Why do Students Leave Education Early? - Theory and Evidence on high school dropout rates

Sofie J. Cabus and Kristof De Witte

TIER WORKING PAPER SERIES
TIER WP 13/01

Why do Students Leave Education Early? - Theory and Evidence on high school dropout rates*

Sofie J. Cabus[‡] Kristof De Witte[§]

Abstract

This paper contributes to the growing literature on school dropout by proposing and empirically testing a theoretical framework on the enrollment decision of youngsters in secondary education. The model relates school dropout to time preferences, motivation, opportunity costs, and policy measures, and is empirically tested on longitudinal data of about 5,000 Dutch vocational students. We evaluate the enrollment decision of students for (1) different intensity levels of dropout prevention policy, and for (2) different levels of economic development. The results indicate that the model can accurately predict actual enrollment rates over the period 2000-2011. Using the model to forecast the level of school dropout in the Netherlands by the year 2020, it is observed that a very strict dropout prevention policy could yield nearly maximum enrollment rates (i.e., 97%) in schools. However, the annual budget for a similar dropout prevention policy is estimated at €574 million or 0.10% of the Dutch GDP.

Keywords: Dropout prevention; Economic modeling; Secondary education

JEL-code: I21, I28

*We are grateful to Emma García, Wim Groot, Henriëtte Maassen van den Brink, Patricia Ellman, and Catharina Solondz for useful comments and discussions. The authors acknowledge financial support of NICIS.

†Top institute for Evidence Based Education Research, Maastricht University, Kapoenstraat 2, 6200 MD Maastricht, the Netherlands, s.cabus@maastrichtuniversity.nl.

‡Corresponding author.

§Top institute for Evidence Based Education Research, Maastricht University, Kapoenstraat 2, 6200 MD Maastricht, the Netherlands, k.dewitte@maastrichtuniversity.nl; and Faculty of Economics and Business, University of Leuven (KUL), Naamsestraat 69, 3000 Leuven, Belgium, kristof.dewitte@econ.kuleuven.be.

1 Introduction

A common policy focus in the European Lisbon Agenda, the European Horizon 2020 Strategy, and the US No Child Left Behind Act, consists of reducing early school-leaving. This makes school dropout as one of the most central themes in the national policy agendas of developed countries. An early school leaver (or school dropout) is a youngster below the age of 23 who leaves education without a higher secondary certificate. Previous literature heavily discussed that school dropouts face significant economic and personal risks (e.g. Psacharopoulos and Layard, 1979; Psacharopoulos, 2007; Rumberger, 2011). It has been argued that particularly migrant students, boys, vocational students and pupils from disadvantageous backgrounds are at-risk for school dropout (e.g. Rumberger, 1983, 2004; Levin, 1987).

Early school leaving has multiple origins. Previous work points to the dynamic and cumulative process in which students accumulate problems before leaving education early. This process is known as the process of 'student attrition', which has various underlying factors. For example, students are more likely to drop out of school if they have one or more years of retention in grade (Planc et al., 2005), or suffer from bad health shocks (Lleras-Muney, 2005; Albouy and Lequien, 2009; Powdthavee, 2010). In addition, Attwood and Croll (2006) and Henry (2007) also argued the importance of lack of interest in schooling, as revealed by truancy behavior. And, among others, Spady (1970), Tinto (1975), Jencks and Mayer (1990), Wenger (2002) and Anderson (2010) point to the influence of the student's commitment to the school, peers, and teachers, and his/her motivation.

Furthermore, the economic cycle has its influence on school dropout. 'At-risk' students are more likely to drop out of school in times of an economic revival in response to the increased availability of (temporary) job market opportunities (e.g. Light, 1995; Shavit and Müller, 1998; Gangl, 2002; Allensworth, 2005). This is particularly observed among vocational students with internships, as they have a closer connection with the labor market compared with their peers enrolled in academic educational tracks.

This paper contributes to the growing literature on school dropout by proposing, and empirically testing, a theoretical framework designed to model enrollment of youngsters in secondary education. The model relates school dropout to time preferences, motivation, aspirations, opportunity costs, and policy measures. This contrasts with the previous literature, which has mainly focussed on what risk factors correlate with early school-leaving. Theoretical frameworks to model the school attendance decision of students in secondary education are rare. To the best of our knowledge, there is only one earlier paper by Funkhouser (1999). He discusses a reduced-form model for school attendance decisions in relation to economic conditions in Costa Rica. Our model differs from that of Funkhouser (1999), as he works with repeated cross-sections, and starts from time allocation within a household. This paper uses a panel data set to study the decision of individual students to leave education early. Furthermore, in his model Funkhouser (1999) takes the household as the main decision unit.

In our model, we suggest a focus on the individual student rather than the household, as, particularly in developed countries, students may act autonomously, independently, and often contrary to parental wishes.

Although, to the best of our knowledge, this is the first attempt to explicitly model the dropout decision of students, it is related to earlier work. Light and Strayer (2000) have investigated to what extent college quality influences graduation rates. The authors model the individual student's decision to attend college at two subsequent points in time. The decision to stay in higher education in period 2 is intertwined with the quality of college in period 1. Light and Strayer (2000) find that the likelihood to graduate from college increases with college quality, if the ability of the student matches his/her college choice.

Keane (2002) has also constructed a discrete choice model for the decision to attend higher education (college) in the US. In a similar way to Cameron and Heckman (1998), the model of Keane (2002) accounts for financial aid and borrowing constraints, factors highly associated with the attendance decision of college students.

Our model was inspired by the work of Cameron and Heckman (1998) and Keane (2002): we used a two-period discrete time model, in which students make a decision to stay at school conditional on their time constraint, motivation constraint, job market aspirations, and current (school or national) policy measures. The policy variable allows us to explicitly model, and forecast, the influence of dropout prevention policies on early school-leaving.

The theoretical model was empirically tested on a cohort of Dutch vocational students. This cohort of students is referred to as VOCL99 (*Voortgezet Onderwijs Cohort Leerlingen*). The longitudinal VOCL99 data consist of about 5,000 vocational students who were followed over the years 1999-2005. The data include information on students' school and labor market experiences.

