

---

# Leaving Early: The Determinants of Student Non-Completion in Croatian Higher Education

MAJA MIHALJEVIĆ KOSOR\*

Faculty of Economics, University of Split  
Split, Croatia

Izvorni znanstveni rad  
UDK: 37.011.3-052(497.5)  
doi: 10.3935/rsp.v17i2.913  
Priljeno: studeni 2009.

*Similar to other European higher education systems, the Croatian higher education framework has experienced some remarkable changes since the 1990s as the number of students enrolled increased by more than 80 percent in the last 15 years. Following this expansion, in many countries there has been an increase in non-completion rates, especially during the first year of studies. To gain more insight, this paper presents the theoretical background for examining student non-completion at the first year of studies and identifies the main determinants. A model of student non-completion is developed and adjusted to the Croatian context. Using a rich dataset on eight cohorts of one large Croatian higher education institution, the empirical work analyses the effects of student's personal characteristics, his/her peers, parental qualifications, previous and current schooling characteristics on the probability of non-completion of the first year of studies. This research is also relevant and applicable in other countries experiencing these inefficiencies in student non-completion.*

**Key words:** Higher education, First-year non-completion, Student Integration Model, Logistic regression.

## INTRODUCTION

The goal of this research is to increase knowledge on the determinants of student non-completion focusing on early leavers in Croatian higher education (HE). There are three broad categories of students that can be labelled as early leavers in this research. These students may be the ones who withdrew from the higher education institution (HEI) before the end of their first year of study; who failed the first year at the HEI

or those who successfully completed their exams, but did not return for the second year. The reasons for focusing on students leaving before completing their first year of studies are twofold. Firstly, relevant research in the US and UK on non-completion revealed that most of the student non-completion occurs during the first year of study (e.g. Tinto, 1982, 1993 for the US; and Smith and Naylor, 2001 and Arulampalam et al., 2005 for UK). This year has also been referred to as the »make or break year« in

---

\* Maja Mihaljević Kosor, Ekonomski fakultet / Faculty of Economics, Matice hrvatske 31, 21 000 Split, Hrvatska/Croatia, majam@efst.hr

The author would like to thank Nick Adnett, Jean Mangan and Geoff Pugh, Staffordshire University, for useful comments and suggestions on an earlier draft of this paper, as well as three anonymous referees.

HE. In Croatian HE, although the number of students enrolled increased by 80 percent in the last 15 years (Central Bureau of Statistics, 2008), only about two thirds of the enrolled students eventually, over a longer period of time, completed their studies and obtained degrees (Šošić, 2004, Babić et al., 2006, Matković, 2009). Secondly, it is necessary to distinguish between non-completion of the first year and at a different time in the programme, as the determinants for first year non-completion may differ from those on later non-completion (Arulampalam et al., 2005). From here, a model of student non-completion in the first year of studies is developed. This is also the first research of this kind for Croatian HE.

Analysis in this area is of significant policy interest, as non-completion may have an adverse effect on the efficiency and effectiveness of a higher education (HE) system. Furthermore, due to the greater access into HE there is already a widespread concern over the rising rates of non-completion identified in European and US education systems. In many OECD countries enrolment rates to HE more than doubled during the last thirty years. Following this expansion, in many countries there has been an increase in non-completion rates. This led some of the recent reports on the future of the European HE systems to stress the need for its reform and to specifically highlight student non-completion as problematic (in Aghion et al., 2008; van der Ploeg and Veugelers, 2008).

This paper is organised as follows: Section 2 presents an overview of the development of the theory of student non-completion, highlighting central features of the models pertinent to non-completion research, and indicating some of the limitations of those models. A critical examination of previous empirical work follows in section 3, with a focus on the more widely used variables contributing to an increased

probability of non-completion. In section 4, a model is developed and in section 5 it is used in estimating the probability of non-completion of the first year of studies for one large Croatian HEI. A discussion of the wider policy context and general conclusions finally follow in section 6.

## **THE THEORY OF STUDENT NON-COMPLETION AND ITS APPLICATION**

In this section theoretical studies and conceptual frameworks dealing with non-completion are considered. An overview is provided focusing on two strands: Tinto's (1975) theory and model of non-completion from the education literature, which is the most widely used model, and Becker's (1964) and Stratton et al.'s (2004) model from the human capital theory. These conceptual frameworks help in identifying variables likely to influence student non-completion for the empirical part of this research.

One of the most extensive theoretical studies of non-completion in tertiary education is that by Tinto who developed a Student Integration Model (1975), which was further extended in 1988 and 1993. Tinto (1975) relates dropout to a student failing to socially and academically integrate at university, i.e. »... the process of dropout from college can be viewed as a longitudinal process of interactions between the individual and the academic and social systems of the college during which a person's experiences in those systems (as measured by his normative and structural integration) continually modify his goal and institutional commitments in ways which lead to persistence and/or to varying forms of dropout« (in Woodley, 2003). Therefore, the explanation in Tinto's Student Integration Model (SIM) of non-completion or persistence depends on the quality of the match between the student and the institution. In

the model, students are viewed as entering the HEI with certain personal characteristics and with their goals, commitments and intentions. Attributes such as age, gender, family characteristics, student ability and prior schooling contribute to student's goals and commitments. Moreover, goals and commitments are continually modified as a student progresses through his/her studies leading to persistence or to non-completion where student's academic and social experiences will either help the student to integrate into the HEI and to persist, or the opposite may happen and the student will not complete the programme. Integration in the SIM depends on two factors: the match between the student and the HEI and the social interactions between the student and others at the institution (other students and the HEI staff). In the SIM model, academic integration is primarily determined by the student's academic performance and his/her level of intellectual development, whereas social integration is primarily a function of the extent and quality of peer group interactions and the student's interaction with the HEI's staff (Tinto, 1975). According to Tinto, the higher the level of integration, the more likely is the student to persist.

