.126	0.106
Rsn Major Pers Realization	0.204	0.259	0.232
Rsn Major Other	0.078	0.069	0.074
Rsn Univ Best for course	0.418	0.373	0.396
Rsn Univ Free	0.174	0.203	0.189
Rsn Univ Reputation	0.232	0.254	0.243
Rsn Univ Other	0.176	0.169	0.172
Only in probability of entry equation			
Prep Course	0.585	0.613	0.599
ENEM	89.987	83.208	86.540

For dummy variables, the mean is equal to the proportion of individuals with that characteristic.

TABLE 6. Coefficients of entry equations (Model I)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Gender	0.561 **	1.694 ***	2.382 **	4.482 ***	0.768	0.711	1.941 ***	1.105	2.081 ***
	(0.273)	(0.495)	(0.952)	(0.773)	(0.469)	(0.761)	(0.341)	(1.348)	(0.583)
White	0.123	0.085	0.046	-0.055	0.343 **	-0.013	0.114	0.189 ***	0.084
	(0.100)	(0.076)	(0.047)	(0.073)	(0.152)	(0.068)	(0.095)	(0.070)	(0.077)
Work	-0.016	-0.300 ***	-0.351 ***	-0.221 **	0.373 ***	-0.100	-0.071	0.023	-0.356 ***
	(0.108)	(0.087)	(0.074)	(0.096)	(0.121)	(0.087)	(0.095)	(0.130)	(0.096)
Age1	-0.603 ***	-0.866 ***	-0.298 *	-0.010	-1.158 ***	-0.530 ***	-0.781 ***	0.776 ***	-0.743 ***
	(0.224)	(0.164)	(0.174)	(0.197)	(0.278)	(0.195)	(0.183)	(0.202)	(0.197)
Age2	-0.346 **	-0.799 ***	-0.273 *	-0.041	-1.035 ***	-0.420 **	-0.645 ***	0.751 ***	-0.540 ***
	(0.173)	(0.140)	(0.165)	(0.173)	(0.244)	(0.173)	(0.143)	(0.176)	(0.168)
Age3	-0.083	-0.599 ***	-0.358 **	-0.067	-0.690 ***	-0.397 **	-0.396 ***	0.486 ***	-0.304 *
	(0.164)	(0.139)	(0.165)	(0.171)	(0.241)	(0.173)	(0.141)	(0.170)	(0.164)
Prim Sch Priv	-0.053	0.296 ***	0.479 ***	0.104	0.051	0.271 ***	0.283 **	0.129	0.213 **
	(0.145)	(0.101)	(0.074)	(0.099)	(0.177)	(0.095)	(0.130)	(0.102)	(0.093)
Prim Sch Mixed	0.057	0.085	0.261 ***	0.055	0.131	0.079	0.105	-0.018	0.104
	(0.146)	(0.104)	(0.080)	(0.104)	(0.178)	(0.103)	(0.132)	(0.111)	(0.100)
Sec Sch Priv	0.162	0.266 ***	-0.535 ***	-0.402 ***	-0.008	-0.499 ***	0.152	-0.646 ***	0.096
	(0.136)	(0.092)	(0.062)	(0.092)	(0.163)	(0.086)	(0.125)	(0.096)	(0.093)
Sec Sch Mixed	-0.160	0.118	-0.535 ***	-0.378 **	-0.308	-0.773 ***	0.012	-0.980 ***	-0.094
	(0.189)	(0.146)	(0.119)	(0.158)	(0.257)	(0.155)	(0.176)	(0.190)	(0.147)
Sec Sch Tech	0.284 **	-0.075	-0.124 *	0.371 ***	-0.005	0.182 *	0.425 ***	0.414 ***	0.145
	(0.122)	(0.093)	(0.067)	(0.091)	(0.213)	(0.101)	(0.131)	(0.130)	(0.113)
Other Major	0.566 ***	0.844 ***	0.381 ***	0.827 ***	0.477 ***	0.717 ***	0.854 ***	0.110	0.614 ***
	(0.160)	(0.104)	(0.083)	(0.121)	(0.162)	(0.093)	(0.124)	(0.107)	(0.141)
Reg Fee	-0.845 ***	-0.674 ***	-0.268 *	-0.511 ***	-0.575 *	-0.913 ***	-0.771 ***	-0.214	-0.556 ***
	(0.150)	(0.138)	(0.158)	(0.154)	(0.300)	(0.187)	(0.124)	(0.244)	(0.120)

TABLE 6. Coefficients of entry equations (Model I, cont.)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Prep Course	0.187 *	0.194 ***	0.200 ***	0.242 ***	0.227 *	0.144 **	0.122	0.012	0.294 ***
	(0.097)	(0.070)	(0.044)	(0.072)	(0.119)	(0.065)	(0.089)	(0.076)	(0.074)
ENEM	0.024 ***	0.038 ***	0.084 ***	0.088 ***	0.034 ***	0.057 ***	0.039 ***	0.133 ***	0.069 ***
	(0.003)	(0.006)	(0.009)	(0.007)	(0.005)	(0.007)	(0.004)	(0.008)	(0.005)
ENEM * Gender	-0.010 **	-0.020 ***	-0.025 **	-0.047 ***	-0.011 *	-0.009	-0.023 ***	-0.013	-0.028 ***
	(0.004)	(0.006)	(0.010)	(0.008)	(0.006)	(0.008)	(0.004)	(0.013)	(0.007)
Constant	-1.686 ***	-4.384 ***	-10.316 ***	-9.908 ***	-4.260 ***	-7.189 ***	-4.470 ***	-17.482 ***	-7.208 ***
	(0.389)	(0.530)	(0.944)	(0.770)	(0.540)	(0.686)	(0.405)	(0.897)	(0.508)
Number of obs.	120,058								

Estimations include the explanatory variables listed in Table 4. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 7. Marginal effects on the probability of entry (Model I)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Gender	-4.78 **	-2.04	-0.75 *	0.53	-2.75 *	-1.75 ***	-1.81	-0.71 ***	-7.80 ***
	(2.20)	(1.29)	(0.40)	(0.84)	(1.41)	(0.53)	(1.80)	(0.18)	(1.20)
White	2.58	1.36	0.36	-0.67	3.73 **	-0.12	1.97	0.49 ***	1.22
	(2.09)	(1.20)	(0.36)	(0.90)	(1.54)	(0.61)	(1.62)	(0.17)	(1.11)
Work	-0.33	-4.63 ***	-2.53 ***	-2.59 **	4.61 ***	-0.87	-1.22	0.06	-5.00 ***
	(2.27)	(1.28)	(0.49)	(1.08)	(1.56)	(0.74)	(1.63)	(0.36)	(1.28)
Age1	-12.73 ***	-15.30 ***	-2.54	-0.12	-15.98 ***	-5.20 **	-14.39 ***	1.71 ***	-11.35 ***
	(4.68)	(3.00)	(1.59)	(2.43)	(4.27)	(2.11)	(3.36)	(0.40)	(3.13)
Age2	-7.37 **	-14.25 ***	-2.34	-0.50	-14.70 ***	-4.25 **	-12.10 ***	1.64 ***	-8.49 ***
	(3.66)	(2.67)	(1.53)	(2.13)	(4.04)	(1.95)	(2.78)	(0.30)	(2.79)
Age3	-1.78	-10.99 ***	-3.00 **	-0.82	-10.54 ***	-4.04 **	-7.65 ***	0.94 ***	-4.93 *
	(3.49)	(2.63)	(1.51)	(2.11)	(3.99)	(1.94)	(2.75)	(0.28)	(2.73)
Prim Sch Priv	-1.12	4.72 ***	3.52 ***	1.26	0.58	2.36 ***	4.90 **	0.35	3.10 **
	(3.03)	(1.59)	(0.51)	(1.18)	(2.00)	(0.80)	(2.23)	(0.27)	(1.34)
Prim Sch Mixed	1.21	1.30	1.78 ***	0.66	1.52	0.64	1.77	-0.04	1.49
	(3.08)	(1.59)	(0.54)	(1.25)	(2.05)	(0.84)	(2.22)	(0.28)	(1.43)
Sec Sch Priv	3.43	4.19 ***	-4.62 ***	-5.08 ***	-0.09	-4.89 ***	2.62	-2.08 ***	1.40
	(2.87)	(1.43)	(0.58)	(1.19)	(1.89)	(0.90)	(2.14)	(0.37)	(1.35)
Sec Sch Mixed	-3.32	1.80	-4.62 ***	-4.81 **	-3.28	-7.01 ***	0.20	-2.81 ***	-1.32
	(3.90)	(2.26)	(0.94)	(1.91)	(2.60)	(1.23)	(2.97)	(0.46)	(2.06)
Sec Sch Tech	6.00 **	-1.20	-0.94 *	4.80 ***	-0.06	1.71 *	7.80 ***	1.29 ***	2.16
	(2.58)	(1.46)	(0.50)	(1.25)	(2.44)	(0.99)	(2.51)	(0.46)	(1.71)

TABLE 7. Marginal effects on the probability of entry (Model I, cont.)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Other Major	11.95 ***	15.26 ***	3.30 ***	11.55 ***	6.13 ***	7.71 ***	16.40 ***	0.31	9.71 ***
	(3.33)	(2.04)	(0.81)	(1.92)	(2.29)	(1.19)	(2.51)	(0.31)	(2.38)
Reg Fee	-17.11 ***	-9.56 ***	-1.95 *	-5.59 ***	-5.68 **	-6.24 ***	-11.84 ***	-0.54	-7.49 ***
	(2.78)	(1.69)	(1.06)	(1.50)	(2.49)	(0.93)	(1.66)	(0.56)	(1.47)
Prep Course	3.95 *	3.11 ***	1.54 ***	2.91 ***	2.57 *	1.27 **	2.12	0.03	4.28 ***
	(2.06)	(1.12)	(0.33)	(0.85)	(1.34)	(0.57)	(1.55)	(0.20)	(1.06)
ENEM	0.40 ***	0.44 ***	0.55 ***	0.74 ***	0.33 ***	0.47 ***	0.47 ***	0.34 ***	0.79 ***
	(0.04)	(0.04)	(0.04)	(0.06)	(0.04)	(0.04)	(0.04)	(0.02)	(0.06)
Number of obs.	120,058								

* significant at 10%, ** significant at 5%, *** significant at 1%.