Using the data we can estimate the influence of individual preferences on school dropout. The results and model assumptions can be used to forecast the influence of policy measures on school dropout rates in a given state of economic development. The outcomes of the model accurately predict actual enrollment rates over the period 2000-2011. They further suggest that, by intensifying the policy measures, the Dutch government should be able to reach the European Horizon 2020 goals. Finally, it is observed that a very strict dropout prevention policy could yield nearly maximum enrollment rates (i.e. 97%) in schools by the year 2017. However, the annual budget for such a similar dropout prevention policy amounts to €574 million or 0.10% of the Dutch GDP.

This paper proceeds as follows. The theoretical framework is outlined in Section 2. The empirical application, together with the data, and the descriptive statistics, are subject of Section 3. The results are presented in Section 4, and Section 5 concludes.

2 Theoretical Framework

2.1 Model set-up

This section presents a theoretical model that explains the dropout decision of students from various angles, and which is in line with the previous literature on school dropout. The discrete choice model builds on the previous work of Cameron and Heckman (1998) and Keane (2002), both of whom have focussed on the college attendance decision of students. One major difference between enrollment in college and secondary education is that students in secondary education are obliged to attend school until the compulsory education age has been reached. Notwithstanding the compulsory nature of school attendance in secondary education, students can take unauthorized absence from school for various reasons (Attwood and Croll, 2006; Henry, 2007). Unauthorized school absenteeism will be discussed in-depth within the scope of 'time allocation'. In the previous literature, lack of motivation to attend school has been highlighted as a major determinant of dropping out of school (see, among others, Adams and Becker, 1990; Jencks and Mayer, 1990; Wenger, 2002; De Witte and Rogge, 2009; Anderson, 2010). Within the scope of 'motivation and aspirations', we position intrinsic- and extrinsic motivation in our theoretical framework. The former type of motivation can be especially influenced by job market aspirations (Coleman, 1984; Ruhm, 1997; Light, 1998, 1999; Shavit and Müller, 1998), whereas the latter type of motivation can be encouraged by parents, peers, or teachers.

After incorporating the time constraint, the motivation constraint, job market aspirations, and dropout prevention policy into the theoretical model, we are able to derive a value function (i.e. the value of education). The value function reveals the main drivers of school dropout behavior – and which particular determinants prevent students from dropping out of school.

Time allocation

We use a discrete time model ($time = \{period1, period2\}$) to address the cumulative process of student attrition as referred to by, among others, Tinto (1975), Finn (1989), and Newman et al. (1992). This indicates the lack of interest in schooling, or disengagement from school, which eventually leads to school dropout. Finn (1985) refers to this process as 'withdrawal' of the student from school.

Students below the compulsory education age have to attend school, otherwise they have an unauthorized absence from school (also called truancy). Assume that, in each time period, school-age students make a decision on the time they devote to study ($0 \leq s \leq 1$), and the time they play truant ($0 \leq h_a \leq 1$). The utility from school attendance is denoted by (ϕ), and the utility from extracurricular activities by (a_1). The total study time is the sum of

compulsory school attendance ($s_{compulsory}$) and the hours of study at home ($s_{homework}$) (e.g. time spent on homework or study for a test):

$$s = s_{compulsory} + s_{homework}. \tag{1}$$

Note that, by including $s_{homework}$ in equation (1), we consider study time at home as an important – non-negligible – part of the study curriculum, as in many EU Member States, such as the Netherlands, doing homework or studying for a test are graded. Nevertheless, it is straightforward to delete study time at home from equation (1), if only the compulsory time at school is considered.

The total study time (s) and duration of the truancy spell (h_a) should not be larger than the total time endowment (L). The time constraint for a regular school attendee in each time period reads as follows:

$$L - s - h_a \geq 0, \tag{2}$$

The remaining endowed time is called leisure: $l = L - s - h_a$. For students who have dropped out of school, s is equal to 0 (i.e. no time spent in school and no study hours at home). Consequently, a dropout student in period 1 has the following time constraint:

$$L - h_a \geq 0. \tag{3}$$

At this point, it is necessary to make an important remark on compulsory school attendance in order to understand the difference between equation (2) and equation (3). While students have the obligation to attend school, they might (illegally) skip some classes. Compulsory school attendance can, therefore, be modelled as a continuum from full school attendance ($s_{compulsory} = 1, h_a = 0$), via truancy during some periods ($1 \geq s_{compulsory} \geq 0, 1 \geq h_a \geq 0$), to early school-leaving ($s_{compulsory} = 0$; $h_a = 1$). In this continuum, the total time a student spends compulsorily at school may be smaller than the duration of the truancy spell. The difference between equation (2) and equation (3) is, therefore, the total time devoted to study time at home. Consequently, for any individual, we can rewrite equation (2) as:

$$L - s_{homework} - (s_{compulsory} + h_a) \geq 0, \tag{4}$$

where ($s = 0$) and ($L - h_a \geq 0$) if a student drops out of school.

Motivation and aspirations

Students obtain utility from both intrinsic (denoted by m) and extrinsic motivation (denoted by y). We assume that both m and y follow a standard normal distribution. With respect to the former type of motivation, Ehrenberg and Brewer (1994) and Adams and Becker (1990) discuss general dislike of school and lack of student motivation as important drivers

of the dropout decision. With respect to the latter type of motivation, for instance, parents, peers and teachers have a critical role in influencing school attendance behavior (see, e.g., Jencks and Mayer, 1990; Wenger, 2002; Rumberger and Palardy, 2005). The intrinsic and extrinsic motivation sums to the total motivation, which can be formalized in a motivation constraint.[1] Note that y_1 should be smaller than m_1 for a student to yield a positive utility from motivation with respect to school attendance.

Students receive utility from school attendance if the opportunity cost of schooling is smaller than the motivation to attend school. The opportunity cost of schooling corresponds to the utility (a) from extracurricular activities in and outside the school (i.e. in the first time period a_1) times the time (h) allocated to them: $a_1 h_a$. Algebraically, a student stays in school if:

$$a_1 h_a \leq m_1 - y_1, \tag{5}$$

where the subscript 1 denotes the first time period.