The application of the SIM model was central to a number of studies (Pascarella and Terenzini, 1991; Johnes and McNabb, 2004), and most of these studies supported the validity of the major variables in the model. However, none of the research papers employed the whole range of explanatory variables that Tinto (1993) hypothesised about, nor did they examine their salient features. The main criticism directed at the Student Integration Model is that it is putting too strong an emphasis on social integration with only a modest stress on the

personal characteristics of students. Also, it is mostly focusing on full-time residential students in the United States who have just recently graduated from school, thus making it complicated to apply the model to e.g. adult learners, foreign students and distant learners, or to education systems in other countries since the variables and functional relationships may require a somewhat different specification (Towles and Spencer, 1993; Yorke, 1999; Braxton, 2000).

The next major model of student non-completion was developed by Bean and Metzner (1985). These authors argue that the above models relied too heavily on the socialisation aspect to explain non-completion. A supplementary theory was necessary to explain non-completion for non-traditional students (i.e. married, parents, older, working and part time students) who did not have the opportunity to become socially integrated in the HE community. Bean emphasises the role of intentions (i.e. factors external to the HEI), while Tinto's model is mostly focused on the match between the student and the institution. Bean and Metzner (1985) assume that non-traditional students are more affected by environmental factors that are outside the academic environment (e.g. working hours, family obligations, finances, opportunity to transfer), than by integration into the academic environment. In their model variables that are assumed to have the greatest effect on non-completion for non-traditional students are academic performance, intent to leave<sup>1</sup>, personal characteristics, secondary school attainment, educational goals and environmental variables (in Summers, 2003). A similar model from the education literature is by Cabrera et al. (1993) who combine Tinto's (1980) Student Integration Model

---

<sup>1</sup> It may be argued that this variable is non-explanatory, i.e. it cannot be used in predictions as it is practically an outcome of the process and not an explanatory variable.

with Bean's (1980) Student Attrition Model into a Model of Student Retention. In the Model of Student Retention the most important factors influencing persistence in HE are the student's intention to persist, student's grade point average at the HEI and the institutional commitment to the student (Cabrera et al., 1993).

Following that, the human capital literature is discussed and additional explanation of student non-completion is presented. The major contribution in applying human capital framework to non-completion comes from Stratton et al. (2004, 2008). Stratton et al. (2008) refer to their model as the »human capital model of non-completion«. However, as will be emphasized in section 4, the variables used in Stratton et al. (2008) are already pertinent to the research carried out in the education literature, and found in Tinto (1975, 1993), Bean and Metzner (1985) and Cabrera et al. (1993) along with the empirical studies stemming from these models. Hence, this model is not completely distinguishable from other models in practice.

According to the model of non-completion based on the human capital theory, students will persist in the HEI as long as the present value of expected benefits exceeds the present value of expected costs. The cost and benefits in pursuing a degree can be both pecuniary and non-pecuniary. The major financial benefits may be higher future earnings after completing the programme. Non-financial benefits may include better working conditions and a more satisfying job. Furthermore, education institutions impart good standards of behaviour, socialise people and enable them to become better informed members of the society. Financial costs include tuition fees and forgone earnings, and non-financial costs may be linked to psychological costs of studying, i.e. stress, anxiety, alienation, more effort and less leisure.

In the »human capital model of non-completion« applied by Stratton et al. (2008), although the authors stress income and constraints rather than sociological reasons, their analysis does not seem to lead to a model specification that would employ different variables than the ones previously known in the non-completion theory and practice. More on this model is presented in the next section.

### **A CRITICAL EXAMINATION OF PREVIOUS EMPIRICAL WORK**

Major contributions in theory and estimation of non-completion come from the US and, to a smaller extent, the UK. These education systems are different in comparison to the Croatian HE system and this may limit the applicability of the major findings to our HE system. However, it may be argued that there is sufficient commonality in the major influences on non-completion that can be extended to HE systems in Croatia and other countries. Such a framework is developed in section 4, taking into account the characteristics of the Croatian HE system.

The first study analysed in this section is an assessment of non-completion of the first year of study for the US HE, where Stratton et al. (2008) use a two stage sequential decision logit model, i.e. a switching regression, to determine whether the factors associated with non-completion differ according to initial enrolment intensity (i.e. part time and full time status). The authors specify three equations, where the first equation models the decision to enrol part-time and it is a function of variables known prior to enrolment. The second equation models the decision to leave HE studies conditional upon having enrolled full-time and the third equation models the decision to leave conditional upon having enrolled part-time. Students who are not enrolled one calendar year after their initial enrolment are defined

as non-completers, i.e. the focus of the empirical work is on non-completion of the first year of studies.

The authors find that academic performance, socio-economic background, parental education and economic factors (living in an area with a high unemployment rate) had a significant effect on students who were initially enrolled full-time, whilst racial and ethnic characteristics had a significant effect on part time students. However, the variables the authors use do not include peer effects, which are well established determinants of student non-completion in the economics of education literature. Peer effects are perceived as a group of influences arising largely from social interactions, in which the behaviour of one individual is affected by the behaviour or characteristics of other individuals in the same group. Variables capturing the characteristics of prior schooling and prior ability are also missing. It may be argued that the lack of the above mentioned variables seriously impairs the validity of the estimates. Furthermore, as a measure of student's ability the authors are using the first year grades at the HEI which are self-reported by the students in cases where that data was not available. It may be argued there may be a bias when combining the HEI awarded grades and the self reported ones, as the latter may not be accurately reported. Another concern is the size of the sample of part-time students, i.e. only about 7.5 percent of the students were initially enrolled part-time. However, Stratton et al. (2008) do not report if any tests were made to check, for example, the existence of outliers that might bias the results.

