Academic and Labor Market Success: The Impact of Student Employment, Abilities, and Preferences.

Juanna Schrøter Joensen*†

February 27, 2007

Abstract

This paper analyzes the transition from university education to labor market work. The decision to drop out or continue education until a degree is acquired depends on the student's academic achievement and labor market opportunities with and without a degree. These are explicitly modelled in a stochastic dynamic environment in order to disentangle the channels through which student employment, abilities and preferences affect academic and labor market success. Estimation of the model reveals that a little student employment has a positive impact on academic achievement, while too much has a negative impact. Abilities and preferences are important determinants of academic success. Students with higher academic abilities have significantly lower dropout rates, while students with higher consumption value of university attendance tend to have a higher probability of graduating, but to spend longer time to graduation.

JEL Classification: I21, I28, J24.

Keywords: Student Employment, Investment in Education, Sequential Resolution of Uncertainty, Grade Level Progression.

*Department of Economics, Stockholm School of Economics. Juanna.Joensen@hhs.se.

†Financial support from the Danish Social Science Research Council is gratefully acknowledged. Part of the work for this paper was done while I was a visiting research scholar at the Department of Economics, University of Texas at Austin. I thank the department for its hospitality. I appreciate guidance and useful suggestions from Helena Skyt Nielsen. The paper has also benefited from valuable discussions with Karsten Albaek, Peter Arcidiacono, Jesper Bagger, Russell W. Cooper, Daniel Hamermesh, Michael Svarer, Steve Trejo and Astrid Würtz. Thanks to seminar participants at Stockholm School of Economics, University College Dublin, University of Texas at Austin, University of Aarhus, and the LMDG 2007 (Sandbjerg) and the ESPE 2007 (Chicago) conferences for comments and suggestions. Any remaining errors and omissions are my responsibility.

1 Introduction

Reducing dropout rates in higher education and increasing the speed at which individuals obtain a given educational level are declared social goals in most countries around the world. However, despite the considerable amount of studies and public debate on these issues, not much is known about the impacts of potential policy interventions. Indeed, important policy decisions have been made based on beliefs about potential impacts on dropout rates and times-to-graduation, in particular the premise that student employment adversely affects academic achievement. This paper provides hard evidence regarding the channels through which student employment affects academic and labor market success and the expediency of these prior beliefs.

As all individuals at some point in time are destined to make the transition from full-time education to full-time labor market work, it is also pivotal that this transition is as smooth as possible - both for the individual and society. Working part-time while enrolled in full-time education, henceforth student employment, can be seen as an instrument through which students achieve this goal. Student employment is a common phenomenon among university students. The average 15-29 year old student living in the average OECD country in 2003 was employed the equivalent of 27% of full-time employment while enrolled in education. Student employment is lowest in France (20%) and highest in Denmark and the Netherlands (57% and 53%, respectively).[1] The average US student is working the equivalent of 39% of full-time employment while enrolled in education.[2] Students are thus a considerable part of the (unskilled) labor force across all OECD countries and they provide a flexible reserve of labor as they often work part-time and adjust their labor supply to the current labor market situation.

Student employment is a highly debated issue because of the gains and losses it generates for individuals and the society. Previous studies have found that student employment potentially reduces the cost of the school-to-work transition, as it increases the probability of stable employment as well as earnings - particularly in the early career; see e.g. Light (2001), Hotz et al. (2002), and Häkkinen (2006). On the other hand,

[1] Interestingly, 15-29 year olds in Denmark expect to spend as much as 9.1 years in education, while 15-29 year olds in Netherlands only expect to spend 5.9 years in the educational system in order to obtain a corresponding level of formal education. This indicates the importance of explicitly specifying the environment within which educational and student employment decisions are made.

[2] Cf. OECD Education at a Glance (2005), Table C4.1a.

student employment is found to lower academic achievement by increasing the probability of dropping out and the time-to-graduation; see e.g. Ehrenberg and Sherman (1987), and Stinebrickner and Stinebrickner (2003). In addition to potentially reducing academic achievement, working directly reduces leisure time, which is itself valued by students.

The essential question is whether the gains from student employment are larger than the distortion effect it has on achievement at the university. To assess the overall effect of student employment, this paper answers three essential questions: First, does academic success, measured by acquired grade levels, (positively) affect wages? Second, does the amount of student employment (negatively) affect academic grade level progression? Third, what is the direct effect of student employment on wages? To answer these questions, students' educational and employment decisions are explicitly modeled in a dynamic stochastic environment. Each year individuals make decisions on whether to stay enrolled at the university or to enter the labor market and work full-time given wages that depend on acquired academic degrees and accumulated labor market experience. Individuals who are still enrolled at the university further decide how many hours to work part-time. These decisions are all conditional on the individual's prior academic and labor market performance, as well as expectations about the future. This paper thus explicitly takes the sequential nature of educational and employment decisions into account, and simultaneously models labor market opportunities with and without university degree(s), accumulation of academic capital in terms of grade level progression, and the accumulation of labor market experience through employment decisions while attending the university. Hence, dropout probabilities, times-to-graduation and accumulation of work experience while enrolled at the university are endogenously determined. The dynamics are important for three reasons. First, the dynamics make it possible to separate the effect of the educational environment from the effect of the labor market on the educational and employment choice. Second, the dynamics also allow individuals to learn about their abilities through accumulated course credits. Those who perform worse than expected may find it more attractive to drop out (or switch to another field). Finally, the dynamics make it possible to control for selection into the various stages of the model; see e.g. Keane and Wolpin (1997), and Cameron and Heckman (1998).

Another novelty of this paper is that it explicitly accounts for uncertainty of academic success. Despite the fact that dropout rates are high, traditional human capital investment models ignore the role of uncertainty (and possible failure) in educational investment and assume that an undertaken educational spell is successfully completed with certainty *and* within the ordained time span.[3] However, most students spend excess time-to-graduation and dropout rates are high. In the average OECD country in 2003, 30% dropped out of higher education.[4] Hence, the survival rate of 77% at Danish universities is slightly above the OECD average of 70%. For the US, Altonji (1993) reports that in the National Longitudinal Survey of the High School Class of 1972 (NLS72) sample about 60% of college candidates actually complete college.

The most important issue to address when estimating the effect of student employment on academic and labor market success is self-selection. Selection bias arises if students choose the amount of student employment based on unobservable characteristics that are also correlated with academic and labor market ability. To account for unobserved student heterogeneity I assume that students are drawn from a finite mixture distribution, where each student is assumed to belong to one of a finite number of *types*, each of which has its own distribution with respect to preferences for university education, labor market and academic ability and/or motivation. Each student's type is unobserved to the econometrician, but can be inferred by Baye's rule. This is exploited both in the estimation strategy based on the conditional choice probability (CCP) estimator developed in Arcidiacono and Miller (2007), and to relate student's unobserved type to observed family and other socioeconomic background characteristics. The potential biases arising from self-selection are eliminated if students make their decisions based on their observable characteristics and type.

The model parameters are estimated using register-based Danish panel data. The data covers a random 10% sample of the Danish population. I have individual level data on detailed educational event histories and labor market histories, including actual labor market experience, approximate working hours, unemployment degree, labor income, and wages. I also have information about parents, courses taken in high school and high school grade point average (GPA). Furthermore, data on accumulated course credits has been collected and merged for the particular purpose of this study. In Denmark, there are no tuition fees for university education, the enrolment period is not restricted,

[3]This literature testifying to the importance of education to the structure and evolution of earnings has grown out of the human capital investment literature pioneered by Mincer (1974) and Becker (1964). See Belzil (2007) for an excellent survey of the evolution of the human capital investment literature with particular focus on stochastic dynamic programming models.

[4]cf. OECD Education at a Glance (2005), Table A3.4.

and all admitted students are eligible for a study grant that suffices for necessary costs of living. University admission is to a high extent only conditional on GPA and Math level from high school. Since the individual's municipality of residence as well as the municipality of the attended university are also observed in the data, average local unemployment rates during university enrolment can be constructed (as a proxy for labor market opportunities).

Estimation of the structural model reveals that a little student employment is complementary to academic achievement, while too much is detrimental. On the other hand, student employment increases wages and reduces the job finding cost. Conditional on the student's observed abilities and skills, there is no evidence of differential effects of more study-related jobs nor from working in jobs that require higher skill levels. Abilities and preferences are found to be important determinants of academic success. Observed abilities and skills reduce dropout rates and times-to-graduation, while student types with higher unobserved academic abilities and/or motivation and higher consumption value of university attendance tend to have lower dropout rates, higher Master graduation rates, but also a higher probability of spending excess time to Master graduation.

A deeper understanding of what influences educational and employment decisions as well as academic and labor market success is a matter of concern not only to educational researchers who seek to better understand the determinants and consequences of human capital investments is pivotal in order to make informed policy decisions. It is important for institutional officials who look to improve their admission policies, as well as policy makers who attempt to develop more effective policies to improve the efficiency of higher education and ease the school-to-work transition for highly educated individuals.

The rest of the paper is organized as follows: Section 2 provides some background information and places the paper in the context of previous literature. The data and the empirical regularities present in the data are presented in Section 3. Section 4 presents the dynamic model of educational and student employment decisions, and the econometric techniques used to estimate the model parameters. Section 5 brings about the empirical results. Section 6 provides some evidence on the heterogeneity of the stock of human capital by also considering the field of university education and the types of student employment. Section 7 concludes.

2 Background

The school-to-work transition of university graduates draws together issues such as wage structure, student employment opportunities and academic opportunities and abilities. It emphasizes the attributes of the processes and their linkage as individuals flow from full-time university education to full-time permanent employment through a varied set of intermediate conditions. Thus the dynamics are pivotal, since it is important to allow choices and outcomes at one point in time to affect future choices and outcomes. In order to explore to what extent adverse experiences during transitions have lasting effects on the prospects of those who experience them, it is important to get hard evidence on the presence of state dependence. Potentially damaging events include educational failure, and the outcomes potentially affected by them wages after university exit. The main empirical challenge in determining the effect of state dependence is controlling for (dynamic) selection. For example, any apparent adverse effects of the incidence of early low academic performance and subsequent low labor income may not reflect scarring by adverse previous experiences, but rather joint selection into both states according to intrinsic individual attributes. Individuals with lower ability and/or motivation may be more likely to underachieve both early and later on even if no causal mechanism is present. Despite the difficulty of controlling for selection effects, empirical evidence suggests that early failure in education and in the labor market has important effects (although small) on subsequent successes; see e.g. Ryan (2001). This evidence points towards the potential importance of school-to-work policies and towards the need for hard evidence on student employment as one potential device in accomplishing a smoother transition.

There are several reasons why individuals may choose to enter the labor market prior to graduation. Many students depend on the extra income; however, employment may also be an investment in enhancing labor market skills. Firstly, by working part-time students might improve their interpersonal skills, get familiarized with the labor market and gain a sense of responsibility, all of which are valuable skills in their later career. Second, potential employers might view student employment as a signal of other favorable attributes, such as high motivation and/or ability. Third, student employment might also enhance job search to the extent that labor market contacts improve employment opportunities after graduation. In most fields of study, university education does not prepare students for one specific job and therefore relevant work experience may be essential for finding the first job after graduation. Finally, work experience in the student's field of study may complement the formal education and improve study motivation. This view that part-time jobs can provide students with valuable work experience and income and may even be a stepping stone to better jobs after graduation, primarily stems from the occupational choice and employer-worker matching literature that views it as beneficial labor market search and job-worker matching that leads young workers and employers to better and more durable employment than they otherwise would have found. For example, Topel and Ward (1992) view jobs during the school-to-work transition as part of early career "climbing up the ladder".

There might, however, also be negative effects of student employment. Obviously, there is a time-use trade-off between working and studying. Working during the semester may interfere with learning and academic performance, and may even encourage students to drop out. Since the enrolment period at the university is not restricted, working might also lead to longer times-to-graduation. Garibaldi, Giavazzia, Ichino and Ettore (2007) (and the references therein) provide evidence on excess times-to-graduation in the US. Brunello and Winter-Ebner (2003) study expected times-to-graduation for (Economics and Business) university students in 10 European countries. They find that the fraction of students who expect to graduate at least one year later than the required time ranges from above 30% in Sweden and Italy to almost zero in the UK and Ireland.[5] In my sample of Danish university students, the average time-to-graduation with a Master's degree is almost two years longer than the target duration, and 64% of Master graduates (37% of Bachelor graduates) spend more than one year in excess of the required time to graduate. Long times-to-graduation provide private monetary costs to individuals by shortening their careers after graduation. Brodaty, Gary-Bobo and Prieto (2006) also provide evidence that French individuals with longer than average time-to-graduation have significantly lower wages and employment rates in their early career - indicating a signalling effect of graduating on time. On top of that, there is the direct social cost of publicly funded grants for students.[6]

The literature on the effects of financial incentives on academic performance and

[5] Their finding that excess time-to-graduation is higher in countries where the share of public expenditure relative to total expenditure for tertiary education is higher and with lower college wage gap, points towards longer times-to-graduation in Denmark.