Students can be motivated to achieve academic success by job market aspirations (e.g. Allensworth, 2005). Students with higher job market aspirations attach higher value to education. Conversely, given the signaling function of schooling, students with a better schooling record (e.g. absence of truancy, a higher secondary school certificate), will obtain a better job market position. This is denoted by a higher utility in the second time period (a_2). Denote by (ρ) the discount factor for the value attached in period 2 to school attendance.

Suppose now that students obtain utility from three sources: leisure (l, where $L \geq l \geq 0$); the activities they undertake while playing truant (c); and schooling (ϕ). The utility function is denoted by:

$$\begin{aligned} \lambda &= u(c, l, \phi), \\ &= u(c, l) + \phi. \end{aligned}$$

We assume that the utility function (λ) follows a uniform distribution with minimum value 0 and maximum value 1, and is concave in both arguments. Intuitively, we can explain the relevance of (λ) in the motivation constraint by socioeconomic tendencies: technological change and economic development (i.e. growing GDP) have changed the lifestyle of people over the past decades. Before the Second World War, people had only a little (money for) spare time. However, since the 1960s the global market economy has had a direct positive influence on the sale of consumption goods and/or the availability of leisure time. Given the concavity in c and l, the marginal utility from consumption and leisure decreases if c and/or l increase (see also Glewwe and Jacoby, 2004). Furthermore, students have to attend school

[1] Note that, by construction of the variable m, we consider a rather broad definition of motivation, as it includes intrinsic motivation, extrinsic motivation, and aspirations.

longer to find a connection with the labor market. We assume that the utility from schooling ($\phi \sim N(\mu, 1)$) follows a standard normal distribution. Thus, we can expect that, with an overall higher utility from schooling, in combination with a decreasing marginal utility from consumption and leisure, socioeconomic tendencies positively influence the decision to attend school.

2.2 Decision rule to attend school

Using the explanatory variables within the scope of the time- and motivation constraint, we construct a value function (V), which denotes, in essence, the value attached to a school-leaving certificate conditional on (compulsory) school attendance. In period 1, students maximize the sum of current utility λ, and the discounted utility from the next time period.[2] The value function then equals:

$$\begin{aligned} V_s &= \max_{\{h,m\}} u\left(y_1 + a_1 h_a - m_1, L - s - h_a\right) \\ &\quad + \phi + \rho^{-1} u(a_2, 1). \end{aligned} \tag{6}$$

Note that, in the second time period, students left school and started work such that the following holds: $c = a_2; l = 1$ and $\phi = 0$.

The value function for a student who dropped out of school equals:

$$\begin{aligned} V_0 &= \max_{\{h,m\}} u\left(y_1 + a_1 h_a, L - h_a\right) \\ &\quad + \rho^{-1} u(a_2, 1). \end{aligned} \tag{7}$$

Comparing the two equations brings insights into the decision rule of the student to attend school. A student will attend school if and only if

$$\text{Stay at school if:} \quad \frac{\phi}{\lambda_1} > m + a_1 s. \tag{8}$$

The left-hand side of equation (8) indicates the utility a student obtains from schooling relatively to total utility. $a_1 s$ is the opportunity cost from school attendance (i.e. a student may not participate in extracurricular activities if in school), and m the utility from intrinsic motivation. As long as the relative utility from schooling is larger than the utility from intrinsic motivation and the opportunity costs, students stay at school.

[2] Note that the discounted value can differ between different types of students. For facilitating the interpretation of the value function, here, we assume a similar discounted value for all types of students. However, in the empirical application (see Section **??**), we relax this assumption by including a measure of student-specific discounted value in the estimation.

2.3 Policy influence

The recent focus on educational performance to foster nations' competitiveness in a globalizing world (e.g. the Europe 2020 Strategy in the EU; the No Child Left Behind Act in the US) has placed early school-leaving high on the policy agenda. National governments, municipalities, and schools are gradually putting a strict dropout prevention policy into effect. Examples in this respect are: improved registration of unauthorized school absence, the introduction of community schools to motivate students, or an increased compulsory education age (European Commission, 2011a, 2011b). Similar 'policy' actions influence the decision rule of students to attend school (i.e. equation 8). Denote the impact of policy actions by θ. Allowing for policy influence, the decision rule equals:

$$\text{Stay at school if:} \quad \frac{\phi}{\lambda_1} > (1-\theta)m + a_1 s. \tag{9}$$

In sum, students will stay at school when the benefits from dropping out ($\frac{\phi}{\lambda_1}$) are higher than the efforts (i.e. direct costs m, and the opportunity costs $a_1 s$ due to school attendance). Using the motivation constraint in equation 5, one can see that, owing to a stricter policy, the utility from intrinsic motivation decreases. A more strict dropout prevention policy (θ) lowers the costs of intrinsic motivation and, thus, increases the probability that students attend school.

3 Empirical test for the decision rule

The empirical strategy

We can empirically test the decision rule in equation (8). This proceeds in two steps. In the first step, we estimate a discrete choice model (e.g. a probit model), in which the dependent variable equals the ex-post observation on whether students have dropped out of school, and the independent variables consist of proxies for motivation, time, aspirations, and a discount factor. We then may write:

$$\Pr(certificate_{i,1} = 1 | s, h_a, m, \phi, \rho, y), \tag{10}$$

where the outcome variable is the attendance decision of individual $i \in \{1, 2, ...n\}$ in period 1; s is a vector of study hours; h_a a vector of time allocated to extracurricular activities; m a vector that measures the utility from the intrinsic motivation of students; ϕ a vector that captures the utility from compulsory school attendance; ρ a vector that measures the job market aspirations of youngsters; and y a vector that measures the utility from extrinsic motivation. From this probability model, we can reveal, for the average student, the risk associated with each independent variable of becoming a school dropout.