For UK HE, Smith and Naylor (2001) examine the probability of non-completion for the entire 1989/90 entry cohort of full-time students in the »old« universities. In their dataset, these students had either completed their studies at the end of 3 or 4 years of study or left the programme prior

to completion. The last data is available for 1993 and the model is estimated using a probit model. Smith and Naylor (2001) find that the probability of non-completion is influenced significantly by prior schooling, personal characteristics and the characteristics of the department and the university. Another limitation of the model is again the lack of a variable for peer effects. This variable is usually constructed as student's individual ability (proxied by A-levels in the UK HE system) relative to the ability of student's peers at the same university or course (also A-levels).

Using the same dataset, Arulampalam et al. (2005) investigate the probability of student non-completion of the first year of study using the data for nine cohorts (1984/85-1992/93) of full-time undergraduate students in the »old« universities in the UK. The authors use the logit model to analyse the determinants of non-completion probabilities during the first year of studies only and focus on the results for the 1992/93 entry cohort which is also the last one for which the data is available. The variables included are related to student's personal characteristics, prior schooling, socio-economic background, enrolment status, course and university characteristics and peer effects. Their results show that students with lower prior attainment (also based on A-level score) are more likely not to complete their first year. The probability of non-completion for students with higher and lower prior attainment increases with the extent of in-class heterogeneity (determined also by prior performance), where the median ranked students are less at risk of non-completion.

Another related study for UK HE is that by Johnes and McNabb (2004) who investigate student non-completion in 1993. Disaggregating the data to the level of the individual full-time student and matching it with institutional information, the authors



were able to distinguish between the two types of non-completion, voluntary and involuntary, and include it in their model. The estimation technique used is the multinomial logit where the dependent variable captures one of three outcomes: completion of studies, voluntary withdrawal and involuntary withdrawal. Other variables used are student's personal characteristics (age, marital status, nationality, residency), type of secondary school attended, prior attainment (A-levels or Scottish Highers), a variable capturing peer effects and gender specific peer effects, degree subject and the characteristics of the university attended (university income from research grants, library expenditures per student, staff-student ratio, teaching and research quality assessment). In terms of peer effects, this variable is constructed as student's individual ability (proxied by student's A-level score) relative to the mean ability of students at the same university and on the same course. A squared term of the peer effects is also included to address the issue of non-linearities. The authors find that peer effects have a significant effect on non-completion and a student is more likely to leave HE studies if his/her prior grades are better relative to those of other students at his/her university or degree course. This is a surprising finding and at odds with the theoretical basis. The squared peer effects term was however insignificant. Furthermore, the probability of non-completion (both voluntary and involuntary) is negatively related to the performance prior to university entry (A-levels or Scottish Highers), i.e. prior attainment. The distinction made in this study between the voluntary and involuntary withdrawal from the theoretical point of view may be unclear. Involuntary withdrawal (i.e. often considered as academic failure) may sometimes be the result of personal and financial problems. In this case, student non-completion may be an outcome of the above mentioned factors rather than a lack of ability.

However, as noted above, the authors have not addressed this concern or the issue of transfers to other courses/HEIs.

Overall, from this review of recent models of non-completion in HE it may be concluded that there are a number of different approaches in estimating this probability. This particularly refers to the time-frame used in the empirical work where, due to data limitations, there is often a defined cut-off point (usually coinciding with the last data available) after which a student is considered as a non-completer although he/she may be continuing the studies, but taking a longer time to complete (the programme). This is an important issue in the Croatian context. Another related remark is about the cohorts examined in the empirical work. In general, assessing non-completion requires that a certain time passes between student enrolments and (non)completion before any empirical work can follow. As a result, there might be a considerable difference between the actual/current situation in the HE and the one that was examined and from which policy proposals have been developed. This limitation can be addressed to some extent by using the most recent cohorts and examining the first year non-completion. It appears that most of these problems emerge due to limited (national) databases.

Taking into consideration the limitations addressed in the above presented review of the empirical work, the next step is to develop a model that will attempt to address these limitations while taking into consideration the specific situation in Croatian HE. Subsequently, a model of student non-completion of the first year of studies is developed given the specific characteristics of the Croatian HE system.

## **DEVELOPING A MODEL OF STUDENT NON-COMPLETION**

As discussed in section 2, there is a collection of variables generally considered

as relevant. Here the focus is on the basic model of non-completion developed by Tinto. The main independent variables are the academic and social integration of the student. As argued by Tinto (1975), the higher the level of academic and social integration, the more likely is the student to complete his/her HE studies. However, Tinto's theory cannot be adequately represented by the variables used in this empirical work because some of the variables were collected at enrolment before any student integration can occur. This is mostly related to variables capturing personal characteristics. These variables were found important in the previous empirical work examined in section 3, but are lacking from Tinto's initial Student Integration Model. Therefore, the empirical work in this paper is not fully rooted in Tinto's model and the latter only serves as a guide in formulating our model. In the empirical work in this paper academic integration is represented by several variables, namely the student's prior schooling characteristics including the type of secondary school attended, secondary school grades and the score at the entry examination for admission to the HEI. A dummy variable, whether the student studied a related subject in the secondary school is included to proxy whether there is a good match between the student and the HEI course.

The social integration component is primarily captured by the quality and the extent of the peer effects operating between the student and others at the institution. This peer effects variable is constructed as the mean ability of students in the same cohort, and group of lectures as student  $i$  and with the same enrolment status (i.e. part time and full time students are considered separately since, in general, these groups have separate lectures). In constructing peer effects, the total score at admission to the HEI is used as a proxy for ability. This score consists of the secondary school grades and the score in the entry examination. Students are selected

by the HEI based on this score, hence this should be a good proxy for peer influences and aims to address some of the limitations of measures used in the previously presented empirical work. Not including this variable would suggest the model is misspecified, as previous research has shown that peer effects are a vital variable in modelling student non-completion (e.g. in Johnes and McNabb, 2004). Furthermore, Sacerdote (2001) argues that peer effects might be even more important determinants of higher education outcomes than institutional quality. When compared to the previous empirical work, the advantage of peer effects variable in this paper is twofold. First, students in this HEI are randomly (alphabetically) assigned to their course group. There are two large course groups and a student interacts mostly with others in the same group. Secondly, there are no lecturer effects, since these groups are being taught in most cases by the same lecturers.