TABLE 8. Marginal effects on the probability of entry for different ENEM groups (Model I)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
= 0	7.67 **	6.67 ***	0.05 *	0.78 *	1.58 *	0.11	10.18 ***	0.00	0.84 *
	(3.40)	(1.19)	(0.03)	(0.41)	(0.92)	(0.14)	(1.67)	(0.00)	(0.48)
(0, 70]	-0.07	4.50 ***	0.94 **	5.50 ***	0.60	0.33	6.30 ***	0.02	2.15 *
	(1.96)	(1.38)	(0.37)	(1.04)	(0.90)	(0.70)	(1.36)	(0.03)	(1.12)
(70, 80]	-3.91 *	2.14	1.68 **	7.78 ***	-0.61	-0.05	2.91 *	0.03	-0.52
	(2.18)	(1.38)	(0.73)	(1.17)	(1.09)	(0.96)	(1.57)	(0.14)	(1.18)
(80, 90]	-6.33 **	-1.13	1.44 *	6.10 ***	-2.17	-0.93	-1.27	-0.04	-5.59 ***
	(2.63)	(1.38)	(0.84)	(1.01)	(1.37)	(0.84)	(1.93)	(0.29)	(1.13)
(90, 100]	-8.47 ***	-5.25 ***	-0.30	0.66	-4.24 **	-2.41 ***	-6.18 **	-0.49	-12.27 ***
	(3.15)	(1.95)	(0.63)	(1.21)	(1.99)	(0.72)	(2.58)	(0.41)	(1.90)
(100, 110]	-10.31 ***	-9.98 ***	-4.08 ***	-8.22 ***	-6.81 **	-4.45 ***	-11.46 ***	-1.88 ***	-19.29 ***
	(3.65)	(2.99)	(1.26)	(2.57)	(2.96)	(1.52)	(3.37)	(0.38)	(3.17)
(110, 120]	-11.53 ***	-14.36 ***	-9.11 ***	-17.78 ***	-9.46 **	-6.65 **	-16.23 ***	-4.42 ***	-24.77 ***
	(3.97)	(4.00)	(3.01)	(4.17)	(4.05)	(2.84)	(4.10)	(1.23)	(4.34)
Average	-4.78 **	-2.04	-0.75 *	0.53	-2.75 *	-1.75 ***	-1.81	-0.71 ***	-7.80 ***
	(2.20)	(1.29)	(0.40)	(0.84)	(1.41)	(0.53)	(1.80)	(0.18)	(1.20)
Number of obs.	120,058								

* significant at 10%, ** significant at 5%, *** significant at 1%.

TABLE 9. Coefficients of choice equations (Model I)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Gender	1.871 ***	6.802 ***	10.601 ***	-0.599	-2.394 ***	2.777 ***	-4.148 ***	2.325 *	-3.212 ***
	(0.329)	(0.486)	(0.657)	(0.401)	(0.800)	(0.554)	(0.451)	(1.357)	(0.409)
White	0.046	0.010	-0.801 ***	0.006	3.606 ***	0.302	0.347 **	-0.149	0.753 ***
	(0.126)	(0.157)	(0.230)	(0.172)	(0.450)	(0.222)	(0.169)	(0.594)	(0.139)
Work	1.648 ***	1.923 ***	1.844 **	2.053 ***	7.649 ***	3.770 ***	3.145 ***	0.612	1.400 ***
	(0.227)	(0.407)	(0.884)	(0.379)	(0.444)	(0.469)	(0.239)	(1.155)	(0.228)
Age1	-1.992 ***	0.503	8.319 ***	-0.239	-4.509 ***	0.213	-3.155 ***	-22.389 ***	-3.685 ***
	(0.291)	(0.321)	(0.642)	(0.402)	(0.725)	(0.527)	(0.362)	(1.896)	(0.330)
Age2	-0.487 **	0.636 **	7.524 ***	1.231 ***	-1.880 ***	3.310 ***	-1.343 ***	-17.289 ***	-0.785 ***
	(0.207)	(0.270)	(0.586)	(0.333)	(0.599)	(0.425)	(0.277)	(1.774)	(0.261)
Age3	-0.057	-0.318	3.304 ***	0.412	-0.267	1.402 ***	-0.880 ***	-12.706 ***	-0.064
	(0.189)	(0.253)	(0.572)	(0.318)	(0.522)	(0.403)	(0.245)	(1.719)	(0.250)
Prim Sch Priv	-0.751 ***	0.441 **	-1.834 ***	-0.625 ***	0.018	0.554 *	-0.153	-1.810 **	-0.787 ***
	(0.170)	(0.216)	(0.380)	(0.229)	(0.472)	(0.329)	(0.221)	(0.778)	(0.171)
Prim Sch Mixed	-0.415 **	0.313	-1.696 ***	-0.323	1.069 **	-0.094	-0.264	-0.769	-0.482 ***
	(0.168)	(0.230)	(0.403)	(0.239)	(0.477)	(0.355)	(0.230)	(0.822)	(0.182)
Sec Sch Priv	-1.312 ***	-0.783 **	-2.832 ***	-3.133 ***	-3.360 ***	-4.659 ***	-1.235 ***	-2.096 **	-0.837 ***
	(0.225)	(0.399)	(0.509)	(0.295)	(0.535)	(0.394)	(0.258)	(0.861)	(0.199)
Sec Sch Mixed	-1.009 **	0.203	-2.171 *	-1.103 *	-3.059 ***	-1.987 **	-0.343	6.056 *	-0.385
	(0.439)	(0.652)	(1.270)	(0.617)	(0.967)	(0.951)	(0.424)	(3.199)	(0.363)
Sec Sch Tech	1.623 ***	2.404 ***	3.972 ***	3.037 ***	-1.108	0.658	0.689 **	0.040	0.687 ***
	(0.257)	(0.460)	(0.737)	(0.348)	(0.820)	(0.553)	(0.316)	(1.121)	(0.262)
Other Major	0.377	4.452 ***	3.931 ***	2.532 ***	3.158 ***	5.002 ***	1.955 ***	-0.588	0.613 ***
	(0.358)	(0.345)	(0.561)	(0.325)	(0.509)	(0.378)	(0.274)	(1.244)	(0.215)
Reg Fee	-1.619 ***	1.232	-3.983 *	0.769	-9.486 ***	2.920 **	7.215 ***	9.969 ***	3.994 ***
	(0.369)	(0.863)	(2.126)	(0.797)	(2.359)	(1.411)	(0.512)	(2.047)	(0.427)
Prof Father Non-manual	0.437 **	0.659 ***	0.405	0.697 ***	0.450	0.324	0.259	-0.320	0.381 **
	(0.172)	(0.177)	(0.276)	(0.208)	(0.363)	(0.268)	(0.203)	(0.646)	(0.153)
Prof Father Manual	0.583 ***	0.804 ***	1.212 **	1.486 ***	-0.702	0.058	0.358	-1.233	0.874 ***
	(0.219)	(0.292)	(0.513)	(0.307)	(0.740)	(0.493)	(0.303)	(1.126)	(0.237)
Prof Father Other	0.641 ***	0.754 ***	1.456 ***	1.498 ***	-2.887 ***	1.161 ***	0.264	1.127	0.917 ***
	(0.205)	(0.264)	(0.481)	(0.298)	(0.759)	(0.407)	(0.282)	(1.043)	(0.211)

Table 9. Coefficients of choice equations (Model I, cont.)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Prof Mother Non-manual	0.425 **	1.326 ***	1.048 ***	0.696 ***	0.440	0.901 ***	-0.363 *	1.087 *	0.599 ***
	(0.182)	(0.194)	(0.264)	(0.213)	(0.357)	(0.271)	(0.212)	(0.645)	(0.155)
Prof Mother Manual	-0.303	1.700 ***	1.316 **	1.152 ***	2.246 ***	1.918 ***	0.218	3.285 *	0.505
	(0.336)	(0.397)	(0.633)	(0.393)	(0.777)	(0.549)	(0.368)	(1.786)	(0.324)
Prof Mother Housewife	0.562 ***	1.647 ***	1.422 ***	0.831 ***	0.229	0.662 **	-0.801 ***	-0.191	0.491 ***
	(0.186)	(0.197)	(0.273)	(0.224)	(0.378)	(0.274)	(0.223)	(0.676)	(0.165)
Prof Mother Other	0.244	1.758 ***	1.456 ***	0.409	0.378	1.053 ***	0.499 *	1.552	0.396 *
	(0.228)	(0.264)	(0.456)	(0.327)	(0.605)	(0.406)	(0.277)	(1.122)	(0.234)
Rsn Maj Job Mkt	3.273 ***	3.357 ***	2.712 ***	3.979 ***	-30.411 **	1.768 ***	-2.472 ***	0.393	-0.236
	(0.269)	(0.411)	(0.658)	(0.415)	(13.329)	(0.630)	(0.577)	(1.486)	(0.349)
Rsn Maj Soc Cont	-0.382	-9.485 ***	-19.753 ***	-7.390 ***	-9.369 ***	-3.449 ***	-1.716 ***	-3.478 ***	-1.464 ***
	(0.314)	(1.242)	(1.788)	(0.697)	(1.269)	(0.447)	(0.281)	(1.002)	(0.213)
Rsn Maj Pers Real	-1.040 ***	-2.048 ***	-3.687 ***	-1.537 ***	-1.261 ***	-2.368 ***	-1.112 ***	-3.638 ***	-0.294 **
	(0.293)	(0.331)	(0.391)	(0.268)	(0.364)	(0.320)	(0.213)	(0.692)	(0.148)
Rsn Maj Other	3.122 ***	0.449	-0.249	2.397 ***	-5.325 ***	0.889	0.949 ***	-1.453	1.337 ***
	(0.256)	(0.525)	(0.703)	(0.414)	(1.513)	(0.565)	(0.321)	(1.345)	(0.256)
Rsn Univ Free	0.789 ***	-0.435	0.872 *	-3.002 ***	-3.930 ***	0.756 *	0.790 ***	2.739 ***	0.283
	(0.256)	(0.405)	(0.492)	(0.336)	(0.537)	(0.388)	(0.238)	(0.891)	(0.178)
Rsn Univ Rep	0.924 ***	0.483	2.188 ***	-2.254 ***	-3.919 ***	1.172 ***	0.599 ***	2.703 ***	0.035
	(0.253)	(0.336)	(0.401)	(0.281)	(0.457)	(0.335)	(0.224)	(0.764)	(0.159)
Rsn Univ Other	1.198 ***	1.912 ***	1.404 ***	-0.874 ***	-2.191 ***	1.157 ***	0.863 ***	1.453	0.076
	(0.265)	(0.341)	(0.494)	(0.302)	(0.485)	(0.396)	(0.252)	(0.930)	(0.189)
Work * Gender	-0.329	-0.441	1.031	-1.430 ***	1.622 **	-2.097 ***	-0.891 **	-6.255 ***	-2.070 ***
	(0.273)	(0.456)	(0.987)	(0.458)	(0.661)	(0.586)	(0.369)	(1.609)	(0.444)
Sec Sch Priv * Gender	0.991 ***	-0.892 **	2.835 ***	1.090 ***	1.514 **	2.681 ***	1.281 ***	7.456 ***	1.146 ***
	(0.264)	(0.440)	(0.591)	(0.363)	(0.726)	(0.487)	(0.380)	(1.146)	(0.369)
Sec Sch Mixed * Gender	0.635	-1.089	4.153 ***	-0.259	0.824	1.175	1.337 **	3.798	1.022
	(0.528)	(0.746)	(1.450)	(0.783)	(1.441)	(1.246)	(0.634)	(3.896)	(0.664)