In a second stage, using the estimated coefficients of equation (10), we estimate:

$$\text{Educational attainment} = \frac{\Phi(\hat{\phi})}{\lambda_1} - \Phi\left((1-\theta)\hat{m} + \hat{s}a_1\right), \tag{11}$$

where $\Phi(\cdot)$ is a normal probability function. We use the normal probability function to derive educational attainment in explicit form (i.e. enrollment rates). The values for θ, λ_1, and a_1 cannot be estimated from the probability model, but require assumptions. We assume that a_1 follows a uniform distribution (i.e. in the interval [0,1]). A value of 0 would indicate that students do not attach any utility to truancy, whereas a value of 1 suggests that students derive all utility from truancy. As both extremes ($a_1 = 0; a_1 = 1$) are unlikely, we start from the median student (i.e. $a_1 = 0.5$). Other assumptions on the value of a_1 are straightforward to implement. Assumptions on λ_1 (total utility) and θ (degree of policy interventions) are part of a sensitivity analysis in Section 4.

Underlying data and variables

The empirical application relies on rich longitudinal survey data of a cohort of Dutch students (VOCL99, Voortgezet Onderwijs Cohort Leerlingen). The longitudinal survey VOCL99 began in 1999, and is a national representative sample of students in the Netherlands (Kuyper et al., 2003a, 2003b, 2005; Zijsling et al., 2005, 2007). Although VOCL99 is the last but one wave, it is the latest wave where information on early school-leaving is currently available. The VOCL99 data consist of about 19,391 students, attending school in 126 different school locations. However, only students in a pre-vocational track (so-called 'vmbo' in the Dutch education system) filled in the exit survey (i.e. after graduation or school dropout). The previous literature indicates that it is particularly those students who are heavily at-risk of school dropout (Borghans et al., 2000; Borghans, 2007; Onstenk and Blokhuis, 2007). Including only these students limits the sample of VOCL99 students to 5,297 observations (N). Six years after entering secondary education, these students were asked to fill out a questionnaire on past study behavior, school experiences, and current experiences in higher education or in the job market.[3] Most sample students (78.0%) graduated from pre-vocational education in 2003 (i.e. following the regular pattern). 10.9% had already graduated from vmbo one year ahead of schedule, while 9.4% graduated one year late. A remaining 1.7% had two years retention in grade. In total, 2.10% of sample students failed the final vmbo exam.

Table 1 provides some descriptive statistics of the data, while Table 5 in the Appendix presents the relevant questions from the survey. The dependent variable measures whether a student finally obtained a certificate of secondary education (i.e. a vocational, a general education, or a pre-university certificate). Students who left school without a certificate obtain a value of 0, while graduates with an upper secondary certificate obtain a value 1.

[3] Note that, without retention in grade, a regular student takes three years to complete lower vocational education. If a student has successfully graduated from lower vocational education, he/she can start in upper vocational education (so-called 'mbo'). Only the latter course of study can result in a higher secondary diploma, which is necessary to be no longer considered as a school dropout.

From our cohort of secondary school students, 93.8% graduated with at least a pre-vocational certificate.

Next, consider the independent variables. First, the students were asked about the importance of school attendance (y) presenting them with the statement 'Attending school is important'. The respondents could then mark on a four-point Likert scale whether they agreed with the statement or not. More than 60% of the students responded that school attendance was important to very important. Second, the survey asked students about their unauthorized school attendance (h) by two questions: (1) How many times a week were you absent from school over the last school year? And (2) How many hours did you not attend school during the absence?[4] From these two variables, we computed the hours of absence a day per week, which serves as a proxy for h. About one in two students was absent from school for less than one hour a day per week. Only a low share of students (4%) has had been absent from school for more than one hour a day per week. Half of the students respond to study more than one hour a day per week. While those numbers might seem very high, they are in line with other surveys and observations. As a third independent variable, we considered the daily study time at home. This was ascertained by asking the students the question: 'How long did you study a day on average during the final grade?' We observe that 11% of students responded that they did not study after school time, while about 23% of the students studied less than one hour a day after compulsory school attendance. Fourth, students were asked whether they liked to attend school. This serves as a proxy for the intrinsic motivation (m). Almost 45% of the students replied that they did not like or only to a small exent liked to attend school. Fourth, students were asked what in their opinion was the education level necessary to obtain their desired job (ϕ). This serves as a proxy for the utility from school attendance. For most students, the education necessary for the desired future job demands at least vocational education or training (mbo). About 6% of the students responded that they were aiming for a university education. About 2% of students said that they did not need an upper secondary certificate for the job. Finally, to obtain information on the discount variable (ρ), students were asked whether they would like to be in education in four years time. In line with the expectations, this variable is equally distributed between the four response groups.

The data used in the empirical test are survey data. This has some inherent advantages and disadvantages (e.g. Bos et al. 1992). Concerning the advantages, survey data allow us to obtain valuable insights, for instance, into the students' living environment, attitudes, and motivation, which are rarely captured in administrative data. However, a major disadvantage of using surveys, is that the data might be prone to significant measurement errors, for

[4]Note that, because of self-reporting, this variable might be prone to measurement errors. Underlying analysis by Zijsling and van der Werf (2007) indicates that this is not a serious issue in the VOCL, mainly because students were only asked about their truancy behavior after graduation.

example, due to self-reporting of behavior or attrition. Referring to the survey used in this paper, Zijsling and van der Werf (2007) argue that the degree of measurement errors is rather low. Moreover, they show that the data are representative for all Dutch pre-vocational students, and that the individual questionnaire items are valid. A second issue might arise from multicollinearity between the explanatory variables. We can easily check for this by simple pairwise or inter-item correlation tests. Table 4 in the Appendix shows a relatively low, though significant, correlation between the variables.