Along with academic and social integration which are important in modelling non-completion, students enter the HEI with certain personal characteristics, goals and commitments. Therefore, the model also includes a vector of personal characteristics such as gender, age at enrolment, marital status and the place of birth (urban or rural). The rationale for including the dummy variable indicating whether the student comes from an urban or rural area is that students from the latter might find it more difficult to adjust to the new environment. To some extent, this variable is capturing the degree of social integration of the student. A proxy for socio-economic status is also included: the level of educational qualifications of student's parents. It is expected that more educated parents have a positive effect on their child's persistence in HE. Additionally, there are variables on other current schooling characteristics, such as the enrolment status, average grades obtained during the first year and a dummy

variable indicating if the student is paying the tuition fee or is exempt.

Student (non)completion is analysed for one large Croatian HEI in the field of social sciences. All students were on the same study programme with minor, if any, changes in study conditions for cohorts across time. Eight consecutive cohorts of first-year students are examined starting with the 1995/96 entry cohort up to the 2002/03 cohort. On average, there was a slow increase in the size of these cohorts over time by approximately 8 percent per year. The last available year in the dataset is 2005. This time frame allows students from the last cohort (2002/03) three years to complete the first year of studies, since in Croatia it may take some time to progress through the first year of studies. The advantage of examining first-year non-completion is that more cohorts may be used and more information may be exploited than would otherwise be feasible. As suggested by relevant literature, this is also the time when most non-completion occurs.

The model of student non-completion has the following form:

$$NC_{-1_{ij}} = \beta_0 + X_i\phi + F_i\varphi + S_i\gamma + \beta_1 P_i + \beta_2 P_i^2 + \beta_3 E_i + C_i\psi + T_i\lambda + \varepsilon_i$$

In the dataset the variable  $NC_{-1}$  is the dependant variable equal to one if student did not complete his/her first year/level of studies at the HEI, and zero otherwise. **Student's personal characteristics** ( $X$ ) include gender, age at enrolment, marital status, and a place of birth (urban or rural). **Family characteristics** ( $F$ ) include parents' educational qualifications where this variable serves as a proxy for socio-economic background. A hypothesis here is that students whose parents are more highly educated are less likely to leave HE and some evidence of this effect can be found in Ermish and Francesconi (2001), Johnes and McNabb (2004), Black et al. (2005), Ishi-

tani (2006), Ortiz and Dehon (2008), and Stratton et al. (2008).

A set of variables for **previous schooling characteristics** ( $S$ ) includes the type of secondary school that the student attended and if the student studied subjects that were related to his/her present subject area. Students' prior ability and to some extent also his/her commitment to the goal of studying at the HEI is proxied by the score at the admission exam (max. value 600) to the HEI. Furthermore, a score based on secondary school grades (max. value 400) is also used as a measure of the student's prior ability. Many studies have established a negative relationship between secondary school attainment and the probability of non-completion (Arulampalam et al. 2005; Johnes and McNabb, 2004; Stratton et al., 2008) and that relationship is also expected here. The academic **peer effects** ( $P$ ) are captured by the mean ability of students in the same group of lectures  $j$  as student  $i$ , where this ability is proxied by the obtained total score at the admission exam. A squared measure of academic peer effects variable is used to allow for the non-linear nature of academic peer effects. This squared term is expected to have a negative sign, so that the positive peer effect becomes smaller with high group mean ability (Light and Strayer, 2000; Johnes and McNabb, 2004).

Several dummy variables are used to indicate the student's **current schooling characteristics** ( $C$ ) i.e. if the student is enrolled full-time or part-time, and is he/she paying tuition fees or is exempt and his/her average grades obtained during the first year of studies (GPA1). However, some students leave their first-year studies before they can obtain any grades so there is no information on this variable for them and they are dropped from the estimation. This only occurs in 80 cases and given the size of the dataset it does not present a major limitation. Furthermore, the results do not vary much between the specifications



with and without the GPA1 variable. Therefore it was decided to use the model with GPA1 to gain more insight on the effect of grades on non-completion. In previous research on the effects of enrolment status on the probability of non-completion there is evidence that suggests that full-time students are more likely to persist in HE. Also included are the dummies for the **year of student's enrolment** (*T*), equal to one if a student enrolled in the HEI in that year and zero otherwise. The years of enrolment, as previously noted, are from 1995-2002

where the omitted category is 2002 as the last year of enrolment in the dataset. There are no multiple observations in the dataset for students who are repeating a year at the same institution. The characteristics of the dataset and the results of estimation are presented next.