[6] Danish university education is almost exclusively publicly funded, and public subsidies directly to students make up more than 30% of total public expenditures on university education. This is the equivalent of 0.85% of GDP, cf. OECD Education at a Glance (2005).

times-to-graduation provides ambiguous evidence.[7] However, most of this literature does not control adequately for confounding unobservable factors. Garibaldi et al. (2007) use tuition discontinuities to identify the effect of university tuition on times-to-graduation of students at Bocconi University in Milan. They find that a 1000 Euro increase in tuition reduces the probability of late graduation by 6 percentage points (with respect to an average graduation probability of 80%). Like in the basic model of this paper, an increase in tuition is equivalent to a reduction in study grants. Not much is known about the optimal length of the period for learning the required skills to obtain any given degree. However, Garibaldi et al. (2007) argue that in the presence of market imperfections (like public subsidies to education) private student incentives do not lead to socially optimal times-to-graduation. Long times-to-graduation are considered a waste of potential high skill labor (for society), and particularly in a welfare state with an ageing labor force it is seen as pivotal to get highly skilled university graduates out on the labor market at a faster rate; see e.g. Velfærdskommisionen (2005). These issues overlap with the possible vanishing of unskilled jobs in high wage countries. The potential cause of an excision of unskilled job opportunities for young individuals is typically seen as technological change, intensified by the export of unskilled work to lower wage countries as a result of continuing globalization: both these factors appear to be causing a general upskilling of the job structure in high-wage countries. Brunstein and Mohler (1994) also find that the labor market increasingly demands adaptable and flexible workers with high levels of academic and technical skills. At the same time there has been a general rise in higher educational participation, and the labor market for university graduates makes up an increasing share of the total labor force; see e.g. Bacolod and Hotz (2006), who find that both educational attainment and student employment have increased over the last decades. In Denmark, the fraction of university graduates doubled from 1980 to 2004. In the sample period 1994-2004, around 40% of high school graduates pursued a university education.

Much research has been devoted to estimating the effects of working while in high school.[8] However, even though working while attending education at a later age is much more common, much less research has been devoted to estimating the effects of working

[7] Garibaldi et al. (2007) provide an excellent survey of this literature.

[8] See e.g. Eckstein and Wolpin (1999), who find that student employment has a negative impact on the performance of high school students. However, they find that restricting student employment would have small effects on graduation rates. Ruhm (1997) provides a thorough survey of the potential effects of student employment on subsequent labor market success.

while enrolled in higher education. Furthermore, work experience acquired shortly before graduation from the university might be more important for university graduates' labor market success than work experience acquired earlier. Previous literature examining the effects of working while in higher education is divided into two branches: studies of the effects of student employment on academic success and studies of the effects on labor market success. Overall, there seems to be a consensus that student employment lowers academic performance and increases subsequent labor market performance. The main estimation issue is to disentangle to what extent the estimated effects are causal or simply reflect the persistent role of unobservable differences in abilities and preferences that influence the likelihood of student employment, academic success and subsequent success in the labor market. There are a few studies that try to address the endogeneity of student employment. In the branch of literature examining the impact on academic success, Stinebrickner and Stinebrickner (2003) use an instrumental variable (IV) approach where they exploit the variation in mandatory work-study programs at Berea College to identify the causal impact of student employment on academic performance. In their sample, all students work at least 10 hours a week, but some students have the option to work additional hours. They find that working additional hours significantly reduces academic success. Ehrenberg and Sherman (1987) find that working less than 25 hours a week has no impact on grade point average, however, it does increase the probability of dropping out and reduces the probability of graduating on time. They also find evidence that dropout rates and times-to-graduation are only adversely affected if students work in off-campus jobs. They estimate a static recursive model. Although their model does take the simultaneity of student employment and educational choices into account, it does not control adequately for unobserved individual heterogeneity. Hence it is difficult to conclude whether these effects reflect causal impacts of student employment on outcomes.

In the branch of literature examining the impact on labor market success, Hotz et al. (2002) find that the positive effect of student employment on wages diminishes substantially when controlling for unobserved individual heterogeneity. Furthermore, they show that the effect becomes statistically insignificant when also dealing with dynamic selection.[9] Häkkinen (2006) uses average local unemployment rate during university enrolment as an instrument for acquired work experience during university

[9] See e.g. Heckman (1981) and Cameron and Heckman (1998) for a thorough discussion of dynamic selection bias and a detailed description of the estimation strategy applied by Hotz et al. (2002) to correct for this bias.

enrolment. She shows that student employment increases earnings and employment rates one year after graduation, but that the effects become statistically insignificant in later years.

This paper joins the two branches of previous literature and contributes by unraveling the mechanisms through which student employment affects academic- and labor market success. Previous literature has claimed that student employment can reduce the costs of the school-to-work transition. This paper provides a direct estimate of these costs in terms of job finding costs, as well as the benefits in terms of higher wage offers and the extent to which student employment complements or detriments academic success, which in turn increases future wages.

3 Data

For the estimation of the model a very rich register-based panel data set comprising a random 10% sample of the Danish population is used. The data set is hosted by the Danish Institute of Governmental Research (AKF) and it stems from Statistics Denmark, who has gathered the data from different sources - mainly administrative registers.

The data contains observations on labor income (earnings), gross income, net income, and wages for the period 1984-2004. All incomes are observed at year-end and deflated to real values measured in year 2000 DKK using the average consumer price index, $PRIS8$, available from Statistics Denmark.

Complete detailed educational event histories are observed for each individual for the period 1978-2005. These comprise detailed codes for the type of education attended (level, field, and educational institution) and the dates of entry and exit, along with an indication of whether the individual completed the education successfully, dropped out or is still enrolled as a student. Furthermore, data on accumulated course credits during the period 1995-2005 have been collected for the particular purpose of this study.[10]

Since educational event histories are available on a monthly basis and accumulated course credits, income, and the other socioeconomic background variables are available on a yearly basis, I have to take a stance on the timing of the educational events. I

[10] I am currently attempting to further augment the data with information on grades from university courses.

assume that an individual is in full-time education if the individual is enrolled at the university more that half of the year, i.e. if the educational spell (lasting more than a year) begins in June or before or ends in July or thereafter. 92% of individuals enter the university in September, hence are registered as being in full-time education from the beginning of the year after they actually enroll. Final exams are usually in January and June, however the individuals accumulate on average 1 course credit in the year prior to the year when they are coded to enter the university. I assume that these course credits are from passing rare December exams and transfer them to the following year. Furthermore, I assume that there are equally many December exams each year of the university study spell and therefore transfer equally many course credits for each year until university exit.[11]

The data set also contains information on course choices in high school and high school GPA.[12] The GPA is a weighted average of the grades at the final exams of each course. All high school courses can be obtained on three different levels referring to the difficulty of the course: low, medium, and high. Both the quality of the courses and the GPA are comparable across high schools, since the control of the high school is centralized at the Ministry of Education. Furthermore, all high school students within each high school cohort are faced with identical written exams, and the oral exams and the major written assignments are evaluated both by the student's own teacher and an external examiner assigned by the Ministry of Education.

Danish university entrants are solely screened on high school GPA and course choices. Hence, it is possible to construct individual choice sets quite precisely, since information about admission requirements is also available in book form and has been merged to the register-based data. Students are admitted to courses leading to a Bachelor or a Master degree. The European Credit Transfer and Accumulation System (ECTS) is used to proxy grade level requirements in terms of course credits. To successfully complete a Bachelor's degree, 180 ECTS have to be accumulated in one major (and possibly also a minor) field. Likewise, to obtain a Master's degree, 300 ECTS have to be accumulated. Most programs are designed so that it is possible to graduate in five years.[13] In principle, however, students can stay enrolled as long as they wish.

[11] A detailed description of the coding of educational events and the course credit accumulation is provided in Appendix A.

[12] In Denmark a numerical grading scale system is used. The possible grades are 00, 03, 5, 6, 7, 8, 9, 10, 11, 13, where 6 is the lowest passing grade, and 8 is given for the average performance.

[13] Medicine is an exeption, requiring six and a half years, since the last year and a half consists of

Note that there are neither Ivory League effects nor tuition fees for university education in Denmark, and all students above 18 receive a study grant from the government, which covers living expenses.[14] Students living with their parents receive a reduced grant, but the grant is independent of parental income, educational effort and achievement as long as the student is less than one year behind scheduled study activity.[15] All enrolled students are eligible to collect the study grant for a maximum of 70 months - which is the target duration to acquire a Master's degree plus one year. The study grant also depends on income from student employment. If yearly student earnings exceed a certain threshold, then the study grant is reduced. Note that during the observation period there was a change in financial aid rules that is incorporated into the estimation strategy: the threshold level for permissible student earnings was raised from $48,000$ DKK to $61,000$ DKK in 1996, while the maximal study grant remained unchanged at around $48,000$ DKK per year.[16] This threshold change may increase the students' incentive to work more and is incorporated into the structural estimation strategy. The effect of this change is discussed further and evaluated in Section 5.4.

3.1 Sample Selection

Among the gross population of high school graduates who are eligible to enter a university education, I select those who initially enrol in a university education when they are between 18 and 22 years old. Since course credit data is only available from 1995 onwards, I only select students who enter the university in September 1994-1996. These university entrants are observed until the end of 2004. The sample comprises $2,129$ individuals - amounting $19,349$ observations of individual characteristics, choices and outcomes over time.

3.2 Descriptive Statistics

The average individual in the sample enters the university in 1995 and is 21 years old at the time of initial enrolment. Most individuals enter a university in one of the two largest cities: 48% enroll in Copenhagen and 24% in Aarhus. The prerequisites from

---

mandatory vocational training.

[14] Until 1996, this age limit was 19 years.

[15] 99% of university students in the sample do not live with their parents.

[16] The exchange rate on December 31, 2000 was 8.0205 DKK/USD and 7.4631 DKK/Euro.

high school of the average individual is a GPA of 9, a Math level of 2.1, a Science level of 1.9, a Social Science level of 1.8 and a Language level of 2.2. 48% are females. Table 1 displays descriptive statistics of the estimation sample separately for university dropouts, Bachelor, and Master graduates. 23% of university entrants drop out, 22% acquire a Bachelor's degree, while 55% acquire a Master's degree as their highest completed university degree.[17] The top part of Table 1 presents student characteristics at university entry. It is seen that Master graduates have significantly higher GPA, Math and Science level from high school.[18] Dropouts, however, do not seem to have disadvantaged observable characteristics compared to Bachelor graduates.

The middle part of Table 1 concerns achievement during university enrolment, and reveals that dropouts on average stay enrolled at the university almost as long as graduates although they accumulate fewer course credits and work more each enrolment year. Master graduates stay enrolled for 6.5 years on average in order to obtain a Master's degree that requires 5 years of full-time study. Bachelor graduates have even longer excess times-to-graduation as the Bachelor's degree requires 3 years of full-time study and they on average stay enrolled for 6.9 years. This may be because many Bachelor graduates are Master dropouts or still enrolled by the end of the sample period. The fact that Bachelor graduates at university exit on average have accumulated 8 course credits more than required to obtain the Bachelor's degree could indicate that many of them are dropouts from the Master program. Figure 2 further reveals that 8% are still enrolled 10 years after university entry leading to some right cencored observations.

For each year of enrollment, dropouts accumulate fewest course credits each year, while for graduates: the higher the highest attained degree, the more course credits accumulated each year. Dropouts accumulate on average the equivalent of 23 ECTS, Bachelor graduates 39 ECTS, and Master graduates 47 ECTS. For accumulated student employment the reverse holds true. On average, students accumulate the equivalent of one year of full-time labor market experience through student employment. Dropouts tend to work more during their studies, accumulating almost half a year of full-time work experience more than the average student during university enrolment.

[17] Note that among the 23% who drop out of university education, subsequently 3% acquire a short cycle higher education and 6% acquire a medium cycle higher education - primarily as school teachers and nurses.

[18] This is in accordance with the literature on ability sorting across levels of education, see e.g. Willis and Rosen (1979), Cameron and Heckman (1998) and Card (1999).