4 Results

The first stage of the identification strategy involves the estimation of equation (10). To facilitate the interpretation, we have standardized the variables before estimating the probit regression. The results of the probit estimation are presented in Table 2. The probability of obtaining a certificate increases with study hours ($\hat{s} = 0.0524$, although not significant), utility from intrinsic motivation ($\hat{m} = 0.1061$, significant at the 1%-level) and extrinsic motivation ($\hat{y} = 0.2432$, significant at the 1%-level). It is remarkable that the extrinsic motivation has a larger influence on the probability to obtain a certificate than the intrinsic motivation. The perceived utility from school attendance increases the probability of obtaining a higher secondary degree by 18% ($\hat{\phi} = 0.1876$, significant at the 1%-level). On the other hand, the higher the value attached to future education, the less likely it is a student has obtained his/her certificate in period 1 ($\hat{\rho} = -0.0827$, significant at the 5%-level). Finally, and perhaps most importantly, the number of hours spent on extracurricular activities is negatively associated with lower educational attainments ($\hat{h}_a = -0.0817$, significant at 1%-level). This finding is in line with the previous literature indicating that unauthorized school absenteeism increases the likelihood of dropping out of school (Schaefer and Millman, 1981; Henry, 2007; Sutphen, 2010; Cabus and De Witte, 2012).

In the second stage of the empirical strategy, the estimated coefficients from Table 2 are used to predict the graduation rates in equation (9). The results are presented in Table 3 for different assumptions on the total utility λ, and the degree of policy interventions θ. By modelling θ in a second-order polynomial function ($\theta + \theta^2$), we assume that additional policy measures have an accelerating effect on early school-leaving.[5] The variable θ ranges from 0, i.e. total lack of policy on early school-leaving, to 1, i.e. an extremely strict policy on school dropout. Concerning the variable λ, we argued before that λ is concave in c and l. Therefore, lower values of λ correspond to higher values of consumption and leisure, and thus to a higher GDP. While, theoretically, λ can range from 1 (i.e. the numerair) to 0 (i.e. a total inelastic utility from additional consumption and leisure), in Table 3 we only present values of $\lambda \geq 0.6$. Smaller values of λ are unrealistic and deliver unreliable results.

The assumptions on λ are in line with theories on human capital accumulation, economic

[5] In contrast, a first-order function in θ assumes that the influence of θ is constant for each marginal increase.

Table 1: Descriptive statistics of a sample of pre-vocational students in the VOCL99 survey.

	N	5,297	%
educational attainment			
certificate		4,971	93.8%
dropout		323	6.1%
scale school attendance (y)			
0 (not important)		454	8.6%
1 (little important)		1,398	26.4%
2 (important)		2,315	43.7%
3 (very important)		1,081	20.4%
hours of truancy a day (h)			
0 hours		2,555	48.2%
0-1 hours		2,423	45.7%
1-2 hours		205	3.9%
half day		21	0.4%
full day		6	0.1%
study time a day (s)			
No study time		567	10.7%
less than 1 hour		1,242	23.4%
1 hour or more		2,457	46.4%
motivation (m)			
0 (don't like attendance)		737	13.9%
1 (only like attendance a little)		1,568	29.6%
2 (like attendance)		2,230	42.1%
3 (like attendance very much)		707	13.3%
education for job (ϕ)			
No education		114	2.2%
pre-vocational (vmbo)		261	4.9%
general (havo)		78	1.5%
pre-university (vwo)		15	0.3%
vocational (mbo-1)		70	1.3%
vocational (mbo-2)		338	6.4%
vocational (mbo-3)		784	14.8%
vocational (mbo-4)		1,075	20.3%
adult vocational (hbo)		1,951	36.8%
university		320	6.0%
other		203	3.8%
scale discount factor (ρ)			
0 (don't like education in 4 years)		1,396	26.3%
1 (only like education a little in 4 years)		1,403	26.5%
2 (like education in 4 years)		1,202	22.7%
3 (like education very much in 4 years)		1,061	20.0%

y: Importance of school attendance measured by a Likert scale.
h: Extracurricular activities measured by hours of work a day.
s: Study time measured by the additional hours of study a day after compulsory school time.
m: Utility from intrinsic motivation for schooling measured by a Likert scale.
ϕ: Utility from school attendance measured by the minimum qualifications for the desired future job.
ρ: Discount factor measured by the likelihood to be in education in four years' time.

Table 2: Estimation results of a probit regression as in equation (10).

				Number of obs=	3,968
				LR chi2(6)=	207.07
				Prob > chi2=	0.0000
Log likelihood = -791.34112				Pseudo R2=	0.1157

$Certificate$	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
Time for study - $\hat{s}$	0.0524	0.0341	1.53	0.1250	-0.0145	0.1192
Hours to extracurricular activities - $\hat{h}_a$	-0.0817	0.0238	-3.43	0.0010	-0.1284	-0.0350
Intrinsic motivation - $\hat{m}$	0.1061	0.0380	2.79	0.0050	0.0316	0.1805
Discount factor - $\hat{\rho}$	-0.0827	0.0349	-2.37	0.0180	-0.1510	-0.0144
Perceived utility from school - $\hat{\phi}$	0.1876	0.0288	6.52	0.0000	0.1312	0.2441
Extrinsic motivation - $\hat{y}$	0.2432	0.0401	6.06	0.0000	0.1646	0.3219
constant	1.7004	0.0375	45.330	0.0000	1.6269	1.7740

1. The values of the variables have been standardized before estimation of the Probit regression.
2. Lower and upper bound of a 95%-confidence interval.

development and technological change (Nelson and Phelps, 1966; Schultz, 1967; Becker, 1965, 1992; Mincer, 1974). In the early 1900s, consumption and leisure were highly valued, as people worked hard and for many hours a day. As such, one may reasonably argue that utility from consumption and leisure was very high (λ near 1) and enrollment in school very low. Socioeconomic tendencies over the past century have created welfare, and the (household) income has increased substantially. Nowadays, a considerable share of the population works in sedentary jobs for about eight hours a day, leaving more time for people to attain a proper education. Technological change has also changed the demand from low-skilled to high-skilled workers. One may reasonably argue that, today, the value attached to future education is higher than it has been over the past century. As a result of these tendencies towards knowledge, a drop in λ increases school enrollment substantially.