### Characteristics of the Dataset

The dataset used in estimating the probability of student non-completion is discussed in this section. The variable de-

Table 2.  
Descriptive Statistics

<b>VARIABLE</b>	<b>Obs.</b>	<b>Mean</b>	<b>Std. dev.</b>
Non-completion_1	3310	0.332	0.471
<b>Personal characteristics (X)</b>			
Age	3308	19.956	3.235
Gender	3310	0.609	
Urban	3302	0.962	
Married	3303	0.327	
<b>Previous schooling characteristics (S)</b>			
Gymnasium	3214	0.568	
Other (OMITTED)	3214	0.393	
Studied related subject	3213	0.307	
Sec. school grades	2869	244.647	74.769
Admission exam	2805	375.253	56.441
<b>Family characteristics (F)</b>			
F_Uni. Or non-university college degree	2749	0.419	
F_Secondary school (OMITTED)	2749	0.562	
F_Basic or no school completed	2780	0.079	
M_Uni. Or non-university college degree	2780	0.301	
M_Secondary school education (OMITTED)	2780	0.562	
M_Basic or no school completed	2780	0.137	
<b>Current schooling characteristics (C)</b>			
Fee status	3310	0.728	
Full time student	3310	0.879	
Current GPA1	3096	3.014	0.579
<b>Peer effects (P)</b>			
Peers_1	3310	614.835	65.067
Peers_1sq	3310	382254.8	63497.52
<b>Year dummies (T)</b>			
Enrolled in 1995	3310	0.093	
Enrolled in 1996	3310	0.118	
Enrolled in 1997	3310	0.135	
Enrolled in 1998	3310	0.115	
Enrolled in 1999	3310	0.126	
Enrolled in 2000	3310	0.126	
Enrolled in 2001	3310	0.130	
Enrolled in 2002 (OMITTED)	3310	0.157	

scriptions of first-year students from all cohorts (N=3310) are presented in Table 1 (in Appendix) and descriptive statistics are presented in Table 2.

The majority of students in the dataset enrolled as full-time students (around 88 percent), 73 percent of students were paying a tuition fee and the majority of students have parents who completed secondary school. Around 61 percent of students in the dataset are females. Furthermore, the average student in the dataset was 19.9 years old at the enrolment in their first level of studies. This is above the standard entry age to HE which is from 18 to 19 years in Croatia. The age at enrolment was particularly high in the first four cohorts from 1995-98 and the mean value is above 20. This may be related to economic and political circumstances in the country at that time. The war ended in 1995, youth unemployment rate was relatively high and the possibility of finding employment with only secondary school qualifications was relatively low. This may have led more mature individuals to enrol in HE programmes at that time.

One of the limitations of this dataset is that the information on the marital status and enrolment status (full or part-time) is more recent i.e. it may be the case that the student was not married in the first year of studies, but later. Since it is impossible in the dataset to discern when the change in status occurred, only the last data on these two variables is used. Therefore, care should be taken in interpreting any particular set of results. However, when the variable on enrolment status is not used and when peer effects are based only on average ability, the main results of the empirical work do not change considerably which suggests that this variable may not present a major problem.

An additional limitation of the dataset are the missing values i.e. regression results are based on only 2207 individuals (in Table 3), although there are 3310 individuals in

the overall dataset (in Table 2). Statistical software such as Stata ignores missing values and uses only complete observations on all variables for a certain student. Therefore, if any of the observations for the individual are missing (e.g. his/her admission exam score, or secondary school grades), that subject is omitted from the analysis. This in turn leads to a decrease in the sample size and may introduce bias in the results, hence there is a need to investigate this issue further. The variables with most missing values are for parental education, admission exam score and secondary school grades. The literature on the analysis of partially missing data is quite recent. Early contributions come from Little and Rubin, (1987) and Rubin (1987). The mechanisms that lead to missing data commonly fall into three standard categories, namely missing completely at random (MCAR), missing at random (MAR) and not missing at random (NMAR). A more detailed discussion of this theory is beyond the scope of this paper, but can be found in Little and Rubin (2002). When the missingness mechanism is MCAR or MAR, then the missingness is observable and the mechanism leading to the missing data does not need to be modelled. Upon examining the dataset, no reason could be found to consider the data is NMAR. A random pattern in missingness will produce unbiased estimates (Baltagi, 2001), hence missing data does not present a problem in the empirical part of this research.

In this type of analysis where the dependant variable is binary and where the interest lies in assessing how each explanatory variable influences the probability of an outcome (completion/non-completion), the two most frequently used models are the binary logit and probit. From the empirical standpoint, logit and probit generally lead to similar conclusions for the same dataset (Long and Freese, 2005). For the empirical work the decision is to use logit.

## RESULTS

This section presents the results for the model of the probability of student non-completion, and also addresses the limitations of the approach, including problems related to missing data. The results of the logit estimation of the probability of student non-completion of the first level of studies are presented in Table 3. Given that the lo-

gistic regression is non-linear, the estimated parameters do not provide a directly useful information for understanding the relationship between the independent variable and the outcome. Therefore, marginal effects are also calculated and presented in the output. They are evaluated at the mean of other variables and a change of 0-1 for dummy variables.

Table 3.  
Regression Results for Logit Estimation of Non-completion of the First Year of Studies

VARIABLE	Coeff.	(z-stat.)	Marg. eff.
Constant	0.571	(0.32)	.
<b>Personal characteristics (X)</b>			
Age	0.107***	(5.06)	0.019
Gender	0.125	(1.11)	0.022
Urban	0.343	(1.14)	0.055
Married	-0.531	(-1.25)	-0.081
<b>Previous schooling characteristics (S)</b>			
Gymnasium	-0.669***	(-4.20)	-0.121
Studied related subject	-0.389**	(-2.34)	-0.066
Sec. School grades	-0.007***	(-8.09)	-0.001
Admission exam	-0.002	(-1.40)	-0.0003
<b>Family characteristics (F)</b>			
F_Uni. or non-university college degree	-0.483***	(-3.92)	-0.083
F_Basic or no school completed	0.064	(0.30)	0.011
M_Uni. or non-university college degree	-0.311**	(-2.33)	-0.033
M_Basic or no school completed	-0.194	(-1.09)	-0.053
<b>Current schooling characteristics (C)</b>			
Fee status	0.252	(1.57)	0.043
Full time student	-1.053***	(-3.49)	-0.221
Current GPA1	-0.134	(-1.21)	-0.024
<b>Peer effects (P)</b>			
Peers_1	-0.003	(-0.42)	-0.0005
Peers_1sq	0.000	(0.76)	0.000
<b>Year of enrolment (T)</b>			
Enrolled in 1995	-1.632***	(-5.65)	-0.194
Enrolled in 1996	-0.696***	(-2.94)	-0.106
Enrolled in 1997	-0.407**	(-2.04)	-0.066
Enrolled in 1998	-0.250	(-1.08)	-0.042
Enrolled in 1999	-0.392*	(-1.87)	-0.064
Enrolled in 2000	0.129	(0.37)	0.023
Enrolled in 2001	-0.482**	(-2.19)	-0.077

N=2207

LR $\chi^2$  =346.58, df=24

Log-likelihood: -1101.97

Pseudo R<sup>2</sup>=0.14

Notes: Significant at \*\*\*1%, \*\*5%, \*10%.