The last part of Table 1 shows earnings differences by level of university education. It reveals that the monetary incentive in terms of degree premiums exists for all levels of university education. Particularly, the premium to Master graduation is high. The hourly wage for Master graduates is 40 DKK higher than for Bachelor graduates, who in turn have 5.5 DKK higher wages than dropouts. The Master's degree premium already exists in the first year after graduation and seems to persist throughout the early career. Interestingly, dropouts earn more than Bachelor graduates just after university exit. This could be an indication that dropouts learn that they have relatively high (unskilled) labor market abilities, relatively low academic ability and/or place lower utility value on university attendance. All in all, it seems that the primary monetary value of a Bachelor's degree is the option value associated with pursuing further university education.[19]

Annual labor market experience while enrolled in full-time education is shown in Figure 1. Average student employment tends to increase monotonically with time since initial enrolment. The proposed model has several explanations for this pattern, as it predicts that students will increase their labor supply - both as they accumulate more academic and labor market capital, cf. Section 4. Figure 1 further illustrates the importance of explicitly modeling the decision process, since forward looking individuals who perceive their probability of graduating as small might be more likely to work. The figure shows that dropouts tend to work more hours. However, among those who graduate, those who acquire a higher degree tend to work less during the first enrolment years, but to increase their student employment more over time. This makes the selection issue problematic and underlines the importance of controlling for (dynamic) selection, since the positive relationship between highest acquired degree and student employment might reflect inherent differences in ability and/or motivation rather than the acquisition of skills that are complementary to academic achievement. The estimated model will take this selection into account.

Figure 2 shows the transition patterns over time after initial university enrolment. It shows that more than 70% of the enrolled students are still in full-time university education five years after initial enrolment, although the required time to acquire a Master's degree is only five years, and only 55% of the students in the sample obtain a Master's degree. The figure also displays the hypothetical 'straight way' through

[19] This option value of education was first noted by Weisbrod (1962) and first treated in a dynamic model of educational choice by Comay, Melnik and Pollatschek (1973).

university, i.e. the transition pattern of university students if they were to acquire the same level of highest completed education, but without any dropouts and graduates with excess time-to-graduation: 23% would never enter the university, 22% would only spend three years in the university to complete a Bachelor's degree etc. Roughly speaking, the area between the 'straight way' line and the actual transition line represents time "wasted" at the university. The essential question is whether this excess enrolment time really is wasted? Do individuals learn any other skills although they do not produce course credits? The individuals who waste most time in the educational system are those who never complete university education. The explanation could be that there is high technological uncertainty associated with the educational investment, which makes it valuable for individuals to start investing in a university education in order to get more information about the value of the investment, i.e. ones academic abilities and preferences.[20] It could also be that bad (unskilled) job opportunities induce some individuals who otherwise not would choose to enter the university to start a university education because of the low opportunity costs.[21] These individuals may be more prone to drop out because they get good (unskilled) job offers, e.g. through student employment. In order to answer all these questions and distinguish between the impact from the educational environment and the labor market, the environment within which the educational and employment decisions are made needs to be explicitly modeled.

Figure 3 displays the university-to-work transition separately for individuals choosing each of the five possible education and employment alternatives of the model proposed in Section 4. Initially around 40% of students choose each one of the alternatives involving working less than the equivalent of 10 hours a week, while around 10% of students choose each one of the alternatives involving more working hours. Figure 3 and Table 2 in combination display how the university entrants flow from university education to full-time labor market work through various intermediate transitions between the education and student employment alternatives. This is important for the identification of the parameters of the structural model.

Figure 4 shows the average course credit accumulation over time since university entry for those who are still enrolled at the university. It shows that students who

[20] A detailed discussion of this type of option value created because of technological cost uncertainty of an irreversible potential investment can be found in Dixit and Pindyck (1994).

[21] The (youth) unemployment rate peaked in 1993. In 1994 it was close to 12% and by the end of the sample period it had fallen to around 6%.

work part-time or less do not seem to perform worse academically - if anything they perform better. However, students who work more than part-time have very low academic achievement. The figure further indicates that dynamic selection is important to consider since course credit accumulation seems to be decreasing over time, indicating that those who stay enrolled longer are those with lower academic achievement.

3.3 Correlations in Data

Reduced form results (not tabulated here, but available upon request) confirm previous findings that student employment increases dropout rates and time-to-graduation, while it increases employment probability and earnings - particularly in the early career. As the complete educational and labor market histories of individuals are observed in data, it is possible to control for prior work experience and educational achievement, i.e. stock of labor market and academic capital at university entry. These are not found to affect the impacts of student employment significantly.

First of all, it is noted that neither GPA, high school courses, wealth, assets, liabilities, nor parental wealth and income affect the amount of student employment significantly *ceteris paribus*. On the other hand, prior labor market experience and the average local youth employment rate during enrolment are significantly positively correlated with student employment.

Dropout Rates Estimating a logit model with an indicator for dropping out as the binary dependent variable and labor market experience acquired while enrolled in full-time education as the explanatory variable of primary interest, reveals that marginally increasing student employment increases the probability of dropping out. This estimate is not significantly changed if controlling for proxies for initial abilities (high school GPA and course levels) and other predetermined variables such as age and/or stock of labor market experience at university entry.

Time-to-graduation More student employment is associated with significantly longer time-to-graduation for all academic grade levels. Regressing time-to-graduation on accumulated student employment experience reveals a positive correlation for Master graduates. The estimated correlation is slightly higher when proxies for initial abilities and academic skill set (high school GPA and course levels) are included in the regression. A higher Math level particularly seems to reduce time-to-graduation. This indicates that Master graduates who work part-time are positively selected in terms of observed ability and skills. Student employment seems to increase time-to-graduation more for Bachelor graduates than for Master graduates.

Identical conclusions are obtained by estimating Cox proportional hazard models. This also reveals that working less than part-time significantly reduces the time-to-graduation, while working more than part-time significantly increases the time-to-graduation for Master graduates. The same holds true for Bachelor graduates, while dropouts who work more drop out earlier. This is consistent with the preposition that highly motivated and able students might be more likely to work a moderate amount of hours while enrolled at the university, while forward looking students who perceive their graduation probability as low might be more likely to work more hours.

Highest acquired degree Estimating an ordered logit model where the dependent variable is the highest acquired degree reveals that student employment is negatively correlated with highest completed degree. The marginal effect of student employment on the probability that the highest completed degree is a Bachelor's degree is positive, and on the probability that it is a Master's degree it is negative. These estimates are not significantly changed when also controlling for other observed individual characteristics.

Post Degree Earnings First of all, work experience acquired during university education is found to be more important for university graduates' labor income than work experience acquired earlier. Accumulated student employment experience increases earnings and employment rates one to six years after graduation, but this positive correlation diminishes with time since graduation, and becomes statistically insignificant three years after graduation. The estimated coefficients are slightly, but not significantly, higher when controls for level and/or field of completed university degree are included in the regression. A similar pattern is found for employment rates.

Instrumental Variables Estimation For all estimations, I have instrumented accumulated student employment with average local youth unemployment rate during enrolment, which can be seen as a proxy for students' employment opportunities. The instrument is not that strong - t-statistics around 3.[22] The validity of the instrument relies on the implicit assumption that the average local youth unemployment rate only affects academic and labor market outcomes of each individual through acquired student employment. In accordance with Häkkinen (2006), the estimated effects from these IV estimations are less precisely estimated and smaller in absolute magnitude, however, they have the same sign.

Having established the empirical regularities in the data, the next section presents the structural model that puts these correlations through an in-depth analysis in order to disentangle the channels through which they operate.

4 Basic model setup

Educational and employment choices are made within given institutional settings. This section presents the applied structural estimation approach which requires an explicit specification of the educational environment in terms of grade level progression, choice sets and graduation requirements, as well as the labor market environment in terms of hours and wage opportunities. Although all decisions in the model are taken by individuals and the model is estimated using individual level panel data, the individual i subscripts are suppressed throughout the following section to make notation less cumbersome.

The focus of this paper is on the transition from full-time education to full-time employment conditional on university enrolment. Hence, educational and working hour choices and outcomes are modeled from initial university entry until exit. All students have entered the labor market T years after initial university enrolment.[23] Individuals initially enroll in a university education at time $t = 0$ given ability endowment $A_0 = A$ and skill set $K_0 = K$, accumulated course credits $G_0 = 0$ and consequently formal educational level $E_0 = 0$, and (unskilled) labor market experience H_0. In each post university entry period $t = 1, ..., T$ individuals $i = 1, ..., N$ have the options to stay in full-time education, $D_t = 1$ and supply $h_t \in \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}\right\}$ hours of labor, or to drop out and/or graduate and start working full-time on the labor market, $D_t = 0$ and

[22] According to Staiger and Stock (1997), a good rule of thumb is that the instrument is weak if the t-statistic is below $\sqrt{10}$.

[23] Nine years after university entry only 8% are still enrolled in full-time education, cf. Figure 1. I choose $T = 9$ when solving the model (for the policy simulations).

$h_t = 1$, receiving wages conditional on their stock of human capital. Offered wages depend on formal educational level, E_t, and accumulated labor market experience, $H_t = H_{t-1} + h_{t-1}$:

$$\ln W_t = w\left(H_{t,}E_t, \varepsilon_t^w\right) \tag{1}$$

where ε_t^w is an idiosyncratic labor market productivity shock. All stochastic components are revealed at the beginning of the decision period, for example, individuals know ε_t^w when making their decisions at time $t+1$, but not when making decisions prior to period $t+1$. The econometrician does not observe ε_t^w - neither before nor after decisions are made.

The fact that educational interruptions seem to be more prevalent in Danish data than in US data points towards estimating a discrete choice type model allowing individuals to reenter full-time education after a spell of full-time employment.[24] However, conditional on university enrolment, transitions from full-time work back to full-time university education are very infrequent as only 1% of observed transitions are from full-time employment back to full-time education, cf. Table 2. Therefore, I find it reasonable to treat full-time labor market work as an absorbing state.

The value of university attendance consists of both the current consumption value of education and the potential for increased future earnings. The detailed educational event history data makes it possible to model important institutional features of grade level progression in some detail. Grade level progression depends on the individual's ability, skill set, prior academic achievement, the degree of participation in the labor market, and time since university entry.[25] Individuals who are investing in university education accumulate course credits following the law of motion:

$$G_{t+1} = G_t + g\left(A, K, G_t, E_t, h_{_t}, t, \varepsilon_t^g\right) \tag{2}$$

where ε_t^g is an idiosyncratic academic achievement shock. This can be thought of as unexpected factors that make students particularly much study motivated (or the

[24] For example, Belzil and Hansen (2002) report that 85% in the NLSY sample have never experienced school interruptions while in the Danish data fewer than 30% never experience any interruptions.

[25] The grade level progression function can be thought of as a production function of academic capital, with accumulated course credits measuring the amount of new academic capital acquired. A complete specification of this production function would include the amount and quality of instruction time, the amount of time spent studying and the usage of complementary inputs. Unfortunately the only proxies for time allocation available in the data are the speed of completing a given degree (relative to others completing the same degree) and the amount of labor market work.

opposite), such as a very inspiring lecturer (or sudden illness).

Graduation at level j requires that a given number of credits are acquired, $G_t \geq \overline{G}^j$, $j \in \{1, 2\}$, where level 1 refers to a Bachelor's degree and level 2 to a Master's degree. Since grade level progression is probabilistic, time-to-graduation, $\widetilde{t}_j$, is a probabilistic outcome that can be influenced by employment decisions.

The impact of student employment on grade level progression and earnings is ambiguous as it will depend on the estimated parameters of the model, cf. the discussion in Section 2. Presumably, the less time spent working the more time the student will have to study and the better the academic achievement while in education, $\frac{\partial G_{t+1}}{\partial h_t} < 0$, hence lower dropout probability and/or shorter time-to-graduation with a given degree level. However, to the extent that employment while in education helps develop work habits or merely signals other attributes such as high motivation and/or ability that make potential workers more attractive to employers, expected wages will be positively related to accumulated student employment experience, $\frac{\partial W_{t+\tau}}{\partial h_t} > 0, \forall \tau > 0$ or equivalently $\frac{\partial W_t}{\partial H_t} > 0$.