The results in Table 3 suggest that, holding λ constant, policy interventions have an accelerated impact on the reduction of school dropout. For example, under the assumption of an average utility of consumption and leisure of ($\lambda = 0.6$), an increase in the level of the dropout prevention policies by a quantile (from 0.25 to 0.5) reduces school dropout by about 3 percentage points from 15% to 12%. Nevertheless, a very strict policy might keep almost all students at school (i.e. under $\theta = 1.00$ and $\lambda = 0.6$, the school dropout rate equals only 2.3%). School dropout declines by about 14.8 percentage points as result of the introduction of a strict dropout prevention policy (i.e. from ($\theta = 0.00$) to ($\theta = 1.00$)). On the other hand, the utility from consumption and leisure has a more pronounced effect on enrollment rates. For example, if utility from consumption and leisure decreases (e.g. by an increase in GDP) from ($\lambda = 1$) to ($\lambda = 0.6$), the graduation rates soar under a constant policy from 19.3% to about 82.9%.

Table 3: School attendance rate based on decision rule of equation (9).

	$\lambda = 1$	$\lambda = 0.9$	$\lambda = 0.8$	$\lambda = 0.7$	$\lambda = 0.6$
$\theta = 0.00$	19.3%	29.9%	43.1%	60.2%	82.9%
$\theta = 0.25$	21.4%	32.0%	45.2%	62.3%	85.0%
$\theta = 0.50$	24.4%	35.1%	48.3%	65.3%	88.1%
$\theta = 0.75$	28.7%	39.3%	52.5%	69.6%	92.3%
$\theta = 1.00$	34.1%	44.7%	58.0%	75.0%	97.7%

where the function of θ is denoted by a second-order polynomial $(\theta + \theta^2)$.

Table 3 summarizes the results under different assumptions on λ and θ. In 2005, the Dutch dropout rate amounted to 13.5% (Eurostat). As shown in Table 3, the resulting school attendance of 86.5% corresponds to $\lambda = 0.6$ and $\theta = 0.30$. Under a stable real GDP in the period 2005-2008 (i.e. λ remains 0.6), we observed in the Netherlands an increase in the school attendance rate to 88.6%. This corresponds to $\theta = 0.52$.

While the underlying data for the empirical analysis originate from 2005-06, it is clear that they can be used for long-run predictions of the effects of dropout prevention policy on early school-leaving. Figure 1 plots the actual and predicted school dropout rates for the Netherlands. The predictions are presented for an increase in the prevention policy intensity θ by 5 percentage points each year, and under a ceteris paribus assumption for λ.[6] From Figure 1 we observe that the predictions generated by our model fit the observed data extremely well. Our model is, therefore, able to explain the observed increase in school enrollment rates in the period 2000-2010.

The available budget for Dutch dropout prevention policy has increased from €313 million in 2008 to €400 million in 2011. Every 5 percentage points increase in θ has, therefore, cost about €29 million. Extrapolating this insight, we can forecast that, if policymakers aim for a maximum dropout prevention policy intensity (i.e. $\theta = 1$), ceteris paribus, the total budget is estimated €574 million per annum. Expressed as a percentage of GDP, this corresponds to a budget (in Purchasing Power Parity in 2005 international $) rising from 0.07% in 2011 to 0.10% in 2017.

5 Conclusion and Discussion

This paper has contributed to the growing literature on school dropout by proposing and empirically testing a theoretical model of the decision making of students to stay at school

[6] As GDP (PPP, constant 2005 international $) increased in the Netherlands from 2000 to 2011 on average by only 0.67% (World Bank indicators), it can be assumed that λ stayed nearly constant during this period.

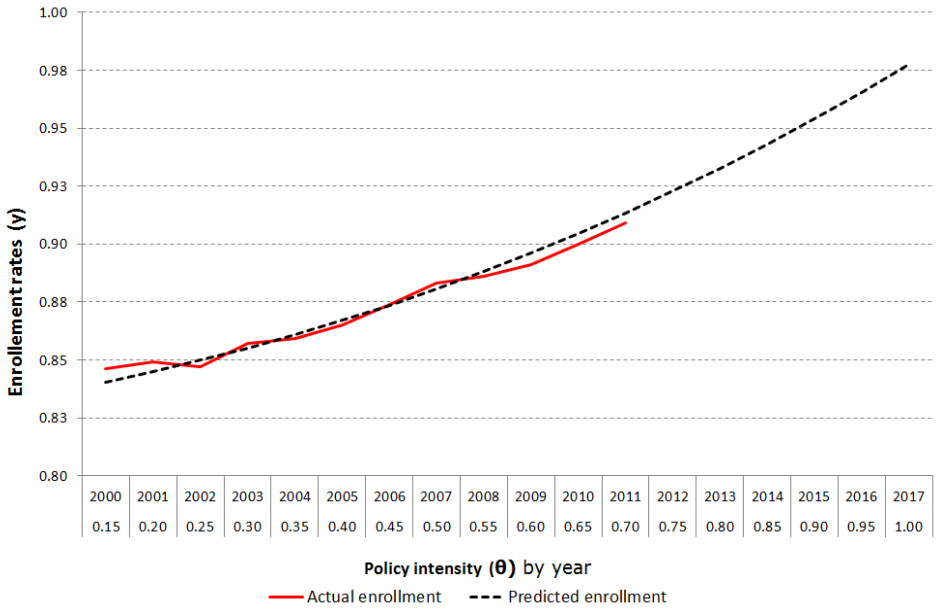


Figure 1: Actual enrollment rates (Eurostat, 2012) and predicted enrollment rates for the Netherlands, by policy intensity (θ) over the years 2000-2008. Note: λ was observed to correspond to 0.6.

in secondary education. The model is inspired by the work of Cameron and Heckman (1998) and Keane (2002). In line with the literature on educational achievement, in general, and school dropout, in particular, we construct a two-period discrete time model in which students make a decision to stay at school conditional on their time constraint, motivation constraint, job market aspirations, and current (school or national) policy measures. Thus, we start by modifying the theoretical framework of Keane (2002) to match the dropout decision of secondary education students.

The way the model is constructed allows us to distinguish between the influence of policy making and economic growth on early school-leaving. As economic development increases the importance attached to a higher secondary degree, an increasing GDP reduces the occurrence of early school-leaving. On the other hand, a stricter policy with respect to early school-leaving reduces school dropout, but it requires that scarce resources are allocated to this policy.