The results in Table 3 are discussed for each group of variables in turn. Unless otherwise explained, statistical significance referred to is of the coefficients at the 1 percent level.

Out of the variables capturing student's personal characteristics ( $X$ ), only age at enrolment had a statistically significant effect on non-completion, the positive effect suggesting that more mature students have a higher probability of leaving the HE programme in the first level of study. A similar finding is also present in the studies by Smith and Naylor (2001) and McGivney (2003), to mention a few, where it is suggested that the reasons for higher non-completion rates of mature students are, in most cases, financial difficulties, responsibility for dependents or academic difficulties, given that these individuals have spent some amount of time outside the education system. A marginal increase in age at enrolment from the average of 19.9 years is associated with a 2.6% increase in non-completion.

For the characteristics related to previous schooling ( $S$ ), attending a gymnasium (grammar school) has a negative and statistically significant effect on non-completion in comparison to individuals who completed other types of secondary schooling (the omitted category). The marginal effects indicate that students from gymnasiums have a 0.12 lower probability of non-completion. This negative effect corresponds to the initial hypothesis presented in section 4, and a similar result has also been found in Italian HE (O'Higgins et al., 2008; Di Pietro and Cutillo, 2009). Also, studying a related subject in secondary school decreases the probability of non-completion by 0.066 (although this is only significant at the 5 percent level). Having good secondary school grades is highly statistically significant and a marginal change in secondary school grades is associated with a 0.001 decrease in the probability of non-

completion. However, the score on the admission exam to the HEI is not statistically significant, indicating that secondary school grades are a better predictor of student persistence in HE.

Examining the proxies for socio-economic background ( $F$ ), there is a negative and statistically significant effect of having a father who completed HE in comparison to a father who completed only secondary school. The same is found for highly educated mothers (although this is only significant at the 5 percent level). Having a highly educated father decreases the probability of non-completion by 0.083 in comparison to the omitted category, holding other variables to their mean. Having a highly educated mother decreases the probability of non-completion by 0.033 (at the 5 percent level).

Within the vector of current schooling characteristics ( $C$ ), being a full-time student has the expected negative sign and is statistically significant at the 1% level. For full-time students, the probability of non-completion is 0.221 lower than for part-time students, holding other variables at their mean. This may be linked to Tinto's model where the extent of student integration plays an important role in determining student completion. It may be argued that student integration is greater for full-time students and that this may serve as an explanation for the negative sign of the variable. The variable on GPA1 is not statistically significant but has the expected negative sign suggesting that students who are obtaining good grades are less likely to leave the HEI. The coefficient on the fee status has a positive sign, suggesting that students paying the tuition fee are less likely to complete the first year of studies, possibly due to financial difficulties. However, this variable is not statistically significant.

The variables on peer effects ( $P$ ) were not found to be statistically significant in the model. Given that this variable is based

on student ability, this result may suggest that a different specification of peer effects may be more appropriate when modelling student non-completion. However, no other variable could be constructed from the available data, hence that line of inquiry could not be followed.

From the above table it may be noted that the dummy variables for the year of enrolment are all negative (except for students enrolled in 2000) in comparison to the omitted cohort of 2002 and these are statistically significant for students enrolled in 1995, 1996 (at the 1% level), 1997, 2001 (at the 5% level) and 1999 (at the 10% level). These findings suggest that students enrolling in those years were significantly more likely to complete their first level of studies in comparison to the last cohort in the dataset of 2002. There are various possible reasons for this. It may be that, despite the data set allowing at least three years for students to complete their first year (even for the final 2002/03 cohort), there are still some who will complete their first year in more time. However, perhaps a more likely reason, given that the effect occurs throughout the period, is that the increasing numbers in HE is affecting completions rates.

## CONCLUSION

The focus of this research was on first year leavers, estimating the probability that a student would drop out of university during his/her first year of study. This year has been identified as the »make or break year« i.e. it is the time when most of non-completion occurs. Non-completion was investigated for eight consecutive cohorts of Croatian students (from 1995-2002) from one large HEI.

The empirical work presented here established that mature students are more likely not to complete their studies, suggesting that any policy designed to reduce non-completion rates should be particularly sensitive to

these students in order to help them progress. On the other hand, attending a gymnasium, having high secondary school grades, studying a related subject in secondary school, being a full time student, and having parents with a university degree has a negative and significant effect i.e. discourages non-completion. The statistically significant effect of secondary school grades suggests that this variable is a good predictor of student persistence in HE, and this may contribute to the discussion of the proposed reforms in the access to HE in Croatia.

However, the limitations of the empirical work need to be addressed given that a large-scale student-level dataset for Croatian students in the HE system is not yet available. Hence, at this point in time, the individual student level information is limited to a single HEI and cannot be used to draw conclusions about the entire HE system. There is also a problem of missing data discussed in section 4.

Student non-completion can be a useful indicator of the internal efficiency of an HE system. However, the reasons for non-completion are varied and non-completion should not be seen only as a failure of individual students. It may also indicate that the education system is not meeting the needs of its customers, though due to the data limitations this line of inquiry could not be followed. Furthermore, the dataset does not offer information on the reasons for student non-completion and it was not possible to follow individual students across HEIs to distinguish between transfer students and dropouts. Both limitations in the availability of data and techniques currently preclude such analysis. However, this research offered a new insight into the characteristics of students who are not completing their first year of studies. It is also important for policy reasons as it may help to identify students who are more at risk of leaving the HEI before obtaining a degree.