4.1 Individuals' Optimization Problem

Individuals who are enrolled at the university face the five mutually exclusive and exhaustive alternatives: $(D_t, h_t) \in \left\{(0,1), (1,0), \left(1, \frac{1}{4}\right), \left(1, \frac{1}{2}\right), \left(1, \frac{3}{4}\right)\right\}$. Since the choice set has the same cardinality, $k \in \{0, 1, 2, 3, 4\}$, in each enrollment period, the discrete choices can be defined by the multiple indicator function $d_t = (d_t^0, d_t^1, d_t^2, d_t^3, d_t^4)$ where $d_t^k = \mathbf{1}$ [alternative k chosen at time t]. Individuals make a sequence of discrete choices $\{d_t\}_{t=1}^T$ and are assumed to maximize expected present value of future utility:

$$\max_{d_t^k} E\left[\sum_{\tau=t}^{\infty} \beta^{\tau-t} \sum_{k=0}^{4} d_\tau^k U_\tau^k (S_\tau)\right] \tag{3}$$

for $t \in \{1, ..., T\}$. The state variables $S_t = (X_t, \varepsilon_t)$ include all the information known to the individual at time t and affecting their choices. Let $X_t = (A, K, G_t, E_t, H_t)$ be the state variables observed by both the individual and the econometrician, and $\varepsilon_t = (\varepsilon_t^w, \varepsilon_t^g, \varepsilon_t^0, \varepsilon_t^1, \varepsilon_t^2, \varepsilon_t^3, \varepsilon_t^4)$ be the random elements of the state vector revealed to individuals but not the econometrician.[26] It is assumed that the choice specific preference

[26]Note that the only endogenous state variables are H_t and G_t. Highest acquired degree, E_t, also evolves over time as a consequence of individuals' choices and outcomes, but does so as a surjective

shocks, $\{\varepsilon_t^k\}$, $k \in \{0,1,2,3,4\}$, $t \in \{1,...,T\}$ are independently and identically type I extreme value distributed, $F\left(\varepsilon_t^k\right) = \exp\left(-e^{-\varepsilon_t^k}\right)$. Since the shocks in period t are revealed before choices in that period are made, but are unknown before t, individuals observe the state, S_t, form their expectations about future realizations of the random elements of the state vector, and then make choices, $d_t = (D_t, h_t)$. It is assumed that the current utility of an individual facing state variable S_t from choosing alternative k is additively separable in X_t and ε_t^k, i.e. $U_t^k(S_t) = U_t^k(X_t) + \varepsilon_t^k$. Maximization of (3) is achieved by choosing the optimal sequence of feasible control variables d_t.

The optimization problem (3) can be rewritten as a dynamic programming problem via Bellman's principle of optimality. The value function, $V_t(S_t)$, is defined to be the maximal expected present value at time t, given the state, S_t, and the discount factor, β:[27]

$$V_t(S_t) = U_t^k(S_t) + \beta E\left[V_{t+1}(S_{t+1})\right] \tag{4}$$

It completely summarizes optimal behavior from period t onward, and is a function of a current utility component and a future expected utility component. Consequently, it can be written as $V_t(S_t) \equiv \max_k V_t^k(S_t)$, where $V_t^k(S_t)$ denotes the alternative specific value functions:

$$V_t^k(S_t) = U_t^k(S_t) + \beta E\left[V_{t+1}(S_{t+1})\,|S_t, d_t^k = 1\right]. \tag{5}$$

One great simplification provided by the assumptions of additive separable utility, conditional independence, and i.i.d. type I extreme value distribution of the ε_t^k's is that the $E\max$ in (5) becomes a closed form expression:[28]

$$\begin{aligned} E\left[V_{t+1}(S_{t+1})\,|S_t, d_t^k = 1\right] &\equiv E\left[\max_\kappa V_{t+1}^\kappa(S_{t+1})\,|S_t, d_t^k = 1\right] && (6) \\ &= E\left[\ln\left(\sum_\kappa \exp\left(V_{t+1}^\kappa(X_{t+1})\right)\right)|X_t, d_t^k = 1\right] + \gamma && (7) \end{aligned}$$

where γ is Euler's constant and $V_{t+1}^\kappa(X_{t+1})$ is the expectation of the alternative κ specific value function given current observed state, X_t, and alternative, k. Consequently, these assumptions obviate the necessity of numerically computing multivariate integrals and

function of G_t. Hence, I only have to keep track of the law of motions for H_t and G_t, respectively.

[27] This approach dates back to Bellman (1957). See e.g. Adda and Cooper (2003) or Stokey and Lucas (1989) for an excellent presentation of dynamic programming.

[28] Consult McFadden (1978, 1981) for details and a derivation of this result.

greatly reduce the computational burden.

In order to estimate the model, it is necessary to adopt explicit forms of the wage offers and the wage equation (1), the grade level progression function (2), as well as the value of university attendance. The following subsections provide detailed specifications and discussions of the educational and labor market environment. In this discussion, for parsimony, I assume that error terms are independent across all periods of the model. Then in the estimation Section 4.2 I relax this assumption and allow the error terms to be correlated by allowing for persistent unobserved individual heterogeneity through mixture distributions.

Labor Market Opportunities Wage offers are assumed to depend on accumulated work experience, H_t, highest acquired degree, E_t, and an idiosyncratic labor market productivity shock, ε_t^w. Log wages are given by, $\ln W_t \equiv w_t$, where:

$$w_t = \alpha_0 + \sum_{j=1}^{2} \alpha_j \mathbf{1}\left[E_t = j\right] + \alpha_3 H_t + \alpha_4 H_t^2 + \varepsilon_t^w \tag{8}$$

This choice of modeling is consistent with the notion and empirical evidence of the importance of degrees acquired as opposed to time spent in education, typically known as *sheepskin effects*. Those who acquire a degree are typically found to earn more than those who have attended education for the same amount of time, but failed to acquire a degree. Furthermore, this approach takes the nonlinearities in rates of returns to educational investment into account.[29]

All individuals receive a wage offer each period and then decide their degree of labor market participation. The average number of hours worked per week is proxied by accumulated labor market experience in the year, and can take one of the five discrete values: $h_t \in \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\}$.[30] Each enrollment period, university students choose one of the four alternatives: not to work, $h_t = 0$, or work part-time, $h_t \in \left\{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\right\}$. After university exit, all individuals work full time: $h_t = 1$. The labor market is assumed to be an absorbing state, hence, labor market opportunities are only explicitly specified during university enrolment.

[29] See e.g. Hungerford and Solon (1987), and Jaeger and Page (1996) for evidence on sheepskin effects, and Heckman, Lochner and Todd (2006) for documentation of the importance of nonlinearities.

[30] This is equivalent to the average number of hours worked per week (each week during the year) taking one of the five discrete values: $h_t \in \{0, 10, 20, 30, 40\}$.

Educational environment To successfully complete a year of university education an individual must accumulate 6 course credits (equivalent to 60 ECTS). A course credit is acquired if a passing grade is received in the course. Accumulating a total of 18 course credits (in field f) is the requisite for obtaining a Bachelor's degree (in field f). Having accumulated the 18 course credits it takes to get a Bachelor's degree ($E = 1$) accumulating additional 12 course credits gives a Master's degree ($E = 2$). Hence, acquiring a Bachelor's degree (in field f) gives the student the option to study further to obtain a Master's degree (in field f), which in turn gives the option to study further to obtain a PhD degree.[31] A level j degree is acquired if $G_t \geq \overline{G}^j$, i.e. highest completed degree corresponds to:

$$E_t = \begin{cases} 0 & , \ if \ G_t < 18 \\ 1 & , \ if \ 18 \leq G_t < 30 \\ 2 & , \ if \ 30 \leq G_t \end{cases} . \tag{9}$$

Let g_t be the discrete variable denoting the number of course credits obtained from time t to time $t + 1$, and let G_t denote the total number of course credits accumulated up until time t. Consequently, $G_{t+1} = G_t + D_t g_t$. It is assumed that accumulation of academic capital depends on initial ability, A, and skills, K,[32] as well as acquired university degrees, E_t, previously accumulated course credits, G_t, years since initial university enrolment, t, and hours worked on the labor market, h_t. Accumulated course credits, G_t, capture the self-productivity of course credits, i.e. course credits produced in one period augment academic capital (measured by course credits) attained in later periods.[33] It is implicitly assumed that no incremental academic capital is produced in a course if the student fails it. This choice of modeling can be thought of as if there is an underlying latent variable, g_t^*, determining the number of course credits that reflects

[31] Less than 2% of university students enroll in a PhD program. All PhD students at Danish universities get scolarships corresponding to the starting wage for a well-qualified state employed Master graduate. Furthermore, the PhD study entails a certain amount of predetermined TA and/or RA work. Therefore, I choose to view enrolment in a PhD program as an occupational choice following Master graduation and do not explicitly model school-work choices during this period.

[32] High school GPA is used as a proxy for innate ability, A, and an indicator for whether the student has high level high school Math is used to proxy initial skills, K, since they are found to be the strongest predictors of academic success. Joensen and Nielsen (2008) also find that high level Math has a positive causal impact on earnings that mainly runs through the increased probability of acquiring a higher education. Likewise, Albæk (2006) finds that a higher high school Math level increases the probability of university graduation.

[33] See e.g. Cunha, Heckman, Lochner and Masterov (2006) for evidence and details on the self-productivity of skills in the technology of skill formation.

the incremental academic capital produced in the year. Higher levels of g_t^* mean that the individual has accumulated more academic capital during the year:

$$\begin{aligned} g_t^* &= \gamma_1 A + \gamma_2 K + \gamma_3 \mathbf{1}\left[E_t = 1\right] + \gamma_4 G_t + \gamma_5 t \qquad (10) \\ &\quad + \gamma_6 \mathbf{1}\left[h_t \geq \frac{1}{4}\right] + \gamma_7 \mathbf{1}\left[h_t \geq \frac{1}{2}\right] + \gamma_8 \mathbf{1}\left[h_t \geq \frac{3}{4}\right] + \varepsilon_t^g \\ &= g\left(X_t\right) + \varepsilon_t^g \qquad (11) \end{aligned}$$

Although g_t^* can take many different values, passing each course during the year is only contingent on earning a requisite amount of academic capital, and g_t can only take eight discrete values: $g_t \in \{0, 1, 2, 3, 4, 5, 6, 7\}$.[34] Therefore the grade level progression is modeled in a qualitatively ordered response framework. In order to estimate the model, some assumptions need to be made on the unobserved academic achievement shocks. It is assumed that ε_t^g are i.i.d. logistically distributed. Consequently, the probability of producing $g_t = g$ course credits between time t and $t+1$, $P\left(g_t = g | S_t, d_t^k = 1\right)$ is given by the ordered logit specification for $k \in \{1, 2, 3, 4\}$. Note that an individual who is not enrolled at the university obviously does not accumulate any course credits, and consequently $P\left(g_t = 0 | S_t, d_t^0 = 1\right) = 1$.

The impact of hours worked on accumulated course credits is interpreted as the extent to which student employment is detrimental to academic achievement. Note that if the parameters governing the effect of hours worked on grade level progression are zero, γ_6=γ_7=γ_8=0, student employment has no impact on academic achievement.[35]

[34] I recognize the potential importance of grades and the student's placement in the grade distribution. However, since I do not have data on grades, it is not possible to model the entire grade distribution in more detail. Nevertheless, because of the substantial degree wage premium and because attainment of university degrees only requires accumulation of a requisite number of course credits, I believe that I capture the most essential features of grade level progression. Consult Eckstein and Wolpin (1999) for a similar model with a more detailed modeling of the entire grade distribution.

[35] Note that there might be very different effects on academic achievement depending on how the working hours are distributed over the year. For example, working the equivalent of 20 hours per week on average over the year can be obtained both by working 25 hours a week during each semester, and by working 40 hours a week during every study break. Although I recognize this fact, I only have information on the total amount worked in the year and cannot distinguish between working during the semester and during the breaks. Likewise, there might be very different effects on academic achievement from working in jobs that directly relate to one's field of study, as there might be different opportunities for such jobs depending on one's field (and city) of study. Unlike Ehrenberg and Sherman (1987) I do not find evidence of differential effects on academic performance of working in on- and off-campus jobs, respectively, after controlling for students' ability and skill sets. Furthermore, descriptives do not show large differences across cities, but some fields stand out in the type of student employment. I will return to this issue in Section 6.

