

## **What is behind University Dropout Decision in Brazil? A Bivariate Probability Model**

**Juliana Guimarães**

*Newnham College, University of Cambridge, U.K.*

Email: jgc29@cam.ac.uk

**Breno Sampaio\***

*Department of Economics, University of Illinois at Urbana Champaign, USA*

**Yony Sampaio**

*Department of Economics, Universidade Federal de Pernambuco*

Email: sampyony@yahoo.com.br

**Abstract:** In this paper we analyze the major determinants of university enrollment and dropout in Brazil. The econometric model consists of two simultaneous equations. The first equation determines whether the student was accepted at the university or not; the second determines students decision to dropout or not from higher education, given they were accepted at the university. Gender, age and marriage influence not only acceptance but also the decision to dropout. As usual family background is also an important determinant of both, and in that matter, family income plays a major role on student's performance and their decision to dropout.

**Keywords:** College Enrollment, Dropout, Brazil

**JEL Classification Number:** I21, J24.

### **1. Introduction**

Dropout from post-secondary schooling is nowadays considered a social problem. It is in fact a major concern in many countries and could no be different at the Brazilian public universities. It is a two-fold problem. It represents the end of students' aspirations to acquire a degree and the corresponding possible repercussions of lower future income, but is also a direct waste of universities resources. Thus, a better understanding of the dropout decision process is crucial for institutions and policy makers in developing specific intervention strategies seeking its reduction.

The phenomenon of dropout at Brazilian public universities is different in many aspects. As public universities in Brazil charge no tuition, students can enroll for free even if they do not attend classes. It is clear that this "fake" enrollment presents a similar cost to the

---

\* Corresponding Author. Email: sampaio2@illinois.edu

institution whether the student drops out or not. As a consequence, two different situations arise: a student may recognize a wrong career choice<sup>1</sup> or the impossibility of attending the university and drops out to apply to another course<sup>2</sup>; or a student may enter in the labor market but continues to enroll occupying a place even if he/she cannot attend classes. If tuition is charged, enrolling has a cost and the student will probably opt to drop out so the two situations end up being similar. This distinction in the institutional setting may change the sign of some key variables found in previous studies that affect the decision to dropout.

For instance, previous literature shows that the dropout decreases as parents' income increase, what is explained, among other causes, by the need of lower income students to take part in the labor market (Bradley and Lenton (2007), Lassibille and Gomez (2008) and Di Pietro (2004)). But it might be that when students are not satisfied with their course choice and have the possibility of applying to another one, dropout is less costly for higher income students who can take a gap year studying and also pay for extra tutoring classes to increase their chances of succeeding in another entrance exam (*vestibular*). In this case, for lower income students, one may argue that it may become more difficult to decide to dropout and try to submit to another entrance exam. Several papers show also that the higher the prior academic performance and the entrance test score, the lower the probability of dropping out (Arulampalam et al., 2004; Lassibile and Gomez, 2008; Murtaugh, et al. 1999; Bradley and Lenton, 2007). The reason for this is that prior performance increases the probability of success (and of returns) and decreases the risk and cost of failure (Eckstein and Wolpin, 1999; Light and Strayer, 2000; Bishop and Mane, 2001). However, when there is no tuition, the sign of this variable may change and in fact be the opposite of previous results. Students with higher grades may easily transfer (given their higher level of human capital accumulation) between courses by applying for the university again.

Given the institutional differences of the Brazilian educational system, the main objective of this paper is to analyze the major determinants of university enrollment and dropout in Brazil and compare with previous results obtained in similar studies. Our results show that gender, age and marriage influence not only acceptance but also the decision to dropout. Contrary to results presented in previous studies, family income is positively correlated with the probability to dropout, which is an expected result given how college admissions are organized in Brazil. Also, there is a positive correlation between entrance test scores

---

<sup>1</sup> An important institutional difference between Brazilian and American universities, for example, is that all students applying for higher education are required to choose their major before they take the entrance exam. Thus, given a choice, students are not allowed to switch majors when enrolled at the university.

<sup>2</sup> If this is the case, the student who decided to dropout must take the university entrance exam again which happens only once a year.

and dropout, which arises from the fact that high ability students have a lower cost of switching majors given they anticipate their higher probability of being accepted for higher education.

The paper is organized as follows. Next section presents the methodology and in the third section describes the data. The fourth section presents the results and in the fifth section concludes.