## BIBLIOGRAPHY

- Aghion, P., Dewatripont, M., Hoxby, C., Mas-Collel, A., & Sapir, A. (2008). Higher aspirations: An agenda for reforming European universities. *Bruegel Blueprint Series*, 5. Available from Bruegel Organisation Web site: [http://www.bruegel.org/uploads/tx\\_btbbreugel/BPJULY2008University.pdf](http://www.bruegel.org/uploads/tx_btbbreugel/BPJULY2008University.pdf)
- Arulampalam, W., Naylor, R., & Smith, J. (2005). Effects of in-class variation and student rank on the probability of withdrawal: Cross-section and time-series analysis for UK university students. *Economics of Education Review*, 24(3), 251-262. doi:10.1016/j.econedurev.2004.05.007
- Babić, Z., Matković, T., & Šošić, V. (2006). Structural changes in tertiary education and impacts on the labour market. *Privredna kretanja i ekonomska politika*, 16(108), 28-65.
- Baltagi, B. (2001). *Econometric analysis of panel data* (2nd ed.). Chichester: John Wiley.
- Bean, J. (1980). Dropouts and turnover: The synthesis of a causal model of student attrition. *Research in Higher Education*, 12(2), 155-187. doi:10.1007/BF00976194
- Bean, J., & Metzner, B. (1985). A conceptual model of nontraditional student attrition. *Review of Educational Research*, 55(4), 485-540. doi:10.3102/00346543055004485
- Becker, G. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. Chicago: University of Chicago Press.
- Black, S., Devereux, P., & Salvanes, K. (2005). Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *The American Economic Review*, 95(1), 437-449. doi:10.1257/0002828053828635
- Braxton J. (Ed.). (2000). *Reworking the student departure puzzle*. Nashville: Vanderbilt University Press.
- Cabrera, A. F., Nora, A., & Castañeda, M. B. (1993). College persistence: Structural equations modeling test of an integrated model of student retention. *Journal of Higher Education*, 64(2), 123-139. doi:10.2307/2960026
- Central Bureau of Statistics Croatia (2008). *Statistical yearbook 2008*. Available from Central Bureau of Statistics Web site: [www.dzs.hr/Hrv\\_Eng/ljetopis/2008/PDF/00-sadrzaj.pdf](http://www.dzs.hr/Hrv_Eng/ljetopis/2008/PDF/00-sadrzaj.pdf)
- Di Pietro, G., & Cutillo, A. (2009). Degree flexibility and university drop-out: The Italian experience. *Economics of Education Review*, 27(5), 546-555. doi:10.1016/j.econedurev.2007.06.002
- Ermisch, J., & Francesconi, M. (2001). Family matters: Impacts of family background on educational attainments. *Economica*, 68(2), 137-156. doi:10.1111/1468-0335.00239
- Ishitani, T. (2006). Studying attrition and degree completion behavior among first-generation college students in the United States. *Journal of Higher Education*, 77(5), 861-885. doi:10.1353/jhe.2006.0042
- Johnes, G., & McNabb, R. (2004). Never give up on the good times: Student attrition in the UK. *Oxford Bulletin of Economics & Statistics*, 66(1), 23-47. doi:10.1111/j.1468-0084.2004.00068.x
- Light, A., & Strayer, W. (2000). Determinants of college completion: School quality or student ability? *Journal of Human Resources*, 35(2), 299-332. doi:10.2307/146327
- Little, R., & Rubin, D. (1987). *Statistical analysis with missing data*. (1st ed.). Hoboken, NJ: Wiley.
- Little, R., & Rubin, D. (2002). *Statistical analysis with missing data*. (2nd ed.). Hoboken, NJ: Wiley.
- Long, J., & Freese, J. (2005). *Regression models for categorical dependent variables using Stata*. College Station, TX: Stata Press.
- Matković, T. (2009). Pregled statističkih pokazatelja participacije, prolaznosti i režima plaćanja studija u Republici Hrvatskoj 1991.-2007. *Revija za socijalnu politiku*, 16(2), 239-250. doi:10.3935/rsp.v16i2.871
- McGivney, V. (2003). Understanding 'persistence' in adult learning, student retention in open and distance learning symposium. Open University, UK. Retrieved September 10, 2005 from Open University Web site: [kn.open.ac.uk/public/index.cfm?wpid=1887](http://kn.open.ac.uk/public/index.cfm?wpid=1887)
- Ministry of Science, Education and Sports (2007). *OECD thematic review of tertiary education, country background report for Croatia*. Zagreb: Ministry of Science, Education and Sports.
- O'Higgins, N., D'Amato, M., Caroleo, F. E., & Barone, A. (2008). Gone for good? Determinants of school dropout in southern Italy. *IZA Discussion Paper* No. 3292. Retrieved April 20, 2009 from Institute for the Study of Labor (IZA) Web site: <http://ftp.iza.org/dp3292.pdf>
- Ortiz, E., & Dehon, C. (2008). What are the factors of success at university? A case study in Belgium. *CESifo Economic Studies*, 54(2), 121-148. doi:10.1093/cesifo/ifn012
- Pascarella, E., & Terenzini, P. (1991). *How college affects students: Findings and insights from twenty years of research*. San Francisco: Jossey-Bass.
- Rubin, D. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