TABLE 9. Coefficients of choice equations (Model I, cont.)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Sec Sch Tech * Gender	-0.501	-0.611	-0.183	-1.673 ***	-1.428	-0.282	0.163	-5.417 ***	-0.428
	(0.309)	(0.527)	(0.900)	(0.441)	(1.114)	(0.696)	(0.452)	(1.669)	(0.499)
Other Major * Gender	0.614	-2.127 ***	-3.804 ***	-0.014	-0.947	-2.193 ***	2.431 ***	-2.748 *	1.275 ***
	(0.393)	(0.386)	(0.654)	(0.392)	(0.712)	(0.474)	(0.370)	(1.625)	(0.369)
Reg Fee * Gender	1.673 ***	4.757 ***	11.622 ***	1.131	7.464 **	3.321 *	0.407	-4.130	1.226
	(0.497)	(1.022)	(2.371)	(0.983)	(3.101)	(1.876)	(0.728)	(2.994)	(0.810)
Rsn Maj J Mkt * Gender	-1.134 ***	-2.432 ***	-0.061	-1.380 ***	-10.954	-0.378	-1.832 *	1.510	-2.236 ***
	(0.325)	(0.495)	(0.822)	(0.529)	(15.883)	(0.816)	(1.108)	(2.028)	(0.782)
Rsn Maj S Cont * Gender	-1.323 ***	4.506 ***	7.607 ***	5.011 ***	-2.598	1.873 ***	1.824 ***	1.825	-0.488
	(0.432)	(1.299)	(1.873)	(0.787)	(2.377)	(0.633)	(0.440)	(1.521)	(0.491)
Rsn Maj P Real * Gender	-0.225	1.556 ***	1.228 **	1.172 ***	1.824 ***	1.465 ***	0.631 *	2.735 ***	0.586 *
	(0.366)	(0.382)	(0.496)	(0.365)	(0.580)	(0.456)	(0.378)	(1.054)	(0.318)
Rsn Maj Other * Gender	-1.340 ***	-0.598	-1.148	-1.511 ***	1.747	-1.174	-0.290	2.994	-0.209
	(0.324)	(0.603)	(0.848)	(0.549)	(1.954)	(0.757)	(0.528)	(1.836)	(0.479)
Rsn Univ Free * Gender	-0.576 *	-1.097 **	-2.564 ***	1.953 ***	-0.400	-0.161	-0.332	-0.102	0.626 *
	(0.310)	(0.468)	(0.607)	(0.442)	(0.827)	(0.532)	(0.426)	(1.300)	(0.353)
Rsn Univ Rep * Gender	-1.237 ***	-1.567 ***	-3.000 ***	1.331 ***	-3.249 ***	-0.215	0.030	-1.530	-0.275
	(0.315)	(0.392)	(0.502)	(0.379)	(0.866)	(0.462)	(0.399)	(1.116)	(0.341)
Rsn Univ Other * Gender	-0.942 ***	-1.177 ***	-1.956 ***	1.842 ***	-0.854	-0.687	0.766 *	-1.039	-0.559
	(0.325)	(0.397)	(0.612)	(0.398)	(0.763)	(0.552)	(0.406)	(1.349)	(0.403)
Constant	-5.743 ***	-11.203 ***	-9.306 ***	-4.425 ***	-7.956 ***	-8.838 ***	-6.131 ***	19.486 ***	-3.539 ***
	(0.408)	(0.568)	(0.892)	(0.531)	(0.980)	(0.712)	(0.473)	(2.265)	(0.395)
Number of observations	120,058								

Estimations include the explanatory variables listed in Table 4. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 10. Marginal effects on choice probabilities (Model I)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Gender	0.73 ***	4.10 ***	13.86 ***	-2.82 ***	-0.78 ***	1.69 ***	-3.04 ***	-10.38 ***	-3.35 ***
	(0.08)	(0.13)	(0.19)	(0.14)	(0.10)	(0.17)	(0.10)	(0.21)	(0.12)
White	-0.20 **	-0.21	-1.42 ***	0.29 *	-0.13	0.61 ***	-0.06	0.49 **	0.63 ***
	(0.09)	(0.14)	(0.23)	(0.15)	(0.11)	(0.18)	(0.11)	(0.24)	(0.12)
Work	0.04	-1.06 ***	16.13 ***	-2.54 ***	0.87 ***	-2.16 ***	1.09 ***	-11.39 ***	-0.97 ***
	(0.09)	(0.17)	(0.51)	(0.18)	(0.21)	(0.29)	(0.16)	(0.40)	(0.17)
Prim Sch Priv	-0.36 ***	-0.18	-0.21	-0.36 *	0.07	0.58 **	-0.62 ***	2.28 ***	-1.20 ***
	(0.11)	(0.19)	(0.33)	(0.19)	(0.14)	(0.26)	(0.16)	(0.33)	(0.18)
Prim Sch Mixed	-0.36 ***	0.32	-0.97 ***	-0.08	0.07	0.32	-0.32 **	1.65 ***	-0.63 ***
	(0.12)	(0.20)	(0.35)	(0.20)	(0.15)	(0.27)	(0.16)	(0.35)	(0.19)
Sec Sch Priv	-3.57 ***	3.80 ***	11.17 ***	-2.14 ***	-1.42 ***	-7.60 ***	-0.46	0.51 ***	-0.30
	(0.49)	(1.48)	(1.04)	(0.26)	(0.28)	(0.80)	(0.33)	(0.14)	(0.21)
Sec Sch Mixed	-4.44 ***	-8.26 ***	13.08 ***	-1.57 ***	-1.07 ***	-5.56 ***	-0.36	8.27 ***	-0.09
	(0.82)	(1.29)	(1.43)	(0.36)	(0.39)	(1.02)	(0.46)	(0.66)	(0.30)
Sec Sch Tech	-0.34 ***	-0.98 ***	21.11 ***	-1.75 ***	-1.34 ***	-4.85 ***	-1.10 ***	-8.93 ***	-1.80 ***
	(0.10)	(0.21)	(0.49)	(0.22)	(0.12)	(0.27)	(0.14)	(0.45)	(0.19)
Other Major	-1.11 ***	-1.04 ***	14.59 ***	-2.33 ***	-1.23 ***	-0.86 **	-0.16	-5.41 ***	-2.45 ***
	(0.09)	(0.25)	(0.62)	(0.27)	(0.13)	(0.38)	(0.18)	(0.50)	(0.17)
Reg Fee	-0.35 ***	0.31	15.08 ***	-3.73 ***	-1.47 ***	-4.43 ***	7.24 ***	-14.11 ***	1.47 ***
	(0.09)	(0.46)	(0.89)	(0.24)	(0.16)	(0.39)	(0.53)	(0.64)	(0.39)
Prof Father Non-manual	0.24 **	0.62 ***	0.01	0.39 **	0.21 *	-0.37	0.15	-1.17 ***	-0.08
	(0.11)	(0.17)	(0.27)	(0.18)	(0.13)	(0.23)	(0.14)	(0.28)	(0.15)
Prof Father Manual	0.33 **	0.87 ***	1.00 *	0.29	0.30	0.03	-0.09	-3.03 ***	0.29
	(0.15)	(0.28)	(0.53)	(0.29)	(0.23)	(0.40)	(0.20)	(0.52)	(0.25)
Prof Father Other	0.30 **	0.94 ***	-0.33	0.69 **	0.52 ***	0.50	0.26	-2.73 ***	-0.15
	(0.14)	(0.24)	(0.43)	(0.27)	(0.20)	(0.34)	(0.18)	(0.43)	(0.21)

Table 10. Marginal effects on choice probabilities (Model I, cont.)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Prof Mother Non-manual	0.11	0.26	0.39	0.61 ***	0.20 *	0.06	-0.09	-1.78 ***	0.24
	(0.12)	(0.16)	(0.26)	(0.19)	(0.12)	(0.21)	(0.15)	(0.27)	(0.15)
Prof Mother Manual	0.14	0.50	0.56	0.38	-0.26	0.17	0.08	-1.48 **	-0.09
	(0.20)	(0.37)	(0.66)	(0.36)	(0.30)	(0.58)	(0.26)	(0.66)	(0.32)
Prof Mother Housewife	0.15	0.46 ***	1.72 ***	0.95 ***	-0.47 ***	-1.10 ***	-0.36 **	-1.88 ***	0.54 ***
	(0.13)	(0.17)	(0.27)	(0.20)	(0.12)	(0.23)	(0.15)	(0.28)	(0.16)
Prof Mother Other	0.17	0.12	0.90 **	0.67 **	0.20	-0.32	-0.03	-1.79 ***	0.08
	(0.15)	(0.26)	(0.43)	(0.27)	(0.20)	(0.36)	(0.20)	(0.44)	(0.22)
Rsn Maj Job Mkt	2.55 ***	1.52 ***	2.43 ***	3.71 ***	-3.32 ***	0.14	-1.71 ***	-4.02 ***	-1.29 ***
	(0.20)	(0.28)	(0.43)	(0.32)	(0.21)	(0.35)	(0.16)	(0.42)	(0.23)
Rsn Maj Soc Cont	-0.17	-2.87 ***	-12.07 ***	-1.06 ***	-1.73 ***	3.12 ***	0.48 ***	13.84 ***	0.45 **
	(0.12)	(0.19)	(0.33)	(0.23)	(0.16)	(0.30)	(0.16)	(0.39)	(0.19)
Rsn Maj Pers Real	-0.49 ***	-0.29 *	-2.97 ***	0.05	0.23 **	-0.14	-0.26 **	3.14 ***	0.72 ***
	(0.08)	(0.15)	(0.23)	(0.16)	(0.11)	(0.19)	(0.11)	(0.24)	(0.14)
Rsn Maj Other	2.36 ***	-0.18	-3.17 ***	1.72 ***	-1.44 ***	-0.10	0.44 **	-1.07 **	1.43 ***
	(0.20)	(0.25)	(0.40)	(0.32)	(0.19)	(0.34)	(0.20)	(0.43)	(0.27)
Rsn Univ Free	0.37 ***	-1.05 ***	-1.54 ***	-2.02 ***	-1.47 ***	1.29 ***	0.54 ***	3.12 ***	0.75 ***
	(0.10)	(0.16)	(0.27)	(0.18)	(0.12)	(0.23)	(0.13)	(0.29)	(0.16)
Rsn Univ Rep	0.14	-0.56 ***	0.28	-1.76 ***	-1.76 ***	1.60 ***	0.44 ***	1.66 ***	-0.04
	(0.10)	(0.15)	(0.23)	(0.16)	(0.11)	(0.19)	(0.12)	(0.24)	(0.14)
Rsn Univ Other	0.38 ***	1.16 ***	-0.86 ***	-0.32 *	-1.18 ***	0.61 ***	0.70 ***	-0.15	-0.35 **
	(0.11)	(0.18)	(0.28)	(0.19)	(0.13)	(0.22)	(0.13)	(0.29)	(0.16)
Number of observations	120,058								

* significant at 10%, ** significant at 5%, *** significant at 1%.