The decision rule derived from the theoretical model is empirically tested on Dutch survey data concerning vocational students. While the data originate from 1999, we are able to explain the observed Dutch school enrollment rates over the period 2000-2010. Moreover, the empirical application can be used to forecast the influence of a more strict school dropout prevention policy or of an increase in economic development. For example, for the Netherlands we observe that the Dutch government can easily reach the European Horizon 2020 target of 8% for early school-leaving. Our results indicate that an increase of one point in

policy strictness costs about €5.8 million.

The theoretical model could be elaborated in several ways. First, the model does not explore other sources of utility arising from, for example, class composition and educational quality. An extension to other time periods is also worth exploring. Second, the empirical model does not include additional control variables. Previous research has argued that, among other school dropout determinants, gender, ethnicity, socioeconomic background, and place of residence play an important role in the decision of students to drop out. By controlling for similar variables, the accuracy of the model might be further improved. Third, the empirical model has only been tested on the Netherlands. To facilitate alternative applications to other countries, the Stata-code is available upon request.

References

[1] Adams, J. L. , Becker, W. E. (1990). Course Withdrawals: A Probit Model and Policy Recommendations. *Research in Higher Education* 31(6), 519-538.

[2] Albouy, V. and Lequien, L. (2009). Does compulsory education lower mortality? *Journal of Health Economics* 28(1), 155-168.

[3] Allensworth, E.M. (2005). Dropout Rates after High-Stakes Testing in Elementary School: A Study of the Contradictory Effects of Chicago's Efforts to End Social Promotion. *Educational Evaluation and Policy Analysis* 27(4), 341-364.

[4] Anderson, D. J. (1998). *'If You Let Me Play': The effects of participation in high school athletics on students' behavior and economic success.* Cornell University, Ph.D. Dissertation.

[5] Attwood, G. and Croll, P. (2006). Truancy in secondary school pupils: prevalence, trajectories and pupil perspectives. *Research Papers in Education* 21(4), 467-484.

[6] Becker, G.S. (1992), Human Capital and the Economy. *Proceedings of the American Philosophical Society* 136 (1), 85-92.

[7] Becker, G.S. (1993). *Human Capital. A theoretical and empirical analysis with special reference to education.* Third Edition. National Bureau of Economic Research, The University of Chicago Press, Ltd. London.

[8] Borghans, L., W. Smits, J.D. Vlasblom, and V. Jacobs. (2000). *Leren en werken in het Nederlandse beroepsonderwijs. Vraag en aanbodontwikkeling voor de BBL 1999–2004* [Vocational education and employment in the Netherlands 1999–2004]. Maastricht: ROA.

[9] Borghans, L. (2007). *Schoolverlaters tussen onderwijs en arbeidsmarkt 2006.* [School-leavers between education and labor market]. Maastricht: ROA.

[10] Bos, K., Ruijters, A.M., Visscher, A.J. (1992). Absenteeism in secondary education. *British Educational Research Journal* 18(4), 381-395.

[11] Cabus, S.J. and De Witte, K. (2012). School Absenteeism and Dropout, A Bayesian Duration Model. *TIER Working Paper Series.*

[12] Cameron, S.V. and Heckman, J.J. (1998). Life cycle schooling and dynamic selection bias: Models and Evidence for Five Cohort of American Males. *Journal of Political Economy* 106(2), 262-333.

[13] Coleman, James S. (1984). The Transition from School to Work." In: *Research in Social Stratification and Mobility 3, edited by Donald J. Treiman and Robert V. Robinson*, 27–59. Greenwich: CT: JAI Press.

[14] De Witte, K. and Rogge, N. (2012), Dropout from secondary education: all's well that begins well. *European Journal of Education* 47 (4), 1-20.

[15] Funkhouser, E. (1999). Cyclical economic conditions and school attendance in Costa Rica. *Economics of Education Review* 14, 31-50.

[16] Gangl, M. (2002). Changing labor market and early career outcomes: Labour Market Entry in Europe over The Past Decade. *Work Employment Society*16(1), 67-90.

[17] Glewwe, P. and Jacoby, H.G. (2004). Economic growth and the demand for education: is there a wealth effect? *Journal of Development Economics* 74, 33-51.

[18] Henry, K.L. (2007). Who's skipping school: characteristics of truants in 8th and 10th grade. *Journal of School Health* 77(1), 29-35.

[19] Jencks, C. and Mayer, S.E. (1990). *The social consequences of growing up in a poor neighborhood.* In L. Lynn Jr. & M. G. H. McGeary (Eds.), *Inner-city poverty in the United States.* Washington D.C.: National Academy Press.

[20] Keane, K.J. (2002). Financial aid, borrowing constraints, and college attendance: Evidence from structural estimates. *American Economic Review* 92(2), Papers and Proceedings of the One Hundred and Fourteenth Annual Meeting of the American Economic Association, 293-297.

[21] Kuyper, H., Lubbers, M.J. & Werf, M.P.C. van der (2003a). *VOCL'99-1: Technisch Rapport.* [Technical report 99-1]. Groningen: GION.

[22] Kuyper, H. & Werf, M.P.C. van der (2003b). *VOCL'99: de resultaten in het eerste leerjaar.* [Results of grade 7]. Groningen: GION.

[23] Kuyper, H. & Werf, M.P.C. van der (2005). *VOCL'99-3: Prestaties en opvattingen van leerlingen in de derde klas van het voortgezet onderwijs.* [Results of grade 10]. Groningen: GION.

[24] Levin, HM. (1987). *New schools for the disadvantaged.* U.S. Department of Education. Office of Educational Research and Improvement.

[25] Lleras-Muney, A. (2005). The Relationship Between Education and Adult Mortality in the United States. *Review of Economic Studies* 72, 189–221.

[26] Light, A. (1995). Hazard model estimates of the decision to reenroll in school. *Labour Economics* 2(4), 381-406.

[27] Light, A and Strayer, W (2000). Determinants of college completion. *Journal of Human Resources* 35(2), 299–332.

[28] Mincer, J. (1974). *Schooling, Experience and Earnings.* Columbia University Press, New York.