- Sacerdote, B. (2001). Peer effects with random assignment: Results from Dartmouth roommates. *Quarterly Journal of Economics*, 116(2), 681-704. doi:10.1162/00335530151144131
- Smith J., & Naylor, R. (2001). Dropping out of university: A statistical analysis of the probability of withdrawal of UK university students. *Journal of the Royal Statistical Society (Series A)*, 164(2), 389-495. doi:10.1111/1467-985X.00209
- Stratton, L., O'Toole, D., & Wetzel, J. (2004). Factors affecting initial enrollment intensity: Part-time versus full-time enrollment. *Economics of Education Review*, 23(2), 167-175. doi:10.1016/j.econedurev.2003.06.002
- Stratton, L., O'Toole, D., & Wetzel, J. (2008). A multinomial logit model of college stopout and dropout behaviour. *Economics of Education Review*, 27(3), 319-331. doi:10.1016/j.econedurev.2007.04.003
- Šošić, V. (2004). Does it pay to invest in education in Croatia? Return to human capital investment as a factor in human resource competitiveness. Retrieved January 30, 2006 from Institute of Public Finance Web site: [www.ijf.hr/eng/competitiveness/sosic.pdf](http://www.ijf.hr/eng/competitiveness/sosic.pdf)
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89-125.
- Tinto, V. (1982). Limits of theory and practice in student attrition. *Journal of Higher Education*, 53(6), 687-700. doi:10.2307/1981525
- Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition*. Chicago: University of Chicago Press.
- Towles, D., & Spencer, J. (1993). The Tinto model as a guide to adult education retention policy. *The Catalyst*, 23(4), 18-22.
- van der Ploeg, F., & Veugelers, R. (2008). Higher education reform and the renewed Lisbon strategy: Role of member states and the European Commission. In G. Gelauff, I. Grilo & A. Lejour (Eds.), *Subsidiarity and economic reform in Europe* (pp. 65-96). Berlin: Springer.
- Woodley, A. (2003). *Student retention in open and distance learning: Conceptualising student drop-out in higher education*. Retention Symposium, The Open University, UK. Retrieved September 10, 2005 from Open University Web site: <http://kn.open.ac.uk/public/workspace.cfm?wpid=1882>
- Yorke, M. (1999). *Leaving early: Undergraduate non-completion in higher education*. London: Falmer Press.

### Sažetak

## RANI ODLAZAK SA STUDIJA: DETERMINANTE NEZAVRŠAVANJA STUDIJA U HRVATSKOM VISOKOM OBRAZOVANJU

*Maja Mihaljević Kosor*

*Ekonomski fakultet Sveučilišta u Splitu  
Split, Hrvatska*

*Hrvatsko visoko obrazovanje doživjelo je velike promjene od 1990-ih, posebice u broju upisanih studenata kojih je danas 80% više nego prije 15 godina. Slične promjene događaju se i u europskom visokoobrazovnom sustavu. U mnogim zemljama posljedica ove ekspanzije je povećanje stope nezavršavanja studija, s naglaskom na prvu godinu studija. Kako bi ostvarili bolji uvid u problematiku, ovaj rad ispituje teorijsku osnovicu u analizi nezavršavanja prve godine studija te identificira glavne determinante. U radu se razvija i model nezavršavanja studija koji je prilagođen hrvatskom visokoobrazovnom sustavu. Baza podataka koja se koristi u analizi obuhvaća 8 generacija studenata jedne velike hrvatske visokoškolske institucije. U empirijskom dijelu istražuju se učinci varijabli kao što su osobne karakteristike studenta i njegovih vršnjaka na studiju, stupanj obrazovanja roditelja, karakteristike srednje škole i fakulteta, na vjerojatnost nezavršavanja prve godine studija. Ovo istraživanje je relevantno i primjenjivo u drugim zemljama koje su suočene s neefikasnim završavanjem studija.*

**Ključne riječi:** visoko obrazovanje, nezavršavanje prve godine studija, *Student Integration Model*, logistička regresija.

## APPENDIX

Table 1.  
Variable Descriptions

VARIABLE	DESCRIPTION
<b>Personal characteristics (X)</b>	
Age	Age of the student at enrolment
Gender	1 if female; 0 otherwise
Married	1 if the student was married; 0 otherwise
Urban	1 if student is from an urban place of living; 0 otherwise
<b>Previous schooling characteristics (S)</b>	
Secondary school type	
Gymnasium	1 if the student attended a gymnasium; 0 otherwise
Other (omitted)	1 if the student attended some other type of secondary schooling; 0 otherwise
Studied a related subject	1 if the student attended a secondary school offering related subjects; 0 otherwise
Secondary school grades	Average secondary school grades (max. 400 points)
Admission exam	Score at the admission exam (max. 600 points)
<b>Family characteristics (F)</b>	
Parental educational attainment	
Father	
F_Uni. or non-university degree	1 if the student's father obtained a university or non-university college degree; 0 otherwise
F_Secondary school (omitted)	1 if the student's father completed secondary education; 0 otherwise
F_Basic or no school completed	1 if the student's father completed basic school education or have no basic school completed; 0 otherwise
Mother	
M_Uni. or non-university degree	1 if the student's mother obtained a university or non-university college degree; 0 otherwise
M_Secondary school (omitted)	1 if the student's mother completed secondary education; 0 otherwise
M_Basic or no school completed	1 if the student's mother completed basic school education or have no basic school completed; 0 otherwise
<b>Current schooling characteristics (C)</b>	
Fee status	1 if the student is paying the fee; 0 otherwise
Full or part-time	1 if the student is enrolled full-time; 0 otherwise
Current GPA1	Average grade for courses at year one taking the value from 2.00-5.00
<b>Peer effects (P)</b>	
Peers_1	Calculated as the mean total score at the enrolment at the HEI (consisting of secondary school grades and entry exam performance) of students on the same course group as student <i>i</i> and at the same level
Peers_1sq	Square of the peers variable
<b>Year of enrolment dummies (T)</b>	
Enrol_year	1 if the student enrolled in that year, 0 otherwise