TABLE 11. Coefficients of entry equations (Model II)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Gender	-0.538 **	-0.676 ***	5.554 ***	3.774 ***	-0.382 *	1.464 ***	0.395 **	-5.053 ***	0.218
	(0.236)	(0.216)	(0.340)	(0.365)	(0.223)	(0.318)	(0.178)	(0.442)	(0.229)
White	0.174	0.131 *	0.084 *	-0.008	0.192	-0.069	-0.030	0.050	0.143 **
	(0.124)	(0.074)	(0.044)	(0.073)	(0.118)	(0.069)	(0.061)	(0.048)	(0.066)
Work	0.027	0.034	-0.161 **	-0.155	0.164	-0.155	0.002	-0.127 *	-0.047
	(0.142)	(0.090)	(0.064)	(0.104)	(0.122)	(0.099)	(0.074)	(0.077)	(0.089)
Age1	0.083	-0.274 *	-0.129	-0.428 **	-0.145	-0.460 **	-0.688 ***	0.072	-0.550 ***
	(0.299)	(0.166)	(0.126)	(0.188)	(0.228)	(0.179)	(0.166)	(0.137)	(0.163)
Age2	-0.014	-0.321 **	-0.039	-0.281 *	-0.242	-0.347 **	-0.458 ***	0.101	-0.437 ***
	(0.225)	(0.131)	(0.115)	(0.151)	(0.187)	(0.141)	(0.096)	(0.119)	(0.130)
Age3	0.241	-0.117	0.032	-0.277 *	-0.172	-0.274 **	-0.323 ***	0.201 *	-0.295 **
	(0.219)	(0.125)	(0.115)	(0.147)	(0.181)	(0.136)	(0.091)	(0.117)	(0.126)
Prim Sch Priv	0.128	0.083	0.095	0.121	0.201	0.046	0.260 *	0.082	0.131
	(0.181)	(0.104)	(0.065)	(0.102)	(0.143)	(0.108)	(0.144)	(0.067)	(0.097)
Prim Sch Mixed	0.039	0.014	0.105	0.021	0.082	0.001	0.165	0.147 **	-0.034
	(0.196)	(0.107)	(0.070)	(0.108)	(0.138)	(0.114)	(0.124)	(0.071)	(0.107)
Sec Sch Priv	0.101	0.259 **	-0.172 ***	-0.165	0.180	-0.046	0.187	-0.134 *	0.296 **
	(0.177)	(0.112)	(0.061)	(0.105)	(0.132)	(0.102)	(0.154)	(0.068)	(0.118)
Sec Sch Mixed	-0.325	-0.038	-0.152	-0.207	-0.083	-0.570 **	-0.039	-0.424 ***	-0.305
	(0.306)	(0.181)	(0.109)	(0.190)	(0.203)	(0.232)	(0.234)	(0.134)	(0.210)
Sec Sch Tech	-0.184	-0.048	0.094	0.361 ***	-0.486 **	-0.085	-0.041	-0.186 **	0.000
	(0.176)	(0.113)	(0.069)	(0.101)	(0.238)	(0.121)	(0.099)	(0.094)	(0.125)
Other Major	0.598 ***	0.309 ***	0.198 ***	0.398 ***	-0.074	0.358 ***	0.205 **	0.342 ***	0.665 ***
	(0.219)	(0.098)	(0.072)	(0.112)	(0.148)	(0.097)	(0.096)	(0.072)	(0.107)
Reg Fee	-1.077 ***	-0.841 ***	-0.473 ***	-0.827 ***	-0.168	-1.727 ***	-0.591 ***	-0.233 *	-0.429 **
	(0.210)	(0.184)	(0.137)	(0.213)	(0.202)	(0.441)	(0.167)	(0.139)	(0.175)

TABLE 11. Coefficients of entry equations (Model II, cont.)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Prep Course	0.259 **	0.017	-0.007	0.196 ***	0.371 ***	0.286 ***	0.106 *	0.301 ***	0.265 ***
	(0.101)	(0.066)	(0.024)	(0.067)	(0.078)	(0.057)	(0.060)	(0.025)	(0.062)
ENEM	0.006 ***	-0.000	0.059 ***	0.048 ***	0.002	0.041 ***	0.005 ***	0.021 ***	0.016 ***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
ENEM * Gender	0.003	0.010 ***	-0.052 ***	-0.037 ***	0.005 **	-0.016 ***	-0.002	0.044 ***	-0.005 *
	(0.003)	(0.002)	(0.003)	(0.004)	(0.002)	(0.003)	(0.002)	(0.004)	(0.002)
Constant	-0.818 **	-1.269 ***	-8.368 ***	-6.287 ***	-2.371 ***	-5.722 ***	-1.657 ***	-5.615 ***	-3.063 ***
	(0.365)	(0.275)	(0.359)	(0.390)	(0.327)	(0.308)	(0.222)	(0.187)	(0.250)
Number of obs.	39,494								

Estimations include the explanatory variables listed in Table 4. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 12. Marginal effects on the probability of entry (Model II)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Gender	-7.08 *	3.21	2.93 ***	2.05 *	1.31	-1.46 **	4.43	-1.47 ***	-3.37 **
	(3.73)	(2.08)	(0.47)	(1.20)	(1.70)	(0.74)	(2.69)	(0.19)	(1.64)
White	4.01	2.27 *	0.62 *	-0.10	2.18 *	-0.59	-0.55	0.14	2.20 **
	(2.85)	(1.25)	(0.33)	(0.96)	(1.29)	(0.60)	(1.14)	(0.13)	(1.01)
Work	0.61	0.61	-1.17 ***	-1.98	2.01	-1.26	0.04	-0.34 *	-0.73
	(3.27)	(1.60)	(0.44)	(1.29)	(1.55)	(0.77)	(1.38)	(0.19)	(1.38)
Age1	1.92	-5.07	-0.96	-5.88 **	-1.86	-4.20 **	-13.15 ***	0.19	-9.18 ***
	(6.90)	(3.08)	(0.97)	(2.68)	(2.96)	(1.71)	(3.08)	(0.35)	(2.81)
Age2	-0.32	-5.86 **	-0.30	-4.02 *	-3.02	-3.29 **	-9.13 ***	0.26	-7.48 ***
	(5.18)	(2.51)	(0.90)	(2.29)	(2.47)	(1.47)	(2.02)	(0.30)	(2.41)
Age3	5.53	-2.23	0.26	-3.97 *	-2.19	-2.66 *	-6.59 ***	0.55 *	-5.21 **
	(5.03)	(2.41)	(0.90)	(2.22)	(2.39)	(1.41)	(1.91)	(0.30)	(2.33)
Prim Sch Priv	2.96	1.46	0.71	1.58	2.31	0.39	4.76 *	0.22	2.06
	(4.20)	(1.82)	(0.47)	(1.32)	(1.62)	(0.91)	(2.60)	(0.18)	(1.51)
Prim Sch Mixed	0.90	0.24	0.78	0.27	0.91	0.01	2.95	0.41 **	-0.52
	(4.53)	(1.84)	(0.52)	(1.37)	(1.52)	(0.95)	(2.24)	(0.20)	(1.60)
Sec Sch Priv	2.34	4.46 **	-1.35 ***	-2.24	2.07	-0.40	3.43	-0.39 *	4.56 **
	(4.11)	(1.90)	(0.50)	(1.45)	(1.49)	(0.90)	(2.79)	(0.21)	(1.77)
Sec Sch Mixed	-7.47	-0.62	-1.20	-2.78	-0.87	-4.11 ***	-0.68	-1.09 ***	-3.96
	(6.97)	(2.88)	(0.83)	(2.45)	(2.10)	(1.45)	(4.06)	(0.31)	(2.58)
Sec Sch Tech	-4.23	-0.84	0.74	5.18 ***	-4.96 **	-0.71	-0.76	-0.48 **	0.01
	(4.04)	(1.95)	(0.56)	(1.55)	(2.09)	(0.97)	(1.81)	(0.23)	(1.97)

TABLE 12. Marginal effects on the probability of entry (Model II, cont.)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Other Major	13.39 ***	5.74 ***	1.61 ***	5.77 ***	-0.86	3.38 ***	3.94 **	1.09 ***	11.89 ***
	(4.64)	(1.91)	(0.62)	(1.76)	(1.67)	(1.01)	(1.90)	(0.27)	(2.15)
Reg Fee	-24.17 ***	-12.28 ***	-3.06 ***	-8.79 ***	-1.89	-8.35 ***	-9.80 ***	-0.59 *	-6.13 ***
	(4.28)	(2.18)	(0.74)	(1.76)	(2.16)	(1.02)	(2.49)	(0.32)	(2.27)
Prep Course	5.99 ***	0.30	-0.06	2.57 ***	4.26 ***	2.36 ***	1.96 *	0.80 ***	4.12 ***
	(2.32)	(1.16)	(0.18)	(0.86)	(0.94)	(0.46)	(1.11)	(0.07)	(0.96)
ENEM	0.17 ***	0.08 ***	0.21 ***	0.37 ***	0.05 ***	0.28 ***	0.08 ***	0.11 ***	0.22 ***
	(0.03)	(0.02)	(0.01)	(0.03)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)
Number of obs.	39,494								

* significant at 10%, ** significant at 5%, *** significant at 1%.