Preferences Individuals choose to divide each year ($L = 1$) between units of time devoted to work and non-work. Non-work time is further divided between leisure and education-related activities. Utility of working, $D_t = 0$, equals wages times hours worked, while utility of attending university, $D_t = 1$, also involves getting a study grant, $\overline{b}$. Furthermore, attending the university is assumed to have both an investment and a consumption value. This approach dates back to Heckman (1976) and is common in the literature; see e.g. Keane and Wolpin (1997), Eckstein and Wolpin (1999), Arcidiacono (2004) and Belzil (2007). Savings decisions are not modelled and are implicitly assumed nonexisting.[36] Hence, each period's consumption is assumed to be the sum of earnings and grants. Utility in period t is assumed to be linear and additive in consumption and the value of university attendance, respectively. The per period alternative specific flow utility can be compactly written as:[37]

$$
\begin{aligned}
U_t^k &= W_t h_t + c\mathbf{1}\left[h_{t-1} = 0\right] h_t \qquad (12)\\
&\quad + b_t D_t + \overline{b}\left(W_t h_t, E_t\right) D_t + \varepsilon_t^k, \\
b_t &= b_t^1 + b_t^k d_t^k \mathbf{1}\left[h_t > 0\right], \qquad (13)\\
b_t^1 &= \beta_0 + \beta_1 A + \beta_2 K + \beta_3 t, \qquad (14)\\
b_t^k &= \beta_1^k A + \beta_2^k K + \beta_3^k t + \sum_{h=1}^{3} \beta_{3+h}^k \mathbf{1}\left[h_{t-1} = \frac{h}{4}\right], \qquad (15)\\
\overline{b}\left(W_t h_t, E_t\right) &= \beta_7 \begin{pmatrix} \left(\mathbf{1}\left[W_t h_t < \overline{h}\right] \overline{b} + \mathbf{1}\left[W_t h_t \geq \overline{h}\right] \underline{b}\right) \cdot \\ \left(\mathbf{1}\left[t \leq 4\right] \mathbf{1}\left[E_t = 0\right] + \mathbf{1}\left[t \leq 6\right] \mathbf{1}\left[E_t = 1\right]\right) \end{pmatrix}. \qquad (16)
\end{aligned}
$$

Utility is normalized to consumption units, where labor income is the numeraire, i.e. the coefficient to earnings, $W_t h_t$, is constrained to one.

The job finding cost an individual who was not employed last period must incur in order to become employed after university exit is captured by c. This cost is linear in the amount of work in the current period, i.e. the cost of finding a full-time job and entering the labor market, $h_t = 1$, is twice the cost of staying enrolled and finding a half-time job, $h_t = \frac{1}{2}$.[38]

[36] This rules out pure income effects on behavior. It does not seem to be a restrictive assumption, since the average student in the sample has a very low wealth of 6571 DKK, i.e. approximately \$819 and €880 in real 2000 amounts. Furthermore, there are no significant correlations between student's wealth and academic and labor market choices and outcomes, respectively, in the data.

[37] See Appendix B for the detailed alternative specific specification of flow utillity.

[38] Alternatively, I tried a specification that allowed costs to differ freely over the four alternatives involving employment, $d_t^k c_k \mathbf{1}\left[h_{t-1} = 0\right], k \in \{0, 2, 3, 4\}$, but this didn't produce significantly different

The utility from attending education in any year also depends on the value attached to effort and learning (relative to earnings). Time devoted to education is valued at b_t DKK per unit of time and specified in equations (13), (14) and (15). University attendance involves psychological effort cost, however, learning may also be valued per se. Therefore the consumption value of university attendance can be interpreted as the value attached to learning net the psychological effort cost incurred by studying. b_t is allowed to depend on ability, skills, and time since university entry.[39] Working might also reduce the consumption value of university attendance if it implies increased effort in learning or if it inhibits participation in study-related (social) activities. This effect might depend on initial ability and skills, years since university entry and whether the student has been able to adjust to the joint activity (measured by the degree of labor market participation in the previous period).

All admitted students are eligible to receive a study grant for a maximum of six years. However, the eligibility period is only four years for students who have not yet acquired a Bachelor's degree. The grant depends on the amount of student employment in the year, since there is a penalty in the study grant for working too much. The amounts chosen match the Danish study grant rules for an average earning student in each of the university attendance alternatives are: $\overline{b} = 50,090$, $\underline{b} = 41,741$, and $\overline{h} = \mathbf{1}\left[year < 1996\right] 48,000 + \mathbf{1}\left[year \geq 1996\right] 61,000$.

The alternative specific preference shocks, ε_t^k, capture the fact that new information about individuals' alternative specific tastes is revealed each period. These taste shocks might affect their alternative specific utilities, and are treated as state variables that are unobserved to the econometrician but revealed to individuals just before they make their choices at time t.

4.1.1 Economics of the model

Being enrolled in a university education provides the students with a direct utility, b_t, a *cost* given by foregone earnings and a *return* given by the higher earnings potential. The cost arises because investment in labor market skills is limited to less than full-time employment, $0 \leq h_t < 1$, while attending the university and accumulated labor

results from this restricted specification.

[39] Including time-since-initial-enrolment effects in the consumption value of attending university is the most direct way to fit the t trend in the university attendance and work choice data. Consult Eckstein and Wolpin (1999) for a similar approach and further discussion.

market experience enhances future earnings, $\alpha_3 > 0$, but at a decreasing rate, $\alpha_4 < 0$. The return occurs since graduating with a Bachelor's degree, $E_t = 1$, shifts the wage profile up by $\alpha_1 > 0$. Acquiring a Bachelor's degree is also valuable because it gives the option of obtaining a Master's degree that further shifts the wage profile up by $\alpha_2 > \alpha_1$. This introduces a trade-off between the time opportunity cost of staying enrolled and the degree premium, i.e. between time invested in enhancing labor market skills (experience) and time and effort invested in producing academic capital (course credits leading to degrees) - both of which enhance future wages.

In the basic model, individuals have heterogeneous initial characteristics, A and K, that affect their consumption value of schooling, b_t, as well as their academic achievement, G_t, and accordingly their (future) labor market opportunities through acquired degrees. Furthermore, they also have heterogenous initial characteristics, H_0, that directly affect their labor market opportunities and thereby their opportunity cost of university attendance. In the extended empirical model with different unobserved types, individuals are also innate heterogenous in their consumption value of education, β_0, and academic abilities, γ_0, that affect their university attendance and academic success and thereby their (future) wages, as well as earnings ability, α_0, that directly affects their wages and their outside opportunities while enrolled at the university.

The basic model provides a number of explanations as to why student employment increases monotonically with time since initial enrollment, cf. Figure 1. *(i)* Since wages increase with work experience, the time opportunity cost of university attendance and consequently the probability of working will rise with time since initial enrollment as work experience is accumulated. *(ii)* Heterogeneity in preferences, abilities and skills implies that those who drop out will be a selected sample. If the dropouts are less prone to student employment, then it will appear as if student employment increases with time since initial enrollment.

Individuals will stay enrolled at the university as long as their expected returns are large enough relative to their costs. The three main incentives driving the university exit decision are: When *(i)* grade level progression becomes impossible or *(ii)* a Master's degree is acquired, $E_t = 2$, the investment value of university attendance becomes zero. Therefore, students who stay enrolled after being unable to produce more course credits or having completed a Master's degree do so only because their consumption value of education is higher than the opportunity cost. *(iii)* Students receive no study grant after being enrolled for six years - or four years if they fail to acquire a Bachelor's degree by then. Therefore, the direct monetary value of university attendance decreases after the study grant eligibility period.

Apart from being very important in driving the exit decision, the probability of grade level progression also controls individuals' expectations about future academic achievement, which is the key uncertain component in the state variable transition. Higher grade level progression induces individuals to stay longer in education, but to spend less time in order to successfully acquire a given degree. Completed degrees affect the probabilities of receiving higher wages, hence they also affect the extent of student employment which in turn affects grade level progression. Hence there is a (risk-return) trade-off in the amount of student employment as it increases the immediate utility through earnings and it increases future wages through increasing labor market experience. However, it can be detrimental to academic achievement, which in turn enhances the probability of higher future wages. Staying enrolled but failing to acquire a degree is very costly in the model because there is no change in the academic capital that enhances wages when no degree is acquired.

4.2 Solution and Estimation

The dynamic programming problem (4) can be solved by backward recursion.[40] All individuals float from full-time education at $t = 0$, $\sum_{i=1}^{N} D_{i0} = N$, to full-time employment at $t = T$, $\sum_{i=1}^{N} D_{iT} = 0$. In periods $t \in \{1, ..., T\}$ individuals make educational and employment decisions. Since the labor market is an absorbing state, $D_{t+\tau} = 0$, $\forall \tau \geq 1$, the problem becomes trivial after university exit, and the conditional value function for the full-time working alternative becomes particularly simple:

$$V_t^0 (S_t) = U_t^0(S_t) + \sum_{\tau=t+1}^{\infty} \beta^{\tau-t} E_{\varepsilon_\tau^w} \left[W_\tau \left(X_\tau^w, \varepsilon_\tau^w\right)\right] \tag{17}$$

where $X_\tau^w = (E_t, H_t + \tau - t)$, $\forall \tau \geq t$. Hence, the only state variable that evolves is accumulated labor market experience and it does so deterministically: $H_{t+\tau} = H_t + \tau - t$, $\forall \tau \geq 1$.

[40] See e.g. Eckstein and Wolpin (1989) and Rust (1994) for details on solving and estimating stochastic dynamic discrete choice models.

The conditional value functions for attending the university are given by:

$$
\begin{aligned}
V_t^k(S_t) &= U_t^k(S_t) + \beta \sum_{g=0}^{7} P\left(g_t = g | S_t, d_t^k = 1\right) \cdot \\
&\quad \begin{pmatrix} \mathbf{1}\left[E_{t+1} < 2\right] E\left[V_{t+1}\left(S_{t+1}\right) | S_t, d_t^k = 1\right] \\ +\mathbf{1}\left[E_{t+1} = 2\right] E\left[V_{t+1}^0\left(S_{t+1}\right) | S_t, d_t^k = 1\right] \end{pmatrix}
\end{aligned} \tag{18}
$$

for alternatives $k \in \{1,2,3,4\}$ where $P\left(g_t = g | S_t, d_t^k = 1\right)$ is the probability of producing g course credits between time t and $t+1$ given in Appendix C. The first part in the bracket refers to students who have not yet acquired a Master's degree, and are still eligible to study next period. The second part refers to students that the model forces to exit the university after Master graduation.

The model has to be numerically solved since an analytical solution is not feasible. To minimize the curse of dimensionality and gain computational feasibility, I adopt the powerful simplification first noted by Rust (1987), and assume conditional independence and that the choice specific components of utility that the individuals do not observe before they make their educational and working choices at time t, $\left\{\varepsilon_t^k\right\}$ for $k \in \{0,1,2,3,4\}$, are also the utility components that the econometrician does not observe (either before or after t), and that the individuals make their choices assuming that these components are i.i.d. type I extreme value distributed. Not only do these (admittedly restrictive) assumptions result in a substantial computational gain in terms of solving the dynamic programming problem (4). The conditional independence assumption, the additive nature of the preference shock in the utility function together with the assumption that this shock follows a type I extreme value distribution also provide a simple analytical form for the conditional choice probabilities, so there is no need to simulate multi level integrals. The conditional probability of choosing alternative k is given by:

$$
P\left(d_t^k = 1 | X_t\right) = \frac{\exp\left(U_t^k\left(X_t\right) + \beta E\left[V_{t+1}\left(X_{t+1}\right) | X_t, d_t^k = 1\right]\right)}{\sum_{\kappa=0}^{4} \exp\left(U_t^\kappa\left(X_t\right) + \beta E\left[V_{t+1}\left(X_{t+1}\right) | X_t, d_t^\kappa = 1\right]\right)} \tag{19}
$$

for any alternative $k \in \{0,1,2,3,4\}$. Note that this is the same form as the upper part of a multilevel nested logit.[41] The difficulty in calculating the current choice probabilities arises because they depend on the future stream of expected utilities. To

[41] See e.g. McFadden (1978,1981) for more details.

compute the potential future stream, the utility of all potential state-choice combinations must be determined. In the estimation, I solve this dimensionality problem using the method developed in Hotz and Miller (1993) that relies on a representation of the value function in which future conditional choice probabilities are treated as data rather than functions of the underlying structural parameters. Since data is available on future choices, these probabilities can be calculated from the sample proportions, and equation (7) can be used in the calculation of the probability statements in (19), which are necessary to evaluate the likelihood function. The conditional choice probabilities are then treated as nuisance parameters in the estimation. Hotz and Miller (1993) show that Bellman's equation (5) can always be written as a function of the per-period utilities and future conditional choice probabilities. Note that given the model assumptions, the term inside the $\ln(\cdot)$ in (7) is the denominator for the conditional choice probability of choosing any of the alternatives k given the state, i.e. $\xi_k(X_{t+1}) = P\left(d_{t+1}^k = 1|X_{t+1}\right) = \frac{\exp\left(V_{t+1}^k(X_{t+1})\right)}{\sum_{\kappa=0}^K \exp\left(V_{t+1}^\kappa(X_{t+1})\right)}$. Particularly, this implies that $\ln(\xi_0(X_{t+1})) = V_{t+1}^0(X_{t+1}) - \ln\left(\sum_{\kappa=0}^K \exp\left(V_{t+1}^\kappa(X_{t+1})\right)\right)$ and that the one-period ahead value function conditional on alternative k chosen this period (7) can be written as:

$$
\begin{aligned}
& E\left[V_{t+1}(X_{t+1})|X_t, d_t^k = 1\right] \qquad (20) \\
= \; & \gamma + \sum_{g=0}^{7} P\left(g_t = g|X_t, d_t^k = 1\right) \ln\left(\sum_{\kappa} \exp\left(V_{t+1}^\kappa(X_{t+1})\right)\right) \\
= \; & \gamma + \sum_{g=0}^{7} P\left(g_t = g|X_t, d_t^k = 1\right)\left(V_{t+1}^0(X_{t+1}) - \ln(\xi_0(X_{t+1}))\right)
\end{aligned}
$$

and is solely a function of the flow utility, $U_t^k(S_t)$, the one-period ahead expected value of exiting the university, $V_{t+1}^0(X_{t+1})$, and the one-period ahead CCP of choosing to exit the university, $\xi_0(X_{t+1})$. $\xi_0(X_{t+1})$ can be thought of as the discrete hazard function for exiting the university. This hazard function represents the probability that an individual with ability A, initial skills K, accumulated course credits G_{t+1}, and accumulated work experience H_{t+1} will exit the university in period $t+1$ (given that the individual has stayed in the university education up until then).