**2. Methodology**

The econometric model<sup>3</sup> consists of two simultaneous equations. The first determines whether the student was accepted at the university or not ( $y_{1i}$ ); the second determines students' decision to dropout or not from higher education, given acceptance at university ( $y_{2i}$ ). Formally speaking, for each  $i = 1 \dots N$ , let  $y_{1i}^*$  and  $y_{2i}^*$  be unobserved latent variables such that:

$$\begin{aligned}
 y_{1i}^* &= \beta_1 x_{1i} + \varepsilon_{1i} \\
 y_{1i} &= \begin{cases} 1, & \text{if accepted at university } (y_{1i}^* > 0) \\ 0, & \text{otherwise} \end{cases} \\
 y_{2i}^* &= \beta_2 x_{2i} + \varepsilon_{2i} \\
 y_{2i} &= \begin{cases} 1, & \text{if accepted at university } (y_{2i}^* > 0) \\ 0, & \text{otherwise} \end{cases}
 \end{aligned}$$

where  $x_{ji}$ ,  $j=1,2$ , are vectors of explanatory variables and the unobserved components  $\varepsilon_{ji}$ ,  $j=1,2$ , are assumed to be mean-zero, bivariate normally distributed with unit variances and correlated across individuals with correlation coefficient  $\rho$ . The correlation coefficient is important since we only observe dropout data for students who did enroll at the university, which are non-randomly selected from the total set of students. Hence  $\varepsilon_{1i}$  is defined over the population of students who are competing for a university place, while  $\varepsilon_{2i}$  is defined only for the subpopulation where  $y_{1i} = 1$ .

Three categories of observations arise: students who did not enroll at the university; students who did enroll at the university and decided not to dropout; and students who did enroll but decided to dropout. In terms of unconditional probabilities, we have

$$\begin{aligned}
 y_{1i} = 0 : \Pr(y_{1i} = 0) &= \Phi[-\beta_1 x_{1i}] \\
 y_{1i} = 1 \begin{cases} y_{2i} = 0 : \Pr(y_{1i} = 1, y_{2i} = 0) &= \Phi_2[\beta_1 x_{1i}, -\beta_2 x_{2i}, -\rho] \\ y_{2i} = 1 : \Pr(y_{1i} = 1, y_{2i} = 1) &= \Phi_2[\beta_1 x_{1i}, \beta_2 x_{2i}, \rho] \end{cases}
 \end{aligned}$$

---

<sup>3</sup> Here we follow the description provided by Di Pietro (2004) and Montmarquette et al (2001).

where  $\Phi$  and  $\Phi_2$  denote the univariate and bivariate standard normal cumulative distribution functions, respectively, and the corresponding log-likelihood function, which is maximized with respect to the parameters  $\beta_1, \beta_2$ , and  $\rho$ , is:

$$\sum_{y_{1i}=1, y_{2i}=1} \ln \{ \Phi_2 [\beta_1 x_{1i}, \beta_2 x_{2i}, \rho] \} + \sum_{y_{1i}=1, y_{2i}=0} \ln \{ \Phi_2 [\beta_1 x_{1i}, -\beta_2 x_{2i}, -\rho] \} + \sum_{y_{1i}=0} \ln \{ \Phi [-\beta_1 x_{1i}] \}$$

### 3. Data

We use a unique data set provided by the Universidade Federal de Pernambuco, which is the major University in Northeast of Brazil. Table 1 presents summary statistics. Among the students accepted at the university, the ones who decided to dropout have lower Entrance Test Score (ETS) and lower GPA. Also, the percentage of students reporting themselves as certain about their career choices<sup>4</sup> is higher among the students who did not dropout. With respect to personal characteristics, among the students who were not accepted for a position at the university there is a slightly higher percentage of females. Among the ones who enrolled, males are clearly the ones most likely to dropout. With respect to marital status, it's worth noticing that the percentage of married students among the ones who dropped out is almost twice as large as the percentage of married students who did not dropped.

When it comes to educational system, there is a large difference in the proportion of students coming from public schools between the ones accepted and not accepted in the university. However, for those accepted, coming from a public school seems not to be affecting the likelihood to dropout. Another important difference between accepted and non-accepted students comes from their family income. Among the students accepted almost 22% had family income above 16 minimum wages, whereas only 14% of the non accepted students presented this level of family income.