TABLE 13. Marginal effects on the probability of entry for different ENEM groups (Model II)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
= 0	-11.63 **	-9.67 ***	5.59 ***	7.84 ***	-3.23	0.77 ***	6.18 **	-0.51 ***	1.61
	(5.13)	(3.41)	(0.37)	(1.27)	(1.99)	(0.23)	(2.83)	(0.06)	(1.71)
(0, 70]	-8.39 **	-1.39	6.66 ***	8.61 ***	-0.52	1.12 ***	4.12 **	-1.34 ***	-0.59
	(3.26)	(1.74)	(0.37)	(0.97)	(1.28)	(0.37)	(2.01)	(0.12)	(1.21)
(70, 80]	-7.69 **	0.91	6.79 ***	8.24 ***	0.35	0.96 *	4.17 *	-1.81 ***	-1.73
	(3.51)	(1.81)	(0.38)	(1.01)	(1.39)	(0.49)	(2.27)	(0.15)	(1.34)
(80, 90]	-7.18 *	2.62	6.12 ***	6.59 ***	0.99	0.38	4.24 *	-2.06 ***	-2.68 *
	(3.80)	(1.96)	(0.41)	(1.09)	(1.57)	(0.60)	(2.52)	(0.18)	(1.51)
(90, 100]	-6.63	4.49 **	4.35 ***	3.41 ***	1.71	-0.82	4.28	-2.07 ***	-3.79 **
	(4.13)	(2.16)	(0.48)	(1.27)	(1.78)	(0.77)	(2.82)	(0.21)	(1.76)
(100, 110]	-6.00	6.54 ***	0.88	-1.76	2.55	-2.91 ***	4.25	-1.58 ***	-5.02 **
	(4.48)	(2.42)	(0.65)	(1.62)	(2.08)	(1.05)	(3.14)	(0.25)	(2.10)
(110, 120]	-5.39	8.49 ***	-4.24 ***	-8.01 ***	3.41	-5.68 ***	4.16	-0.35	-6.21 **
	(4.77)	(2.69)	(0.98)	(2.14)	(2.38)	(1.46)	(3.43)	(0.35)	(2.47)
Average	-7.08 *	3.21	2.93 ***	2.05 *	1.31	-1.46 **	4.43	-1.47 ***	-3.37 **
	(3.73)	(2.08)	(0.47)	(1.20)	(1.70)	(0.74)	(2.69)	(0.19)	(1.64)
Number of obs.	39,494								

* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 14. Coefficients of choice equations (Model II)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Gender	1.534 ***	5.817 ***	8.143 ***	-4.379 ***	-3.010 *	-1.242	-6.168 ***	-8.669 *	-4.047 ***
	(0.552)	(1.035)	(1.848)	(0.991)	(1.624)	(1.473)	(1.248)	(4.739)	(0.991)
White	0.089	-0.071	-2.340 ***	0.475	1.969 **	0.181	0.031	-2.433	0.563
	(0.239)	(0.333)	(0.835)	(0.388)	(0.969)	(0.580)	(0.376)	(2.525)	(0.364)
Work	1.125 ***	1.417 **	0.340	0.050	4.076 ***	1.615	2.008 ***	-16.806 ***	0.063
	(0.344)	(0.715)	(1.857)	(0.853)	(0.957)	(1.261)	(0.549)	(4.379)	(0.552)
Age1	-0.915 *	0.428	8.423 ***	-2.719 ***	-2.811 *	-1.453	-4.614 ***	-52.531 ***	-3.773 ***
	(0.551)	(0.715)	(2.116)	(0.971)	(1.496)	(1.467)	(0.988)	(9.284)	(0.826)
Age2	-0.757 *	0.061	6.862 ***	-0.660	-2.302 *	1.509	-2.631 ***	-47.722 ***	-2.051 ***
	(0.401)	(0.614)	(1.881)	(0.778)	(1.231)	(1.234)	(0.606)	(8.608)	(0.673)
Age3	-0.527	-1.291 **	1.052	-1.179	-1.402	-0.695	-2.305 ***	-39.683 ***	-1.859 ***
	(0.375)	(0.586)	(1.818)	(0.748)	(1.158)	(1.178)	(0.559)	(8.385)	(0.643)
Prim Sch Priv	-0.256	0.390	-0.696	-0.144	1.770	0.472	-0.490	0.087	-1.385 ***
	(0.345)	(0.444)	(1.175)	(0.536)	(1.083)	(0.910)	(0.590)	(3.427)	(0.448)
Prim Sch Mixed	-0.149	1.283 ***	0.704	1.170 **	3.601 ***	2.474 **	0.322	1.787	0.282
	(0.370)	(0.471)	(1.254)	(0.580)	(1.084)	(1.005)	(0.573)	(3.568)	(0.504)
Sec Sch Priv	-0.516	0.591	3.401 **	0.053	-1.433	1.140	0.171	21.745 ***	0.980 *
	(0.398)	(0.775)	(1.475)	(0.744)	(1.207)	(1.122)	(0.657)	(3.870)	(0.589)
Sec Sch Mixed	-0.797	1.430	2.359	-0.105	1.789	4.296	1.035	29.240 ***	2.293 *
	(0.854)	(1.302)	(2.796)	(1.354)	(1.901)	(3.096)	(1.014)	(9.737)	(1.274)
Sec Sch Tech	1.102 **	0.949	4.701 **	2.099 **	-4.171	0.603	0.164	4.966	-0.079
	(0.438)	(1.030)	(1.986)	(0.846)	(2.648)	(1.596)	(0.796)	(6.013)	(0.745)
Other Major	-0.617	2.213 ***	-1.349	-1.730 **	0.779	-0.687	-0.001	-25.782 ***	-2.455 ***
	(0.513)	(0.733)	(1.782)	(0.810)	(1.178)	(1.057)	(0.630)	(3.969)	(0.546)
Reg Fee	-0.843	4.718 ***	9.701 **	6.264 **	0.261	52.150 **	16.634 ***	10.506	8.621 ***
	(0.543)	(1.535)	(4.050)	(2.743)	(2.034)	(25.258)	(2.182)	(8.607)	(1.760)
Prof Father Non-manual	0.531	1.244 ***	1.016	1.081 **	0.817	0.855	0.387	-0.716	0.069
	(0.388)	(0.388)	(0.957)	(0.475)	(0.837)	(0.746)	(0.516)	(2.899)	(0.382)
Prof Father Manual	0.561	1.647 ***	3.258 *	2.686 ***	2.473 *	3.129 **	0.036	7.795	0.461
	(0.485)	(0.620)	(1.683)	(0.763)	(1.332)	(1.473)	(0.741)	(5.454)	(0.624)
Prof Father Other	1.009 **	1.851 ***	3.258 **	3.193 ***	0.662	2.483 **	1.068	10.536 **	0.755
	(0.442)	(0.568)	(1.559)	(0.742)	(1.169)	(1.093)	(0.650)	(4.782)	(0.583)

Table 14. Coefficients of choice equations (Model II, cont.)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Prof Mother Non-manual	0.378	0.321	-0.466	0.722	0.752	-0.202	0.014	-0.452	0.781 *
	(0.429)	(0.421)	(0.945)	(0.498)	(0.800)	(0.736)	(0.545)	(2.994)	(0.401)
Prof Mother Manual	0.379	-0.055	-1.695	0.623	-0.324	-2.535 *	-0.547	-8.692	-0.626
	(0.624)	(0.750)	(1.832)	(0.948)	(1.828)	(1.523)	(0.944)	(6.621)	(0.817)
Prof Mother Housewife	1.162 ***	1.569 ***	1.462	1.759 ***	0.147	0.417	-0.236	-0.234	1.342 ***
	(0.435)	(0.445)	(0.962)	(0.517)	(1.014)	(0.770)	(0.566)	(3.153)	(0.437)
Prof Mother Other	0.644	1.052 **	3.873 ***	2.077 ***	-0.089	-0.029	0.772	-3.273	1.561 ***
	(0.493)	(0.533)	(1.453)	(0.712)	(1.254)	(0.893)	(0.655)	(3.803)	(0.562)
Rsn Maj Job Mkt	2.127 ***	1.505 *	-2.622	1.913 **	-40.932 ***	-3.487 **	-5.526 ***	-20.889 ***	-3.282 ***
	(0.431)	(0.778)	(1.898)	(0.933)	(11.981)	(1.383)	(1.054)	(4.383)	(0.773)
Rsn Maj Soc Cont	1.427 ***	-6.925 ***	-6.604 ***	-0.021	-6.889 ***	4.942 ***	2.072 ***	35.109 ***	2.676 ***
	(0.463)	(1.580)	(1.953)	(1.065)	(1.780)	(1.307)	(0.711)	(5.117)	(0.662)
Rsn Maj Pers Real	0.228	-1.790 ***	-1.218	-0.020	0.070	0.207	-0.747	7.308 ***	1.046 **
	(0.392)	(0.679)	(1.135)	(0.604)	(0.797)	(0.808)	(0.492)	(2.601)	(0.411)
Rsn Maj Other	2.792 ***	-0.887	-1.464	3.271 ***	-6.297 ***	0.735	0.757	-5.197	1.412 **
	(0.432)	(0.981)	(1.986)	(0.959)	(2.044)	(1.322)	(0.737)	(4.237)	(0.660)
Rsn Univ Free	1.104 ***	-0.052	1.031	-2.771 ***	-3.590 ***	1.100	1.246 **	5.626 *	1.602 ***
	(0.379)	(0.732)	(1.430)	(0.763)	(1.099)	(0.975)	(0.582)	(3.114)	(0.485)
Rsn Univ Rep	0.677 *	0.126	1.207	-3.266 ***	-7.282 ***	-0.457	0.996 *	1.453	-0.467
	(0.380)	(0.645)	(1.197)	(0.664)	(1.216)	(0.856)	(0.523)	(2.774)	(0.454)
Rsn Univ Other	1.346 ***	1.222 *	2.633 *	-0.948	-5.194 ***	1.177	1.048 *	3.404	0.359
	(0.373)	(0.708)	(1.392)	(0.708)	(1.172)	(0.964)	(0.578)	(3.044)	(0.496)
Work * Gender	0.295	0.855	5.096 **	1.749 *	3.897 ***	2.140	1.045	13.663 ***	0.188
	(0.440)	(0.846)	(2.072)	(1.000)	(1.472)	(1.533)	(0.842)	(5.203)	(0.905)
Sec Sch Priv * Gender	-0.461	-2.763 ***	-3.511 **	-1.247	-2.490 *	-2.935 **	1.348	-12.851 ***	-1.796 *
	(0.441)	(0.913)	(1.569)	(0.886)	(1.489)	(1.298)	(0.915)	(4.150)	(0.925)
Sec Sch Mixed * Gender	-1.029	-4.082 ***	-5.279 *	-2.466	-6.299 **	-8.757 **	-0.553	-24.797 **	-6.329 **
	(0.942)	(1.563)	(3.047)	(1.771)	(3.097)	(3.933)	(1.563)	(10.451)	(2.660)

TABLE 14. Coefficients of choice equations (Model II, cont.)