The grade level progression probability, $P\left(g_t = g|S_t, d_t^k = 1\right)$, controls individuals' expectation about the one-period ahead state transition and together with the wage equation it also controls the expectation about the one-period ahead value of the full-time labor market alternative. Furthermore, since full-time labor market work is an absorbing state, $V_{t+1}^{0}(X_{t+1})$ simplifies to (17). Finally, this implies that the conditional probability of choosing alternative k that enters the likelihood function is given by substituting (17) and (20) into (19).

This notable simplification is a fortunate product of the additive separability, the conditional independence and the i.i.d. type I extreme value assumptions in the choice specific utility functions, as well as the absorbing labor market state property of the model. In the terminology of Arcidiacono and Miller (2007), the model is said to exhibit the one period dependence (OPD) property, since the current value function only depends on the one-period ahead value of university exit and the probability of choosing to exit the university and start working full time on the labor market.

4.2.1 Estimation

The behavioral parameters of the theoretical model are estimated by a maximum likelihood based procedure. The model basically requires two types of parameters to be estimated: utility function (preference) parameters: $\boldsymbol{\beta}'s$ (and $\boldsymbol{\alpha}'s$), and transition parameters: $\boldsymbol{\gamma}'s$ and $\boldsymbol{\alpha}'s$. The transition parameters are used in forming expectations about uncertain future events, including course credit accumulation and wages. These include the $\boldsymbol{\gamma}$ parameters in the course credit generating process (10) through which students learn about their academic abilities, as well as the $\boldsymbol{\alpha}$ parameters of the wage process (8) that form individuals' expectations about future wages. Note that one important feature of the model is that the wage equation (8) is both part of the law of motion and an important part of utility.[42]

Let $O_{it} = (D_{it}, h_{it}, w_{it}, g_{it})$ denote the vector of observed choices and outcomes for individual i at time t, where the choices are the university attendance indicator and the amount of labor market work, $d_{it} = (D_{it}, h_{it})$, and accepted wages, w_{it}, and accumulated course credits, g_{it}, are the observed outcomes. The likelihood function for the sample of individuals $i = 1, ..., N$ observed from period $t = 0, 1, ..., T_i$ is given by the product over the individual likelihood functions, which is the density for the sequence of observables conditional on the model parameters Θ. Because of the additive separability

[42] This is a standard feature of structural dynamic discrete choice schooling models, see e.g. Belzil (2007) for further discussion and a review of the literature.

and conditional independence assumptions, individuals' likelihood contribution can be decomposed into a product of conditional and marginal densities for each transition. With independent errors across each of the outcomes, the likelihood function can be broken into the three pieces:

$\mathcal{L}_w(\boldsymbol{\alpha})-$ the likelihood contribution of wages

$\mathcal{L}_g(\boldsymbol{\gamma})-$ the likelihood contribution of grade level progression

$\mathcal{L}_d(\boldsymbol{\theta})-$ the likelihood contribution of utility (preferences and choices)

where $\boldsymbol{\theta}=(\boldsymbol{\gamma},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\xi})$. The sample log likelihood function is then the sum of these three components:

$$\begin{aligned}\ln\mathcal{L}(\Theta) &= \ln\prod_{i=1}^{N}\left(\mathcal{L}_g(\boldsymbol{\gamma})\times\mathcal{L}_w(\boldsymbol{\alpha})\times\mathcal{L}_d(\boldsymbol{\theta})\right) \qquad (21)\\ &= \sum_{i=1}^{N}\left(l_g(\boldsymbol{\gamma})+l_w(\boldsymbol{\alpha})+l_d(\boldsymbol{\theta})\right).\end{aligned}$$

Note that the entire set of model parameters enters the likelihood through the choice probabilities and that subsets of the parameters enter through the other structural relationships as well - $\boldsymbol{\alpha}$ through the wage equation and $\boldsymbol{\gamma}$ through the course credit production function. Given the additivity of $\ln\mathcal{L}(\Theta)$, estimation could be carried out by fast sequential maximum likelihood. Since I have data on student employment choices, accumulated course credits, and wages, I can consistently estimate the $\boldsymbol{\gamma}$ and $\boldsymbol{\alpha}$ parameter vectors by maximizing $\mathcal{L}_g$ and $\mathcal{L}_w$ separately. Then using the $\hat{\boldsymbol{\gamma}}$ and $\hat{\boldsymbol{\alpha}}$ parameter estimates, I can consistently estimate the preference parameter vectors $\boldsymbol{\beta}$ by maximizing $\mathcal{L}_d(\hat{\boldsymbol{\gamma}},\hat{\boldsymbol{\alpha}},\boldsymbol{\beta},\boldsymbol{\xi})$. Estimating the parameters stepwise rather than jointly saves significant computational time. The resulting inconsistent standard errors for the preference parameters, due to the estimation error, could be corrected with one Newton step over the whole likelihood, cf. Rust (1994). The rate of time preference is fixed at $\beta=0.95$ in all the estimations.

Note that (additive separability and conditional independence imply that) if there is no unobserved individual heterogeneity in the model, the final estimation step reduces to estimating a multinomial logit of current choices on current flow utility and the discounted one-period ahead expected value of entering the labor market as well as the conditional probability of this event.

4.2.2 Individual Heterogeneity

Given the diversity of individual characteristics at university entry, it is unlikely that individuals have the same sets of preferences for education, skills, abilities and motivation (with respect to educational and labor market work). Hence, it seems important to account for persistent heterogeneity in multiple traits that may themselves be related. A common approach in the literature is to treat these initial traits as unmeasured and drawn from a mixture distribution; see e.g. Keane and Wolpin (1997), Eckstein and Wolpin (1999), and Arcidiacono (2004). It is assumed that there is a fixed number of discrete types of individuals who differ in the parameters that describe their preferences, their academic ability and motivation, and their labor market ability. I adopt this nonparametric approach introduced by Heckman and Singer (1984) and allow for a finite mixture of M types. Each type comprises a fixed proportion, π_m, $m \in \{1, ..., M\}$, of the population.

This way of accounting for unobserved heterogeneity allows for flexible correlation of the errors across the various alternatives as well as correlation over time. This approach also allows me to address two central question: Firstly, who drops out of and who spends a lot of excess time in university education? How do these individuals differ from other graduates (taking the 'straight way') in terms of unmeasured persistent initial traits and how are those traits related to observed family (and other) background characteristics? Secondly, why do these individuals drop out and/or spend so long studying? Which initial traits (if any) are important in explaining the propensity to drop out and/or the excess enrolment time?

In the estimation, wage offers are allowed to differ by unobserved type reflecting persistent differential labor market skills, $\alpha_0 = \sum_{m=1}^{M} \alpha_{0m} \mathbf{1}\left[type = m\right]$, in equation (8). Persistent academic abilities and/or motivation are also allowed to differ by type by introducing, $\gamma_0 = \sum_{m=1}^{M-1} \gamma_{0m} \mathbf{1}\left[type = m\right]$, in equation (10). Likewise, the consumption value of attending the university, $\beta_0 = \sum_{m=1}^{M} \beta_{0m} \mathbf{1}\left[type = m\right]$, in equation (13) is allowed to differ by unobserved type. The likelihood function becomes a finite mixture (or weighted average) of the type-specific likelihoods. Hence, every given type is described by a vector of parameters that are given to them at the time of university entry, corresponding to their labor market skills, academic abilities and/or motivation, and their preferences for university attendance.

To conserve on parameters and avoid identification issues, I consistently only allow for first-order heterogeneity effects. This approach is common in the literature; see e.g. Eckstein and Wolpin (1999). However, there might be type-specific effects of student employment if it for example is more valuable for types who are more likely to drop out.

4.2.3 CCP Estimation with unobserved heterogeneity

The model with unobserved heterogeneity is estimated using the procedure of Arcidiacono and Miller (2007). They extend the class of CCP estimators by adapting the application of the EM algorithm to sequential likelihood developed in Arcidiacono and Jones (2003) to CCP estimators based on Hotz, Miller, Sanders and Smith (1994).

Apart from the huge gain in computational time, two important advantages of this approach are: its ability to account for the role of unobserved heterogeneity in dynamic selection since unobserved heterogeneity can be incorporated into both the flow utility functions and the transition functions *and* its applicability to large populations that are partitioned by unobserved proportions like earnings and academic ability and/or value of university attendance.

Let $\mathcal{L}\left(O_{it}|X_{it}, type_i = m; \theta, \pi, \xi\right)$ be the likelihood of observing choices and outcomes O_{it} for individual i at time t conditional on facing state variable $(S_{it}, type_i = m)$, structural parameters θ and nuisance parameters ξ. The likelihood of any given path of choices and outcomes $O_i = (O_{i1}, ..., O_{iT_i})$ conditional on the observed state $X_i = (X_{i1}, ..., X_{iT_i})$ sequence and unobserved type m, is obtained by forming the product over the T period likelihoods. The sample log likelihood is thus given by:

$$\ln \mathcal{L}\left(\Theta\right) = \sum_{i=1}^{N} \ln \left(\sum_{m=1}^{M} \prod_{t=1}^{T_i} \pi_m \mathcal{L}\left(O_{it}|X_{it}, type_i = m; \theta, \pi, \xi\right) \right). \tag{22}$$

Directly maximizing the log likelihood can be very costly in computational time. However, the EM algorithm simplifies this optimization problem substantially by reintroducing additive separability in the log-likelihood functions. An alternative to maximizing (22) directly is to iteratively maximize the expected log likelihood function, where the n^{th} iteration involves maximizing:

$$\ln \mathcal{L}\left(\Theta\right) = \sum_{i=1}^{N} \sum_{m=1}^{M} \sum_{t=1}^{T} q_{im}^{(n)} \ln \mathcal{L}\left(O_{it}|X_{it}, type_i = m; \theta, \pi^{(n)}, \xi^{(n)}\right) \tag{23}$$

with respect to θ to obtain $\theta^{(n)}$. At each maximization step the probabilities of being each of the unobserved types, $\pi^{(n)} = \left(\pi_1^{(n)}, ..., \pi_M^{(n)}\right)$, are taken as given. $q_{im} \equiv q_{im}\left(O_i, X_i; \theta^{(n)}, \pi^{(n)}, \xi^{(n)}\right)$ is formally defined in (24) and denotes the probability that individual i is of type m given the parameter values $\left(\theta^{(n)}, \pi^{(n)}, \xi^{(n)}\right)$ and conditional on all the data on i's choices, outcomes and characteristics $(O_i, X_i) = (O_{i1}, ..., O_{iT_i}, X_{i1}, ..., X_{iT_i})$. These conditional probabilities of being each of the unobserved types are taken as given and are used as weights in the maximization step. Finally, $\xi^{(n)}$ is the vector of conditional choice probability estimates plugged into the n^{th} iteration and updated as described below in (26).