One important variable in our identification strategy is if the student attended private classes or not before doing the entrance exam. This variable directly affects the entrance exam and, therefore, the likelihood of being accepted for higher education. However, given students' income, it should not have any significant relation to students' decision to dropout. Thus, we include this variable in the first stage as an exogenous variation in students' probability to enroll, but we do not include in the second stage.

---

<sup>4</sup> Students were queried on how sure they were about their career choice.

**Table 1: Summary Statistics**

Variables	Students Applying for University		Students Enrolled at University	
	Students not accepted (N = 74,351)	Students accepted (N =7,582)	Students who did not dropout (N=5,288)	Students who did dropout (N=2,294)
ETS			6.138 (1.03)	5.952 (.909)
1 <sup>st</sup> Semester GPA			7.804 (1.08)	7.124 (1.93)
Decided about Career			.861 (.346)	.835 (.371)
Female	.570 (.495)	.503 (.500)	.550 (.498)	.395 (.489)
Age	20.61 (5.49)	20.06 (4.19)	19.81 (3.96)	20.64 (4.65)
Married	.065 (.246)	.049 (.215)	.038 (.192)	.072 (.259)
Married X Female	.033 (.178)	.019 (.137)	.020 (.139)	.017 (.131)
Metropolitan Area	.857 (.350)	.909 (.287)	.901 (.298)	.928 (.259)
<b>Tests Taken</b>				
0	.527(.495)	.318 (.466)	.319 (.466)	.316 (.465)
1	.237 (.425)	.324 (.468)	.331 (.470)	.310 (.463)
2	.129 (.335)	.209 (.407)	.207 (.405)	.215 (.411)
3	.059 (.236)	.089 (.284)	.088 (.284)	.090 (.286)
>4	.045 (.208)	.058 (.235)	.054 (.227)	.068 (.251)
<b>Public School</b>				
Primary School	.261 (.439)	.157 (.364)	.157 (.364)	.156 (.363)
High School	.322 (.467)	.249 (.432)	.249 (.433)	.247 (.431)
<b>Parents Education</b>				
Illiterate	.089 (.285)	.049 (.215)	.052 (.223)	.040 (.195)
Primary School	.142 (.349)	.088 (.283)	.086 (.281)	.092 (.288)
High School	.362 (.481)	.329 (.470)	.336 (.472)	.313 (.464)
College Degree	.403 (.491)	.531 (.499)	.522 (.500)	.552 (.497)
<b>Family Income</b>				
<5	.411 (.492)	.274 (.446)	.283 (.451)	.254 (.435)
5-10	.332 (.471)	.366 (.482)	.371 (.483)	.356 (.479)
11-15	.107 (.309)	.143 (.351)	.141 (.348)	.149 (.356)
16-20	.062 (.241)	.093 (.291)	.091 (.288)	.097 (.296)
>20	.080 (.272)	.118 (.323)	.109 (.311)	.140 (.347)
Currently Enrolled at HE	.116 (.320)	.169 (.375)	.129 (.335)	.262 (.440)
Private Classes	.546 (.498)	.637 (.481)		

Note: Standard Deviation presented in parenthesis.

#### **4. Results**

Table 2 presents our estimates. The first column brings the results for the drop out decision and the second column deals with the determinants of enrollment decision. The two most important results, contrasting to prior estimates, are that dropout is positively correlated with income and with higher entrance test scores. As posed in the introduction, this particular result is a consequence of a system that charges no tuition to its students, which, upon the recognition of a bad career choice, stimulates the wealthier (given the high costs implied by this decision) and the best (given less able students expect to be accepted for higher education again with smaller probability when compared to the more able students) students to dropout to try another career. We consider this as a very interesting result because of its policy implication for Brazilian universities.

Other results show that females have a lower probability to enroll at the university compared to male students; however, they also have a lower probability of dropping out of the University once they are in. Older students have lower probability of being accepted but higher probability of dropping out. Both results are obtained in the literature however, they might be driven by unobserved variables, such as ability/motivation, since less able students are more likely to repeat grades during primary/secondary education and, as a consequence, are more likely to enter college older than more able students (Sampaio, 2009). Marriage do increase the probability of enrolling, however it also increases the probability of dropout. An interaction of marriage and female indicators tell us that this result is the opposite for females students. Married girls tend to have a lower probability of enrolling, however once they get in they tend not to drop out.