	Technologies	Exact Sc. and Agr. Eng.	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Sec Sch Tech * Gender	-0.216	2.461 **	1.345	-0.002	-1.265	2.371	0.933	-0.763	-0.372
	(0.598)	(1.250)	(2.236)	(1.035)	(3.447)	(2.007)	(1.097)	(6.937)	(1.294)
Other Major * Gender	1.117 *	-0.719	0.463	1.269	-0.200	2.335 *	2.683 ***	8.622 *	3.141 ***
	(0.587)	(0.827)	(1.967)	(0.972)	(1.608)	(1.299)	(0.883)	(4.429)	(0.837)
Reg Fee * Gender	2.066 **	6.026 **	-0.680	0.183	1.715	-3.498	-4.184 *	10.268	-1.977
	(0.831)	(2.764)	(3.819)	(2.521)	(2.956)	(11.152)	(2.481)	(10.493)	(2.016)
Rsn Maj J Mkt * Gender	0.681	0.430	8.385 ***	3.086 **	28.348 **	7.490 ***	0.775	26.006 ***	3.995 ***
	(0.575)	(0.966)	(2.289)	(1.208)	(12.317)	(1.879)	(2.100)	(6.253)	(1.266)
Rsn Maj S Cont * Gender	-3.900 ***	-1.509	-15.469 ***	-4.974 ***	-9.593 ***	-13.600 ***	-3.384 ***	-48.815 ***	-8.438 ***
	(0.692)	(1.774)	(2.386)	(1.322)	(3.303)	(1.831)	(0.986)	(6.639)	(1.171)
Rsn Maj P Real * Gender	-1.623 ***	1.323 *	-2.170	-0.305	0.416	-1.848	0.174	-7.429 **	-0.239
	(0.576)	(0.786)	(1.398)	(0.815)	(1.189)	(1.168)	(0.785)	(3.790)	(0.732)
Rsn Maj Other * Gender	-0.667	1.412	-1.782	-2.405 *	3.260	-2.097	-0.143	4.530	0.347
	(0.571)	(1.140)	(2.332)	(1.243)	(2.642)	(1.844)	(1.102)	(6.017)	(1.078)
Rsn Univ Free * Gender	-1.156 **	-1.344	-2.565	1.454	-0.219	1.000	-0.893	4.200	-0.799
	(0.520)	(0.880)	(1.719)	(0.994)	(1.613)	(1.372)	(0.916)	(4.517)	(0.835)
Rsn Univ Rep * Gender	-1.163 **	-1.991 **	-1.812	2.309 ***	-0.424	1.622	-0.335	0.463	-0.196
	(0.525)	(0.783)	(1.468)	(0.871)	(1.801)	(1.200)	(0.813)	(3.934)	(0.784)
Rsn Univ Other * Gender	-0.455	-0.201	-1.211	2.763 ***	0.640	1.027	0.545	2.705	-0.279
	(0.496)	(0.823)	(1.657)	(0.915)	(1.686)	(1.341)	(0.845)	(4.351)	(0.854)
Constant	-4.144 ***	-6.121 ***	3.351	3.043 **	1.086	3.502 *	-0.999	83.701 ***	2.383 **
	(0.764)	(1.216)	(2.805)	(1.277)	(2.021)	(1.881)	(1.045)	(10.621)	(1.051)
Sigma	0.213 ***								
	(0.036)								
Number of observations	39,494								

Estimations include the explanatory variables listed in Table 4. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 15. Marginal effects on choice probabilities (Model II)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Gender	0.79 ***	4.52 ***	24.14 ***	-2.81 ***	-1.00 ***	0.07	-3.41 ***	-16.95 ***	-5.36 ***
	(0.15)	(0.22)	(0.41)	(0.27)	(0.18)	(0.32)	(0.19)	(0.41)	(0.25)
White	-0.07	-0.02	-2.11 ***	0.62 **	0.43 **	0.14	-0.05	0.23	0.84 ***
	(0.17)	(0.25)	(0.47)	(0.28)	(0.21)	(0.36)	(0.22)	(0.49)	(0.29)
Work	1.04 ***	1.67 ***	1.25 *	0.59	2.88 ***	2.16 ***	2.12 ***	-12.05 ***	0.33
	(0.24)	(0.37)	(0.70)	(0.43)	(0.39)	(0.61)	(0.33)	(0.68)	(0.42)
Prim Sch Priv	-0.41 *	0.24	0.23	-0.28	0.27	0.42	-0.39	2.04 ***	-2.11 ***
	(0.24)	(0.34)	(0.68)	(0.40)	(0.30)	(0.52)	(0.43)	(0.70)	(0.44)
Prim Sch Mixed	-0.49 **	0.48	-0.48	-0.13	0.72 **	0.45	-0.29	0.90	-1.17 **
	(0.24)	(0.37)	(0.73)	(0.43)	(0.32)	(0.57)	(0.39)	(0.76)	(0.47)
Sec Sch Priv	-0.74 ***	-1.38 ***	-1.59 **	-1.43 ***	-1.62 ***	-1.45 **	0.09	7.83 ***	0.29
	(0.25)	(0.38)	(0.69)	(0.43)	(0.38)	(0.57)	(0.42)	(0.71)	(0.42)
Sec Sch Mixed	-0.45	-0.44	-2.14 *	-1.33 **	-0.10	-1.33	0.71	4.53 ***	0.55
	(0.35)	(0.57)	(1.12)	(0.64)	(0.56)	(0.88)	(0.55)	(1.16)	(0.65)
Sec Sch Tech	0.51 *	1.54 ***	7.78 ***	2.05 ***	-1.09 ***	-1.64 ***	-0.18	-7.77 ***	-1.20 ***
	(0.27)	(0.44)	(0.85)	(0.54)	(0.29)	(0.59)	(0.32)	(0.84)	(0.45)
Other Major	-0.51 **	2.00 ***	2.90 ***	-0.76	0.61	4.26 ***	0.74 **	-7.81 ***	-1.44 ***
	(0.22)	(0.52)	(0.94)	(0.49)	(0.38)	(0.78)	(0.36)	(0.83)	(0.46)
Reg Fee	0.04	5.39 ***	-10.98 ***	-1.69 **	-1.23 ***	0.86	18.67 ***	-15.54 ***	4.49 ***
	(0.26)	(1.33)	(1.10)	(0.67)	(0.31)	(1.77)	(3.30)	(1.51)	(1.34)
Prof Father Non-manual	0.38 *	0.99 ***	0.41	0.66 *	0.37	-0.80 *	0.11	-1.66 ***	-0.48
	(0.22)	(0.32)	(0.56)	(0.36)	(0.25)	(0.43)	(0.35)	(0.57)	(0.36)
Prof Father Manual	0.36	0.92 *	1.57	1.90 ***	0.81 *	-0.94	-0.30	-3.95 ***	-0.36
	(0.29)	(0.48)	(0.99)	(0.61)	(0.45)	(0.74)	(0.43)	(0.98)	(0.54)
Prof Father Other	0.52 *	1.13 ***	0.08	2.12 ***	-0.07	-0.83	0.34	-2.97 ***	-0.31
	(0.29)	(0.43)	(0.83)	(0.55)	(0.33)	(0.60)	(0.42)	(0.82)	(0.49)

TABLE 15. Marginal effects on choice probabilities (Model II, cont.)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Prof Mother Non-manual	-0.09	0.21	-0.17	0.65 *	0.12	-0.21	-0.15	-1.34 **	0.98 ***
	(0.22)	(0.31)	(0.54)	(0.35)	(0.25)	(0.42)	(0.41)	(0.56)	(0.34)
Prof Mother Manual	-0.02	0.02	2.62 *	1.59 *	0.21	-0.87	-0.24	-3.17 **	-0.14
	(0.38)	(0.61)	(1.34)	(0.85)	(0.59)	(1.07)	(0.60)	(1.39)	(0.69)
Prof Mother Housewife	0.28	0.71 **	2.64 ***	0.96 **	-0.38	-1.17 ***	-0.76 *	-3.07 ***	0.79 **
	(0.24)	(0.35)	(0.59)	(0.40)	(0.27)	(0.44)	(0.40)	(0.59)	(0.37)
Prof Mother Other	0.07	0.27	2.02 ***	0.82 *	-0.38	-0.58	-0.05	-3.14 ***	0.97 *
	(0.27)	(0.41)	(0.78)	(0.48)	(0.33)	(0.61)	(0.48)	(0.82)	(0.50)
Rsn Maj Job Mkt	2.78 ***	1.57 ***	4.11 ***	6.12 ***	-3.54 ***	0.10	-2.13 ***	-7.39 ***	-1.62 ***
	(0.38)	(0.45)	(0.79)	(0.63)	(0.16)	(0.59)	(0.26)	(0.77)	(0.42)
Rsn Maj Soc Cont	0.22	-3.16 ***	-18.24 ***	-0.82 **	-2.46 ***	1.80 ***	0.95 ***	21.26 ***	0.44
	(0.23)	(0.28)	(0.54)	(0.40)	(0.22)	(0.52)	(0.33)	(0.69)	(0.39)
Rsn Maj Pers Real	-0.32 **	-0.33	-4.33 ***	-0.01	0.24	-0.40	-0.51 **	4.35 ***	1.31 ***
	(0.15)	(0.27)	(0.49)	(0.30)	(0.24)	(0.35)	(0.21)	(0.49)	(0.31)
Rsn Maj Other	3.11 ***	0.41	-5.25 ***	3.23 ***	-1.73 ***	0.01	0.46	-2.38 ***	2.16 ***
	(0.41)	(0.45)	(0.79)	(0.60)	(0.30)	(0.62)	(0.39)	(0.83)	(0.55)
Rsn Univ Free	0.37 *	-1.05 ***	-2.53 ***	-2.72 ***	-1.79 ***	1.47 ***	0.51 **	4.11 ***	1.64 ***
	(0.20)	(0.29)	(0.55)	(0.33)	(0.25)	(0.43)	(0.26)	(0.58)	(0.36)
Rsn Univ Rep	0.12	-1.03 ***	1.07 **	-2.11 ***	-2.61 ***	1.22 ***	0.85 ***	2.76 ***	-0.27
	(0.19)	(0.26)	(0.48)	(0.31)	(0.21)	(0.36)	(0.23)	(0.49)	(0.30)
Rsn Univ Other	0.76 ***	0.62 *	0.16	-0.47	-2.32 ***	0.52	0.59 **	0.37	-0.24
	(0.22)	(0.32)	(0.56)	(0.38)	(0.23)	(0.42)	(0.24)	(0.56)	(0.33)
Number of observations	39,494								

* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 16. Simulated Choice Probabilities (Model II)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Choice probabilities with own entrance probabilities									
Male	2.44 ***	7.41 ***	38.90 ***	7.09 ***	2.74 ***	12.71 ***	2.08 ***	22.41 ***	4.21 ***
Female	1.62 ***	3.01 ***	12.72 ***	9.60 ***	3.73 ***	12.68 ***	6.26 ***	39.91 ***	10.47 ***
Difference	0.83 ***	4.40 ***	26.18 ***	-2.50 ***	-0.99 ***	0.03 ***	-4.18 ***	-17.50 ***	-6.26 ***
Simulated choice probabilities with male entrance probabilities									
Male	2.44 ***	7.41 ***	38.90 ***	7.09 ***	2.74 ***	12.71 ***	2.08 ***	22.41 ***	4.21 ***
Female	1.93 ***	2.89 ***	21.11 ***	12.38 ***	4.00 ***	14.33 ***	7.18 ***	25.02 ***	11.16 ***
Difference	0.51 ***	4.52 ***	17.79 ***	-5.28 ***	-1.26 ***	-1.62 ***	-5.10 ***	-2.61 ***	-6.95 ***
Simulated choice probabilities with female entrance probabilities									
Male	2.43 ***	8.28 ***	28.96 ***	7.57 ***	3.02 ***	13.11 ***	2.43 ***	30.07 ***	4.13 ***
Female	1.62 ***	3.01 ***	12.72 ***	9.60 ***	3.73 ***	12.68 ***	6.26 ***	39.91 ***	10.47 ***
Difference	0.81 ***	5.27 ***	16.24 ***	-2.02 ***	-0.71 ***	0.42 ***	-3.83 ***	-9.84 ***	-6.34 ***

* significant at 10%, ** significant at 5%, *** significant at 1%.

TABLE 17. Marginal effect of gender for interacted variables (Model II)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Average effect	0.79 ***	4.52 ***	24.14 ***	-2.81 ***	-1.00 ***	0.07	-3.41 ***	-16.95 ***	-5.36 ***
	(0.15)	(0.22)	(0.41)	(0.27)	(0.18)	(0.32)	(0.19)	(0.41)	(0.25)
Work	0.95 *	7.20 ***	20.30 ***	-2.69 ***	-1.08 *	-1.73 **	-7.03 ***	-8.10 ***	-7.83 ***
	(0.52)	(0.66)	(0.92)	(0.75)	(0.58)	(0.74)	(0.70)	(0.82)	(0.70)
No Work	0.77 ***	4.02 ***	24.86 ***	-2.83 ***	-0.98 ***	0.40	-2.73 ***	-18.61 ***	-4.89 ***
	(0.14)	(0.22)	(0.45)	(0.28)	(0.18)	(0.35)	(0.18)	(0.46)	(0.27)
Sec Sch Priv	0.64 ***	3.34 ***	25.56 ***	-2.60 ***	-1.02 ***	0.52	-1.96 ***	-19.82 ***	-4.65 ***
	(0.14)	(0.23)	(0.50)	(0.31)	(0.20)	(0.39)	(0.18)	(0.53)	(0.28)
Sec Sch Mixed	0.83	4.12 ***	21.52 ***	-2.06 *	-2.05 **	0.34	-3.89 ***	-12.42 ***	-6.38 ***
	(0.72)	(1.08)	(1.72)	(1.16)	(0.98)	(1.34)	(0.95)	(1.76)	(1.22)
Sec Sch Pub	1.13 ***	7.27 ***	21.34 ***	-3.41 ***	-0.77 **	-1.00 *	-6.64 ***	-11.13 ***	-6.80 ***
	(0.37)	(0.51)	(0.73)	(0.57)	(0.37)	(0.57)	(0.49)	(0.63)	(0.54)
Sec Sch Tech	-0.29	8.06 ***	29.35 ***	-8.85 ***	-1.44 **	-1.98 **	-4.51 ***	-13.45 ***	-6.90 ***
	(0.68)	(0.90)	(1.43)	(1.20)	(0.58)	(0.99)	(0.76)	(1.18)	(0.88)
Sec Sch Not Tech	0.89 ***	4.21 ***	23.68 ***	-2.27 ***	-0.96 ***	0.25	-3.32 ***	-17.26 ***	-5.22 ***
	(0.15)	(0.22)	(0.42)	(0.27)	(0.18)	(0.33)	(0.20)	(0.43)	(0.26)
Other Major	1.55 ***	6.44 ***	17.45 ***	-3.35 ***	-1.83 **	0.38	-3.88 ***	-11.78 ***	-5.00 ***
	(0.55)	(0.86)	(1.20)	(0.86)	(0.75)	(1.06)	(0.88)	(1.19)	(0.80)
No Other Major	0.73 ***	4.35 ***	24.74 ***	-2.76 ***	-0.93 ***	0.04	-3.37 ***	-17.41 ***	-5.39 ***
	(0.15)	(0.23)	(0.43)	(0.28)	(0.18)	(0.33)	(0.19)	(0.43)	(0.27)

Table 17. Marginal effect of gender for interacted variables (Model II, cont.)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Reg Fee	2.50 ***	9.14 ***	15.32 ***	0.89	-0.04	-0.39	-10.89 ***	-7.48 ***	-9.04 ***
	(0.68)	(1.03)	(1.22)	(0.91)	(0.58)	(0.95)	(1.15)	(0.90)	(1.01)
No Reg Fee	0.61 ***	4.02 ***	25.09 ***	-3.21 ***	-1.10 ***	0.12	-2.61 ***	-17.97 ***	-4.96 ***
	(0.14)	(0.22)	(0.43)	(0.28)	(0.19)	(0.34)	(0.17)	(0.44)	(0.26)
Rsn Maj Job Mkt	0.62	1.47 *	28.91 ***	-8.38 ***	0.51 **	-2.61 **	-2.88 ***	-13.21 ***	-4.42 ***
	(0.79)	(0.89)	(1.45)	(1.20)	(0.22)	(1.03)	(0.48)	(1.25)	(0.76)
Rsn Maj Soc Cont	0.59	3.65 ***	8.61 ***	2.85 ***	-0.10	5.87 ***	-2.71 ***	-13.28 ***	-5.48 ***
	(0.44)	(0.48)	(0.87)	(0.71)	(0.34)	(0.94)	(0.65)	(1.22)	(0.72)
Rsn Maj Pers Real	0.25	5.83 ***	23.53 ***	-1.41 ***	-0.72 *	0.49	-3.06 ***	-19.46 ***	-5.46 ***
	(0.24)	(0.44)	(0.83)	(0.51)	(0.39)	(0.60)	(0.35)	(0.81)	(0.52)
Rsn Maj Other	1.53 *	6.54 ***	20.13 ***	-5.44 ***	0.09	-1.63	-5.02 ***	-9.42 ***	-6.79 ***
	(0.83)	(0.82)	(1.37)	(1.05)	(0.53)	(1.07)	(0.78)	(1.35)	(1.01)
Rsn Maj Ability	1.01 ***	4.30 ***	27.67 ***	-3.41 ***	-1.74 ***	-0.73	-3.58 ***	-18.31 ***	-5.22 ***
	(0.17)	(0.30)	(0.59)	(0.36)	(0.27)	(0.44)	(0.25)	(0.58)	(0.34)
Rsn Univ Free	0.56	3.95 ***	21.02 ***	-1.44 **	-0.42	1.42 *	-4.19 ***	-14.10 ***	-6.80 ***
	(0.39)	(0.49)	(0.92)	(0.56)	(0.37)	(0.73)	(0.47)	(0.95)	(0.64)
Rsn Univ Rep	0.33	2.51 ***	23.69 ***	-1.26 ***	-0.54 **	1.55 **	-3.10 ***	-19.00 ***	-4.17 ***
	(0.25)	(0.37)	(0.84)	(0.47)	(0.26)	(0.64)	(0.35)	(0.85)	(0.46)
Rsn Univ Other	0.79 **	5.92 ***	22.10 ***	-0.84	-0.62 *	-0.69	-3.53 ***	-17.27 ***	-5.86 ***
	(0.39)	(0.54)	(0.92)	(0.65)	(0.35)	(0.71)	(0.46)	(0.89)	(0.57)
Rsn Univ Best for Course	1.16 ***	5.32 ***	26.58 ***	-5.11 ***	-1.67 ***	-1.03 **	-3.21 ***	-16.84 ***	-5.22 ***
	(0.21)	(0.34)	(0.62)	(0.44)	(0.33)	(0.46)	(0.27)	(0.61)	(0.38)
Number of observations	39,494								

* significant at 10%, ** significant at 5%, *** significant at 1%.