[29] Nelson, R. and Phelps, E. (1966). Investment in humans, technological diffusion, and economic growth. *American Economic Review* 51(2), 69-75.

[30] Newman, F. (1992). Student engagement and achievement in American secondary schools. New York: *Teachers College Press.*

[31] Onstenk, J. and Blokhuis, F. (2007) internship in the Netherlands: connecting school- and work-based learning. *Education + Training* 94(6), 489-499.

[32] Plank, S., DeLuca, S., Estacion, A. (2005). *Dropping Out of High School and the Place of Career and Technical Education: A Survival Analysis of Surviving High School.* National Research Center for Career and Technical Education, S.P.M.N. & National Dissemination Center for Career and Technical Education, C.O.H.

[33] Powdthavee, N. (2010). Does education reduce the risk of hypertension? Estimating the biomarker effect of compulsory schooling in Engeland. *Journal of Human Capital* 4, 173-202.

[34] Psacharopoulos, G. and R. Layard (1979). Human capital and earnings: British evidence and critique. *Review of Economic Studies* 46, 485-503.

[35] Psacharopoulos, G. (2007). *The cost of school failure – A feasibility study.* European Expert Network on Economics of Education.

[36] Ruhm, C (1997). The Extent and Consequences of High School Employment. *Journal of Labor Research* 16, 293–303.

[37] Rumberger, R.W. (1983). Dropping out of high school: the influence of race, sex, and family background. *American Educational Research Journal* 20(2), 199-220.

[38] Schaefer, C.E. and H.L. Millman (1981). *How to help children with common problems.* New York: Von Nostrand Reinhold.

[39] Schultz, T. (1967). *The Economic Value Education.* New York: Columbia University Press.

[40] Shavit, Y. and Müller, W. (1998). *From school to work. A comparative study of educational qualifications and occupational destinations.* Clarendon Press Oxford, New York, pp.550.

[41] Spady, W. (1970). Dropouts from higher education: an interdisciplinary review and synthesis. *Interchange* 1, 64-85.

[42] Sutphen, R.D., Ford, J.P., and Flaherty, C. (2010). Truancy interventions: a review of the research literature. *Research on Social Work Practice* 20(2), 161-171.

[43] Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research* 45, 89-125.

[44] Wenger, J. (2002). Does the dropout age matter? How mandatory schooling laws impact high school completion and school choice. *Public Finance and Management* 2(4), 507-543.

[45] Zijsling, D.H., Kuyper, H., Lubbers, M.J. & Werf, M.P.C. van der (2005). *VOCL'99-3: Technisch Rapport.* [Technical report 99-3]. Groningen: GION.

[46] Zijsling, D.H. and van der Werf, M.P.C. (2007). *VOCL99-schoolverlaters VMBO. Beschrijving van het bestand.* [Description of the data VOCL99]. Groningen: GION.

Table 4: Correlation matrix(1)

	$certificate$	s	h_a	m	ρ	ϕ	y
$certificate$	1						
s	0.1235 (0.0000)	1					
h_a	-0.1357 (0.0000)	-0.1302 (0.0000)	1				
m	0.1630 (0.0000)	0.2435 (0.0000)	-0.1086 (0.0000)	1			
ρ	-0.0122 (0.3843)	-0.0009 (0.9563)	0.0422 (0.0028)	0.0597 (0.0000)	1		
ϕ	0.1391 (0.0000)	0.0792 (0.0000)	-0.0435 (0.0018)	0.0993 (0.0000)	0.2468 (0.0000)	1	
y	0.1933 (0.0000)	0.3649 (0.0000)	-0.1570 (0.0000)	0.4643 (0.0000)	0.1254 (0.0000)	0.1277 (0.0000)	1

Note: P-values between brackets.

6 Appendix

Table 5: Definition of the variables in the VOCL99 questionnaire.

	Dependent variable (Educational attainment)
Question 1	When did you leave school without a certificate? 1. I don't know anymore 2. In 7th grade 3. After 7th grade 4. In 8th grade 5. After 8th grade 6. In 9th grade 7. After 9th grade 8. In 10th grade 9. After 10th grade (HAVO or VWO) 10. After 10th grade: failed the VMBO exam
	Hours of study (s)
Question 6	How long did you study a day on average during the final grade? 1. I don't know anymore 2. I did not do homework 3. Less than 1/2 hour a day 4. 1/2 to 1 hour a day 5. 1 to 1.5 hours a day 6. 1.5 to 2 hours a day 7. 2 to 2.5 hours a day 8. 2.5 to 3 hours a day 9. 3 to 3.5 hours a day 10. more than 3.5 hours a day
	Hours of truancy (h) - (1)
Question 11	How often were you unauthorized absent from school last year? 1. Never 2. One or several times 3. Every month at least once 4. Every week at least once 5. Every week at least two or three times 6. At least every day
	Hours of truancy (h) - (2)
Question 12	How long were you every time unauthorized absent from school? 1. One or two hours of class 2. A half of school day 3. A total school day 4. Multiple days in a row 5. fluctuating
	Motivation (m)
Question 5.12	Mark the answer best fitting your opinion: In that final grade I most often did like going to school. 1. do not agree 2. agree a little 3. agree 4. totally agree

	Scale discount factor (rho)
Question 17	Below there are 6 choices that represent your future in four years. Tick where appropriate. In four years I would like to... 1. be enrolled in education 2. have paid work 3. be enrolled in education and have paid work 4. live on (unemployment or social) benefits 5. do the housekeeping 6. do volunteer work

	Education for job (phi)
Question 21	What is the educational level necessary for that job? 1. None 2. VMBO (pre-vocational) 3. HAVO (general secondary) 4. VWO (pre-university) 5. MBO 1 (vocational-1) 6. MBO 2 (vocational-2) 7. MBO 3 (vocational-3) 8. MBO 4 (vocational-4) 9. HBO (adult vocational) 10. University 11. Other, namely...

	Scale school attendance (y)
Question 5.17	Attending school is important 1. do not agree 2. agree a little 3. agree 4. totally agree

TIER WORKING PAPER SERIES
TIER WP 13/01
© TIER 2013
ISBN 978-94-003-0048-4