The type school attended has an influence on enrollment but not on the drop out decision. Students that attended public schools (primary and secondary) do have a lower probability of getting in. However, we cannot detect any effect from type of school on dropout decision. Also, students that are currently enrolled in another higher education institution have a higher probability of entering the University, but also, a higher probability of dropping out. Here, clearly the fact the student already have an option, does play a key role in the drop out decision.

**Table 2: Bivariate Probit MLE estimation**

Variables	Prob. to Dropout	Prob. to Enroll	Variables	Prob. to Dropout	Prob. to Enroll
ETS	.176***		<b>Parents Education</b>		
1 <sup>st</sup> .Semester.GPA	-.220***		Primary.School	.139*	-.049
Decided.about.Career	-.080**		High.School	.136*	.001
Female	-.129**	-.049***	College.Degree	.249***	.104***
Age	.052	-.077***	<b>Family Income</b>		
Age <sup>2</sup>	-.001**	.001***	5-10	.044	.149***
Married	.373***	.143***	11-15	.114*	.219***
Married.X.Female	-.367***	-.179***	16-20	.149**	.275***
Metropolitan.Area	.234***	.256***	>20	.299***	.276***
<b>Tests Taken</b>			Currently Enrolled at HE	.369***	.051***
1	.171*	.474***	Private Classes		.136***
2	.231*	.655***	Fall Entry	YES	
3	.171	.688***	Year Fixed Effects	YES	YES
>4	.207	.712***	Major Fixed Effects	YES	YES
<b>Public.School</b>			Constant	-2.076***	-.848***
Primary.School	-.085	-.165***	Constant	-2.076***	-.848***
High.School	.039	.103***			

Note:  $\rho = .673^{***}$  (.182), Log.Likelihood=-27,399.36, N=81,933. \*\*\* indicates  $p < 0.01$ , \*\*  $p < 0.05$  and \*  $p < 0.10$ .

**5. Conclusions**

This paper uses a bivariate probability model to analyze the major determinants of university enrollment and dropout in Brazil. Given the institutional differences of the Brazilian educational system, this paper has a major contribution of highlighting that policies based on results found by American and European researchers could lead to very different outcomes. Our results show that, contrary to conventional wisdom, family income is positively correlated with the probability to dropout. Also, there is a positive correlation between entrance test scores and dropout, which arises from the fact that high ability students, given the recognition of a wrong career choice, have a lower cost implicitly defined by their higher probability of being accepted for a different major on next years' exam. Also, gender and marriage play key roles in both the probability to enroll and to dropout. In terms of academic variables, the types of school attended have a higher impact on the enrollment outcome, but not necessarily on the drop out decision.

**References**

- Arulampalam, W., Naylor, R., Smith, J., 2001, A hazard model of the probability of medical school drop-out in the UK, *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 167 (1), 157-178.
- Bishop, J., Mane, F., 2001, The impacts of minimum competency exam graduation requirements on high school graduation, college attendance and early labor market success, *Labour Economics*, 8 (2), 203-222.
- Bradley, S., Lenton, P., 2007, Dropping out of post-compulsory education in the UK: an analysis of determinants and outcomes, *Journal of Population Economics*, 20, 299-328.
- Di Pietro, G., 2004, The determinants of university dropout in Italy: a bivariate probability model with sample selection, *Applied Economics Letters*, 11, 187- 191.
- Eckstein, Z., Wolpin, K., 1999, Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities, *Econometrica*, 67 (6), 1295- 1339.
- Lassibille, G., Gómez, L., 2008, Why do higher education students drop out? Evidence from Spain, *Education Economics*, 16 (1), 89-105.
- Light, A., Strayer, W., 2000, Determinants of College Completion: School Quality of Student Ability?, *Journal of Human Resources*, 35 (2), 299-332.
- Montmarquette, C., Mahseredjian, S., Houle, R., 2001, The determinants of university dropouts: a bivariate probability model with sample selection, *Economics of Education Review*, 20, 475-484.
- Murtaught, P., Burns, L., Schuster, J., 1999, Predicting the Retention of University Students, *Research in Higher Education*, 40 (3), 355-371.
- Sampaio, B., 2009, Too Late to Change? The effect of age on college students' dropout. Mimeo, University of Illinois.