Appendix A. Model I estimation results for year 2008

Table A1. Marginal effects on the probability of entry, year 2008 (Model I)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Gender	-11.78 ***	-1.96	-0.61	-0.49	-1.03	-1.68 *	-1.23	-0.82 ***	-7.60 ***
	(3.69)	(2.52)	(0.66)	(1.65)	(2.46)	(0.86)	(3.35)	(0.31)	(2.30)
White	1.63	4.36 *	0.20	0.23	4.25	-0.39	0.21	0.21	0.99
	(3.79)	(2.29)	(0.60)	(1.64)	(2.78)	(1.06)	(3.11)	(0.30)	(2.06)
Work	-2.36	-2.13	-1.42 *	-1.19	5.34 *	-1.38	-1.75	0.84	-4.42 *
	(4.18)	(2.63)	(0.85)	(2.00)	(2.94)	(1.28)	(2.99)	(0.68)	(2.48)
Age1	7.88	-19.46 ***	-0.65	-6.27	-21.01 ***	-7.55 **	-17.22 ***	1.26 *	-20.43 ***
	(8.91)	(6.02)	(2.70)	(4.66)	(8.11)	(3.80)	(6.08)	(0.64)	(5.61)
Age2	0.82	-18.26 ***	-0.60	-3.16	-19.33 **	-5.88 *	-8.86 *	1.42 ***	-13.58 ***
	(6.92)	(5.36)	(2.59)	(4.13)	(7.66)	(3.55)	(5.02)	(0.48)	(5.06)
Age3	10.72	-11.78 **	-0.80	-4.57	-14.65 **	-5.61	-2.94	1.24 ***	-8.15
	(6.58)	(5.19)	(2.55)	(4.06)	(7.47)	(3.49)	(4.87)	(0.46)	(4.97)
Prim Sch Priv	-1.82	1.76	2.70 ***	-0.90	0.38	1.85	14.52 ***	0.49	6.25 **
	(5.41)	(3.34)	(0.89)	(2.14)	(3.82)	(1.40)	(4.18)	(0.45)	(2.45)
Prim Sch Mixed	0.37	0.25	1.33	1.37	0.19	0.41	7.02	0.05	2.98
	(5.36)	(3.34)	(0.93)	(2.32)	(3.91)	(1.48)	(4.32)	(0.49)	(2.58)
Sec Sch Priv	3.75	6.65 **	-5.40 ***	-3.94 *	0.18	-4.23 ***	1.61	-1.90 ***	3.76
	(5.31)	(2.85)	(0.99)	(2.10)	(3.59)	(1.54)	(4.37)	(0.60)	(2.40)
Sec Sch Mixed	-16.21 **	1.95	-4.69 ***	-4.54	-2.43	-7.41 ***	-7.32	-3.00 ***	-1.67
	(7.95)	(4.21)	(1.63)	(3.31)	(4.69)	(2.14)	(5.57)	(0.72)	(3.73)
Sec Sch Tech	1.31	-3.14	-0.17	5.74 **	-2.94	2.54	9.07 *	1.80 **	-1.55
	(5.00)	(2.87)	(0.85)	(2.34)	(4.63)	(1.84)	(4.77)	(0.83)	(3.30)

Table A1. Marginal effects on the probability of entry, year 2008 (Model I, cont.)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Other Major	15.68 ***	17.78 ***	3.93 ***	12.54 ***	8.81 *	9.12 ***	18.17 ***	0.48	14.25 ***
	(5.20)	(3.90)	(1.35)	(3.53)	(4.83)	(2.05)	(4.50)	(0.58)	(4.07)
Reg Fee	-27.76 ***	-13.28 ***	-3.32 **	-6.23 **	-1.86	-9.35 ***	-9.89 ***	-0.89	-4.02
	(4.94)	(2.91)	(1.43)	(2.50)	(5.26)	(0.90)	(3.34)	(0.82)	(3.01)
Prep Course	6.11 *	1.74	1.92 ***	4.83 ***	4.64 *	1.72 *	3.68	0.40	3.14 *
	(3.50)	(2.14)	(0.52)	(1.45)	(2.39)	(0.92)	(2.72)	(0.32)	(1.89)
ENEM	0.27 ***	0.37 ***	0.62 ***	0.78 ***	0.31 ***	0.31 ***	0.40 ***	0.40 ***	0.63 ***
	(0.05)	(0.07)	(0.09)	(0.12)	(0.06)	(0.06)	(0.06)	(0.04)	(0.09)
Number of obs.	39,494								

* significant at 10%, ** significant at 5%, *** significant at 1%.

Table A2. Marginal effects on choice probabilities (Model I)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Gender	0.70 ***	3.87 ***	14.67 ***	-2.61 ***	-0.98 ***	1.14 ***	-2.86 ***	-10.12 ***	-3.82 ***
	(0.14)	(0.21)	(0.36)	(0.25)	(0.17)	(0.30)	(0.17)	(0.38)	(0.21)
White	-0.10	-0.25	-1.51 ***	0.55 **	-0.10	0.18	0.01	0.48	0.73 ***
	(0.15)	(0.24)	(0.43)	(0.27)	(0.21)	(0.35)	(0.18)	(0.42)	(0.23)
Work	-0.17	-2.17 ***	23.22 ***	-3.34 ***	-0.26	-3.82 ***	0.96 ***	-13.46 ***	-0.95 ***
	(0.16)	(0.23)	(0.72)	(0.29)	(0.36)	(0.44)	(0.27)	(0.64)	(0.30)
Prim Sch Priv	-0.44 **	0.00	0.45	0.13	-0.07	0.40	-0.77 **	2.36 ***	-2.07 ***
	(0.21)	(0.31)	(0.60)	(0.34)	(0.28)	(0.47)	(0.34)	(0.58)	(0.35)
Prim Sch Mixed	-0.46 **	0.38	-0.94	0.33	0.27	1.09 **	-0.74 **	1.37 **	-1.29 ***
	(0.21)	(0.33)	(0.65)	(0.37)	(0.30)	(0.51)	(0.31)	(0.63)	(0.37)
Sec Sch Priv	-6.15 ***	-5.30 ***	20.73 ***	-2.67 ***	-1.75 ***	-4.62 ***	-1.05	-0.05	0.85
	(1.30)	(1.00)	(1.61)	(0.53)	(0.56)	(0.87)	(0.80)	(0.18)	(0.54)
Sec Sch Mixed	-9.66 ***	-4.49 ***	16.83 ***	-1.79 ***	-1.04 *	-3.46 ***	-2.31 *	5.27 ***	0.66
	(2.46)	(1.60)	(2.20)	(0.66)	(0.62)	(1.10)	(1.21)	(0.80)	(0.67)
Sec Sch Tech	-0.56 ***	-2.29 ***	25.45 ***	-2.21 ***	-1.43 ***	-6.37 ***	-1.02 ***	-9.52 ***	-2.07 ***
	(0.19)	(0.28)	(0.74)	(0.35)	(0.26)	(0.43)	(0.26)	(0.72)	(0.34)
Other Major	-1.26 ***	-2.25 ***	21.98 ***	-3.70 ***	-2.05 ***	-1.52 **	-0.53 *	-8.27 ***	-2.40 ***
	(0.15)	(0.33)	(0.94)	(0.38)	(0.20)	(0.68)	(0.29)	(0.78)	(0.32)
Reg Fee	-0.74 ***	-1.88 ***	23.10 ***	-4.76 ***	-2.15 ***	-5.42 ***	8.81 ***	-17.88 ***	0.90
	(0.14)	(0.33)	(0.81)	(0.34)	(0.21)	(0.56)	(1.11)	(0.73)	(0.66)
Prof Father Non-manual	0.35 *	0.78 ***	0.24	0.69 **	0.47 **	-0.70	-0.07	-1.36 ***	-0.39
	(0.19)	(0.29)	(0.51)	(0.31)	(0.23)	(0.44)	(0.23)	(0.51)	(0.27)
Prof Father Manual	0.56 **	0.76 *	1.41	1.00 **	0.51	0.10	-0.44	-3.51 ***	-0.40
	(0.26)	(0.44)	(0.96)	(0.51)	(0.38)	(0.65)	(0.32)	(0.85)	(0.40)
Prof Father Other	0.42 *	1.05 ***	-0.37	1.94 ***	0.07	-0.59	0.18	-2.35 ***	-0.35
	(0.23)	(0.40)	(0.71)	(0.45)	(0.34)	(0.57)	(0.32)	(0.69)	(0.36)

TABLE A1. Marginal effects on choice probabilities (Model I, cont.)

	Technologies	Exact Sciences	Engineering and Arch.	Natural and Earth Sc.	Arts	Social Sciences	Humanities	Health and Biol. Sc.	Other Health and Biol. Sc.
Prof Mother Non-manual	-0.26	0.22	-0.01	0.36	0.36	-0.19	0.14	-1.33 ***	0.71 ***
	(0.19)	(0.27)	(0.47)	(0.32)	(0.22)	(0.39)	(0.24)	(0.48)	(0.27)
Prof Mother Manual	-0.13	-0.12	2.14 *	0.77	-0.69 *	-1.02	0.26	-1.17	-0.05
	(0.37)	(0.57)	(1.20)	(0.65)	(0.41)	(1.08)	(0.48)	(1.18)	(0.57)
Prof Mother Housewife	0.07	0.68 **	1.76 ***	0.75 **	-0.56 ***	-0.84 **	-0.19	-2.04 ***	0.37
	(0.22)	(0.30)	(0.53)	(0.35)	(0.21)	(0.43)	(0.26)	(0.52)	(0.29)
Prof Mother Other	-0.16	0.40	2.04 ***	0.97 **	-0.20	-0.62	0.02	-3.04 ***	0.58
	(0.23)	(0.40)	(0.71)	(0.47)	(0.30)	(0.58)	(0.32)	(0.69)	(0.39)
Rsn Maj Job Mkt	2.65 ***	1.69 ***	2.99 ***	4.13 ***	-3.32 ***	-0.11	-1.87 ***	-4.86 ***	-1.30 ***
	(0.34)	(0.45)	(0.75)	(0.55)	(0.30)	(0.58)	(0.26)	(0.68)	(0.37)
Rsn Maj Soc Cont	0.24	-2.81 ***	-12.52 ***	-0.34	-1.79 ***	2.07 ***	0.70 **	14.04 ***	0.42
	(0.21)	(0.28)	(0.61)	(0.40)	(0.26)	(0.53)	(0.27)	(0.70)	(0.34)
Rsn Maj Pers Real	-0.32 **	-0.35	-3.22 ***	0.09	0.22	-0.29	-0.41 **	3.27 ***	1.01 ***
	(0.13)	(0.25)	(0.43)	(0.28)	(0.20)	(0.35)	(0.19)	(0.43)	(0.25)
Rsn Maj Other	3.06 ***	0.47	-4.34 ***	2.16 ***	-1.32 ***	-0.18	0.44	-2.11 ***	1.83 ***
	(0.37)	(0.42)	(0.71)	(0.58)	(0.31)	(0.60)	(0.34)	(0.71)	(0.47)
Rsn Univ Free	0.32 *	-0.99 ***	-2.02 ***	-2.11 ***	-1.31 ***	1.42 ***	0.29	3.35 ***	1.06 ***
	(0.18)	(0.26)	(0.49)	(0.32)	(0.21)	(0.41)	(0.21)	(0.50)	(0.30)
Rsn Univ Rep	0.03	-0.89 ***	0.83 **	-1.37 ***	-1.72 ***	1.18 ***	0.63 ***	1.55 ***	-0.23
	(0.16)	(0.25)	(0.42)	(0.29)	(0.19)	(0.35)	(0.21)	(0.42)	(0.25)
Rsn Univ Other	0.71 ***	0.65 **	-0.05	-0.28	-1.47 ***	0.62	0.38 *	-0.32	-0.24
	(0.19)	(0.29)	(0.50)	(0.33)	(0.22)	(0.41)	(0.22)	(0.50)	(0.28)
Number of observations	39,494								

* significant at 10%, ** significant at 5%, *** significant at 1%.