The estimation algorithm is triggered by setting initial values for the CCPs, $\xi^{(1)}$, the sample proportion of each unobserved type, $\pi^{(1)}$, and initial values for the structural parameters, $\theta^{(1)}$. The values for $\theta^{(1)}$ and $\xi^{(1)}$ are obtained from estimating the model without any unobserved heterogeneity, cf. Section 4.2 above. Each iteration in the algorithm has four steps. Given $\left(\theta^{(n)}, \pi^{(n)}, \xi^{(n)}\right)$ it is proceeded as follows:

Step 1 Compute $q_{im}^{(n+1)}$, the probability that each individual i is of type m for all M types conditional on all the data (O_i, X_i) and given parameter values $\left(\theta^{(n)}, \pi^{(n)}, \xi^{(n)}\right)$ as:

$$q_{im}^{(n+1)} = \frac{\prod_{t=1}^{T} \pi_m^{(n)} \mathcal{L}_{imt}\left(O_{it}|X_{it}, type_i = m; \theta^{(n)}, \pi^{(n)}, \xi^{(n)}\right)}{\sum_{m=1}^{M} \prod_{t=1}^{T} \pi_m^{(n)} \mathcal{L}_{imt}\left(O_{it}|X_{it}, type_i = m; \theta^{(n)}, \pi^{(n)}, \xi^{(n)}\right)}. \tag{24}$$

Note the the numerator is the same across all time periods and that the denominator is the same across all time periods and all types. This is essentially Baye's rule, since the denominator is the likelihood of observing the choice and outcome sequence O_i conditional on the observed state sequence X_i for given parameters, and the numerator.

Step 2 Given $q_{im}^{(n+1)}$, the population fraction of type m, $\pi^{(n+1)}$, is updated by averaging the conditional type probabilities for each individual i and each type m over the sample:

$$\pi_m^{(n+1)} = \frac{1}{N} \sum_{\iota=1}^{N} q_{im}^{(n+1)}. \tag{25}$$

Step 3 Given $q_{im}^{(n+1)}$ for each individual i and each type m, together with $\pi^{(n+1)}$ and $\xi^{(n)}$, maximize the expected log likelihood function (23) to obtain estimates of $\theta^{(n+1)}$.

Step 4 Update the conditional choice probability nuisance parameters $\xi^{(n+1)}$ for each individual i and each type m using the conditional likelihood of observing choice $k = 0$ for state variable $(X, type = m)$ when the parameters are $\left(\theta^{(n+1)}, \xi^{(n)}\right)$:

$$\xi_{0Xm}^{(n+1)} = P\left(d_{it}^{0} = 1 | X, m; \theta^{(n+1)}, \xi^{(n)}\right) = \mathcal{L}_0\left(X, m; \theta^{(n+1)}, \xi^{(n)}\right) \tag{26}$$

Arcidiacono and Miller (2007) show that this algorithm converges to a fixed point and is computationally feasible for many problems with the finite time dependence property. Another computational advantage is that the estimation step can be made sequential, since given the probability that individual i is of type m, q_{im}, the likelihood factors as in the case without unobserved heterogeneity (21).

4.3 Identification

Identification of the wage offer and grade level progression functions rests on variation in earnings, hours and course credit data. The problem of identification can be viewed as a sample selection problem since wages are only observed for individuals who choose to work and course credits are only observed for those who are enrolled in a university education. The exclusion restrictions, the functional form, and distributional assumptions embedded in the model serve the same purpose as would a sample selection correction in a two-step or full information estimation procedure.

Regarding the exclusion restrictions, A and K only affect grade level progression, g_t, and do not directly affect wage offers, w_t, - other than through accumulated course credits, G_t, and accumulated work experience, H_t, which they affect indirectly through affecting grade level progression and the consumption value of university attendance, b_t, as well as how this value is affected by the amount of student employment, $b_t^k, k \in \{2, 3, 4\}$.

The $\boldsymbol{\alpha}$ parameters are identified from data on wages and the state variables: highest completed academic degree, E_t, and acquired labor market experience, H_t. Unobserved heterogeneity, α_{0m}, is identified by cross-sectional variation in wages conditional on these state variables at each t. The $\boldsymbol{\gamma}$ parameters are identified from course credit data and the state variables: ability A, skills K, accumulated course credits, G_t, and hours worked during the period, h_t, and the unobserved heterogeneity, γ_{0m}, is identified by cross-sectional variation in acquired course credits conditional on these state variables.

The remaining utility function parameters, $\boldsymbol{\beta}$, are identified based on the principle of revealed preferences. If students behave myopically (i.e. the model is static), then identification of the utility function parameters would come from observing their school-work choices and accepted wages. The dynamic optimization problem resembles a static multinomial logit model with the future component of the value function treated as a known quantity based on the estimated wage parameters, $\boldsymbol{\alpha}$, and the course credit production parameters, $\boldsymbol{\gamma}$, that control the expectation of next period's state variable for given discount factor, and the CCPs, $\boldsymbol{\xi}$, that are treated as nuisance parameters.

5 Basic Model Estimation Results

This section first discusses some of the parameter estimates and their implications for student behavior and outcomes. Second, unobserved student types are related to observed family background characteristics, and the implications of the estimates for dropout rates and excess time-to-graduation behavior are discussed. Third, evidence on the basic model fit to regularities in the data is presented. Finally, the effectiveness of various public policy interventions aimed at improving academic performance are discussed.

5.1 Parameter Estimates

Estimates of the parameters of the structural model are presented in Table 3. The top panel of the table concerns the wage equation parameters, the middle panel the grade level progression parameters, and the bottom panel concerns the remaining utility parameters that only enter the choice probabilities. The parameter estimates from the basic model are presented in column one, and column two presents the parameter estimates from the basic model with unobserved heterogeneity through a mixture of two types.

The parameter estimates reveal that it does pay off to invest in academic capital, as there are sizeable wage premiums to completing a Bachelor's and Master's degree of 16 and 40 percentage points, respectively, compared to university entrants who fail to acquire a university degree. The usual concave impact of labor market experience on wages is also found. Note that the return to an extra year of experience is 9 percentage points, which is very high, but not surprising given the selective sample of relatively highly skilled individuals in their early career.

Students with higher initial ability and skills from high school have higher academic achievement. Likewise, there is evidence of self-productivity of academic skills as both accumulated course credits and acquired degrees have a significantly positive effect on the number of course credits produced in the year. The latter also indicating that those who exercise the option of Master study after Bachelor graduation are those with higher academic achievement. The longer time since initial enrolment the lower academic achievement *ceteris paribus*. Regarding students' employment, it is found that working the equivalent of 10 hours a week significantly increases academic achievement, while working additional hours is detrimental. Working the equivalent of 10 hours per week significantly reduces the probability of not passing any course in the academic year by 7 percentage points, i.e. $P(g_t = 0|\bar{X}, \mathbf{1}\left[h_t \geq \frac{1}{4}\right] = 1) - P(g_t = 0|\bar{X}, \mathbf{1}\left[h_t \geq \frac{1}{4}\right] = 0) = -0.074$, and it increases the probability of passing all the courses in the year by 3 percentage points, i.e. $P(g_t = 6|\bar{X}, \mathbf{1}\left[h_t \geq \frac{1}{4}\right] = 1) - P(g_t = 6|\bar{X}, \mathbf{1}\left[h_t \geq \frac{1}{4}\right] = 0) = 0.032$. Working additional 10 hours per week significantly increases the probability of accumulating zero course credits by 5 percentage points, i.e. $P(g_t = 0|\bar{X}, \mathbf{1}\left[h_t \geq \frac{1}{2}\right] = 1) - P(g_t = 0|\bar{X}, \mathbf{1}\left[h_t \geq \frac{1}{2}\right] = 0) = 0.053$, and reduces the probability of producing six course credits by two percentage points, i.e. $P(g_t = 6|\bar{X}, \mathbf{1}\left[h_t \geq \frac{1}{2}\right] = 1) - P(g_t = 6|\bar{X}, \mathbf{1}\left[h_t \geq \frac{1}{2}\right] = 0) = -0.023$. Hence, working the equivalent of 20 hour per week reduces the probability of not passing any course by 2 percentage points and increases the probability of passing all courses by 1 percentage point. Working more than 20 hours a week has a large negative impact on academic achievement as it increases the probability of failing all the courses in the year by 38 percentage points and reduces the probability of passing all courses by 14 percentage points.

The consumption value of university attendance is found to be increasing in initial ability and Math skills, while it decreases with time since initial enrolment *ceteris paribus*. Student employment is found to decrease the consumption value of university attendance, and this effect primarily goes through lower consumption value for students with higher initial ability and skills, $\beta_1^k, \beta_2^k < 0$, but not for students who have been enrolled at the university for longer, since $\beta_3^k = 0$. However, if the student also worked in the previous year and thereby has been able to adjust to the joint study-work activity, the value of university attendance is significantly higher, $\beta_4^k, \beta_5^k, \beta_6^k > 0$ for $k \in \{2, 3, 4\}$.

The estimates of the parameters in the wage equation revealed that student employment increases wages through acquired experience. The estimate of the job finding cost, c, further reveals that individuals who were not employed in the previous period have significantly lower probability of entering the labor market and start working full time. Finally, a higher level of study grant tends to increase the likelihood that students will stay enrolled at the university.[43]

The model that allows for unobserved heterogeneity through a mixture of two unobserved types reveals that type 1 students have slightly lower (unskilled) labor market ability, $\alpha_{01} < \alpha_{02}$, but the difference is not significant. On the other hand, type 1 students have higher unobserved academic ability and/or motivation, $\gamma_{01} > 0$, and tend to also have a higher consumption value of university attendance, $\beta_{01} > \beta_{02}$. Introducing unobserved heterogeneity does not affect most of the estimates of the model parameters. The only significant change is that the effect of initial ability and skills on the consumption value of education diminishes or disappears. This is not surprising, since type 1 students have significantly higher GPA and Math level from high school.

5.2 Parental Background and Types

Types are treated as unobserved (to the econometrician) in the estimation. However, each individual can be assigned a set of type probabilities, $q_i = (q_{i1}, ..., q_{iM})$, by applying Baye's rule to each individual's contribution to the likelihood function as discussed in the estimation section, cf. equation (24). Family background and other socioeconomic background data observed prior to university entry can then be merged with the estimation sample and related to type probabilities. This approach gives a sense of how family background affects individuals' preferences for university education.

Type 1 individuals who seem to be academic types comprise 28% of the sample. Type 1 individuals are more likely to have mothers with higher education and higher income; however, their fathers do not seem to have significantly different observable characteristics.

Unobserved types play a significant role in explaining dropout and excess time-to-graduation behavior. Table 4 displays marginal effects from logit estimation on four measures of academic achievement. All estimations are first performed only with

[43] Note that β_7 should in theory be identified, but in practice there seems to be too little variation in study grant, $b(W_t h_t, E_t)$, for a given alternative. The reason is that for a fixed h_t there is not much variation in wages from student employment - not even when incorporating the change in treshold for allowed student earnings. Therefore it becomes difficult to distinguish the consumption value effect captured by β_0 from the study grant effect captured by β_7.

parental background characteristics as controls and then the type 1 probability is added as an additional control variable. The four academic outcomes are indicator variables for: *(i)* dropping out of the university, *(ii)* acquiring a Master's degree, *(iii)* spending excess time to Master graduation, and *(iv)* spending more than one excess year to Master graduation. Table 4 reveals that individuals that have a higher type 1 probability have a significantly lower probability of dropping out, a significantly higher probability of Master graduation, and are also more prone to spend excess time-to-graduation. This corroborates the fact that type 1 individuals have higher unobserved academic ability and/or motivation as well as higher consumption value of university attendance. A Wald test of the joint statistical significance of parental background characteristics is performed for all estimations. The tests show that parental background jointly affects the dropout probability, but when controlling for the type 1 probability, the parental background effect is only jointly significant on a 5% level of significance. Likewise, the effect of parental characteristics on Master graduation is significant on a 10% level of significance, but becomes insignificant when controlling for the type 1 probability. Apparently, the unobserved types embody some predictive information about academic achievement that is contained in parental background characteristics. Furthermore, there seem to be some unobserved individual traits that are predictive of academic achievement and not fully captured by family background characteristics, since the inclusion of unobserved types greatly increases the accuracy of predicting these academic outcomes.

5.3 Model Fit

To assess whether the estimated model captures the essential features of the data, the observed and the predicted choice distributions, transitions, dropout rates, times-to-graduations and wages are compared.

Table 5 compares observed and predicted measures of academic and labor market success. The upper part of the table reveals that the model is very precise in predicting the grade level progression probabilities, $P(g_t = g)$, $g \in \{0, 1, ..., 7\}$, as well as the average accumulation of course credits over time, $\bar{g}_t$, $t \in \{0, 1, ..., 9\}$. The basic model also does a good job in predicting the total amount of course credits accumulated, G_{10}, and highest acquired degrees, E_{10}, by the end of the sample period. At last it is seen that both the level of predicted course credits for those who are enrolled at the university and the level of wages after university exit are slightly lower than their observed counterparts.

Table 6 compares observed and predicted choice probabilities - both overall and over time. The overall choice probabilities are almost point on - as are the probabilities of choosing alternatives $(D_t, h_t) \in \left\{\left(1, \frac{1}{4}\right), \left(1, \frac{3}{4}\right)\right\}$ over time. The model underestimates the probability of initially attending the university and not working while overestimating this probability in the later enrolment years. The basic model underestimates the amount of dropouts in the first three years of university enrollment, while it overestimates the amount of students that will drop out right around the years for 'normal' Bachelor and Master graduation. However, when unobserved heterogeneity is introduced the model predicts the timing of university exit very precisely.

5.4 Policy Effects

Having assessed the model fit and determined the extent to which student employment affects academic and labor market success, it would be interesting to evaluate the potential effects that public policies will have on students' choices and outcomes. Although the impact of student employment on academic achievement has been the subject of numerous studies and is still much debated, not much is known about the impact of potential policy interventions. With the structural parameter estimates at hand, it would be possible to determine the extent to which restrictions on student employment would affect the dropout rates and times-to-graduation - either directly by putting restrictions on hours worked while attending university or indirectly by changing the threshold and/or amount of student benefits received if working. The fact that study grants tend to increase the probability of attending the university, cf. Table 3, suggests that policies aimed at reducing study grants and/or introduce tuition costs could reduce times-to-graduation. However, the credibility of these policy simulations would not be high given the present practical problems of estimating the parameter on received study grant. Therefore, providing evidence on such very much coveted policy effects is left for future research.

Nevertheless, the data allows for an evaluation of an actual policy implementation that increased the threshold income students were allowed to earn without being punished by a cut in the benefits. Up to 1996, students were allowed to earn $48,000\ DKK$ a year while receiving full benefits. In 1996 this threshold was raised to $61,000\ DKK$

a year, giving students an incentive to work more while enrolled at the university. The amount of benefits received was unchanged at around $48,000\ DKK$ per year during the whole period. Using $Z_t = \mathbf{1}[year \geq 1996]$ as an instrument for the amount of student employment would identify the effect of student employment on accumulated course credits (and earnings) for those students who were induced to work more because of the higher threshold. However, the change in benefit rules had no significant effect on students' labor supply.[44]

6 Heterogeneous Stock of Human Capital

So far this paper has treated the stock of academic capital and student employment experience as homogenous. However, there is a lot of heterogeneity in earnings, dropout rates, times-to-graduation, and the amount and type of student employment across fields of university education. Table 7 describes some of the traits on which university students seem to differ across fields. The fields are ranked in terms of wages - with the fields furthest to the right being the most lucrative. Table 7 shows that average yearly earnings vary from around 150,000 DKK (Humanities / Arts / Education) to almost 270,000 DKK (Health Sciences).[45] Dropout rates range from a low of 9% for Health Science students to 28% for Business students and 34% for Life Science students. Among those who initially choose the two most lucrative fields (Health Science and Natural Science / Engineering), 88% and 70%, respectively, acquire a Master's degree, compared to only 46% and 40% of those who initially choose Business and Humanities / Art / Education, respectively. Furthermore, the reasons for dropping out, the required skill sets, and employment prospects might also differ across fields. Overall, the more lucrative field, the more course credits are accumulated each enrolment year, the shorter times-to-graduation for any given degree. Accumulated course credits per enrolment year range from less than 4 (Humanities / Arts / Education) to almost 5 (Health Sciences). The fact that times-to-graduation seem to decrease with post degree earnings could suggest the importance of option values. Hence, if there is high post graduation employment uncertainty, the value of staying enrolled is higher because of the value

[44] This result is probably one of the practical obstacles to estimating β_7.

[45] Christiansen, Joensen and Nielsen (2007) also find that the risk properties of human capital returns vary considerably across fields of university education.

of waiting for more information on labor market opportunities.[46] Student employment might be an efficient device to resolve this uncertainty.

Business students stand out regarding student employment. On average they work the equivalent of 27% of full-time employment each enrolment year, which is around twice as much as the average students in the other fields. They also seem to have very different types of jobs, as they to a higher extent work in office and medium skilled jobs and seem to have fairly good opportunities of study-related employment in the private sector.[47] However, fewer of them work in public services and in on-campus jobs compared to the other fields. Health Sciences and Natural Science / Engineering students are different from the other students in that more of them, 17%, have high-skilled student jobs.

6.1 Results for Heterogeneous Human Capital Stock

The basic model is estimated on stratified samples of university entrants by field in order to assess the robustness of its predictions for subsamples of students who seem to have very different traits in terms of educational preferences, abilities, skill sets, and propensity to work. Particularly, Business students seem to be very prone to work many hours - both before, during and after attending the university. Table 8 presents the parameter estimates from the basic model estimated separately on each of the six aggregate fields of university study. Overall, there do not seem to be differential effects of student employment over fields, and basically all the conclusions in the previous Section 5 remain. Regarding the wage profiles, the most lucrative field in terms of wages (Health Sciences) stands out by having an extremely steep wage profile, a very high Master's degree premium and an insignificant Bachelor's degree premium. However, we must bear in mind that the dropout rate and the Bachelor graduate rates are also very low in the Health Science field. Regarding academic achievement it is noteworthy that in the three most lucrative fields (Business, Natural Sciences / Engineering, and Health Sciences) higher initial Math skills are more important for grade level progression than higher general ability, while in the three less lucrative fields (Humanities / Arts / Education, Life Sciences, and Social Sciences) higher general ability is more important

[46] See e.g. Dixit and Pindyck (1994) and Hogan and Walker (2007) for a more thorough treatment of this type of option value arising because of investment return uncertainty.

[47] They work to a much higher extent in shops / hotels / restaurants, finance / telecom / transport companies, and in business / consultancy services.

for grade level progression.

6.2 Discussion and Further Research

Many individuals drop out of university education, and many individuals who ultimately complete a university education spend considerably longer time in doing so than the normal length of the program requires (an excess of two years on average).[48] Consequently, few individuals pursue what in some sense reflects the 'straight way' through the educational system. Some just spend a long time completing the started educational spell, while others switch to another field of education. This switching may lead to more flexible labor market conditions and to better adjustment to labor demand. Nevertheless, changing the field of education may result in a waste of resources: the total time spent in education increases, and if the accumulated academic capital is partly field specific and not transferable to other fields, a part of the accumulated academic capital will also be lost. Specificity of knowledge acquired in particular fields and the importance of prerequisites in fields such as mathematics mean that changing field because of new information about preferences, poor academic performance, or other factors may be very costly. This cost of switching may be even higher in e.g. Denmark than e.g. the US, since most accumulated course credits are not transferable between fields. Hence, changing fields of study during university enrolment (most often) means that one has to start all over again with zero course credits. As Bamberger (1986) also points out, to the extent that individuals do not know if they will be able to and will want to complete a program in a particular field, they must take into account the alternative options that starting a particular program of study might lead to. An interesting extension of the basic model presented in this paper, which introduces choice of field of university education, would be suitable for estimating the degree of substitutability and complementarity of specific skills and academic capital across fields.

The literature treating education as a sequential choice under uncertainty has primarily focused on the choice of level of education, hence treating the stock of human capital as homogenous. Also Keane and Wolpin (1997) treat educational levels as homogenous, although they allow for endogenously determined occupation-specific ex-

[48] The average public cost per student per year in higher education was 125,000 DKK in 2004 and has been fairly stable over the sample period cf. Statistics Denmark and OECD, Education at a Glance (2005). Most of this public cost is for student grants and "taximeter" rates to the educational institution for completed courses and degrees. These rates also vary across fields.

perience. However, it might be important to allow for heterogeneity in the stock of academic capital within levels of education. The numerous transitions between fields also suggest that individuals are uncertain about choice of field of university education. The earnings differences also seem to be even larger across fields of education than across levels, cf. Table 1 and Table 7. Furthermore, the initial endowments of individuals choosing different levels and fields seems to differ (in nontrivial ways).

The present paper is silent about switching costs and specificity of academic capital. Quantifying these effects is beyond the scope of this paper, but would be an interesting subject of future research. There are some studies that quantify the roles of preferences and abilities on choice of college major. Altonji (1993) provides a theoretical model of sequential choice of education in a three period model with three possible education levels: high school, some college and college graduate, and two different college fields: languages and math. He finds that abilities, high school courses, and preferences affect the expected return to college. Arcidiacono (2004) provides structural estimates of a dynamic model similar in structure to Altonji (1993). He allows individuals to choose college quality and field and finds that differences in the monetary returns to education explain little of the ability sorting across fields of major, since it is mainly driven by preferences for particular majors. His model is a three period model with no leeway about the timing of education investment. An extension of the basic model in this paper that allows students to choose field of university education would recognize the fact that investment in education takes time, hence, the time-to-graduation would be stochastic. Another important contribution of this extension would be that neither Altonji (1993) nor Arcidiacono (2004) allow for student employment. In their model, students are either in full-time education working zero hours in the labor market or working full-time on the labor market. This might also be important to the extent that student employment helps resolve technological cost and/or return uncertainty.[49]

7 Conclusion

Despite the fact that reducing dropout rates and times-to-graduation have been declared social goals for many years in many countries, current research does not provide much

[49] Extending the model in this paper to allow for choice of field of university education (and switching between fields) would increase the number of choices by $5 * \#fields$ and the computational burden of solving the model above the current computer power constraint.

evidence on how to obtain these social goals through public policies. This paper provides an in-depth analysis of the channels through which student employment, abilities and preferences affect academic achievement while attending the university and how these in turn affect wages. A thorough understanding of these matters is pivotal in order to be able to construct public policies to achieve the posed social objectives.

The structural model in this paper explicitly takes the simultaneous and sequential nature of educational and student employment decisions and the uncertainty of academic outcomes into account. Furthermore, it allows for unobserved heterogeneity in both utility and transition equations and is thus able to control for (dynamic) self-selection. Estimation of the model reveals considerable positive impacts of observed abilities and skills on academic achievement. Types of students with high and persistent unobserved academic ability and/or motivation and high consumption value of university attendance are much less prone to drop out, much more prone to graduate with a Master's degree and to spend more time obtaining the degree. The latter outcomes would be difficult to alter through public policies - other than policies aimed at changing parental characteristics that seem to affect these persistent student traits. However, the positive impacts of observed abilities and skills might be suggestive of the effectiveness of policies that enhance students' cognitive skills at younger ages.

More than part-time student employment is found to be very detrimental to academic achievement. On the other hand, student employment of a moderate number of hours (less than 20) a week significantly increases academic achievement and future labor market outcomes. Therefore, it can be considered an effective device through which students complement their university education and smoothen their transition from full-time education to full-time labor market work.

References

[1] Adda, J. and R. W. Cooper (2003), "Dynamic Economics: Quantitative Methods and Applications", The MIT Press, Cambridge, MA.

[2] Altonji, J. (1993), "The Demand for and Return to Education When Education Outcomes are Uncertain", *Journal of Labor Economics*, 11(1): 48-83.

[3] Albæk, K. (2006), "Evner, fag, indsats i gymnasiet og beståelse på Samfundsvidenskab", Memo no. 213, Department of Economics, University of Copenhagen.

[4] Arcidiacono, P. (2004), "Ability Sorting and the Returns to College Major", *Journal of Econometrics*, 121(1-2): 343-375.

[5] Arcidiacono, P. and J. B. Jones (2003), "Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm", *Econometrica*, 71(3): 933-946.

[6] Arcidiacono, P. and R. A. Miller (2007), "CCP Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity", manuscript, Duke University.

[7] Bacolod, M. and J. V. Hotz (2006), "Cohort Changes in the Transition from School to Work: Evidence from three NLS surveys", *Economics of Education Review*, 25(4): 351-373.

[8] Bamberger, I. (1986), "Values and Strategic Behaviour", *Management International Review*, 4: 57-69.

[9] Becker, G. (1964), "Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education", Columbia UP: New York. First edition, 1964. Third edition, 1993.

[10] Bellman, R. (1957), "Dynamic Programming", Princeton: Princeton University Press.

[11] Belzil, C. (2007), "The Return to Schooling in Structural Dynamic Models: A Survey", *European Economic Review*, 51(5): 1059-1105.

[12] Belzil, C. and J. Hansen (2002), "Unobserved Ability and the Return to Schooling", *Econometrica*, 70(5): 2075-2091.

[13] Brodaty, T. O., R. J. Gary-Bobo and A. Prieto (2006), "Does Speed Signal Ability? The Impact of Grade Repetitions on Employment and Wages", manuscript, University of Paris I.

[14] Brunello, G. and R. Winter-Ebmer (2003), "Why do students expect to stay longer in College? Evidence from Europe", *Economics Letters*, 80(2): 247-253.

[15] Cameron, S. and J. J. Heckman (1998), "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males", *Journal of Political Economy*, 106(2): 262-333.

[16] Card, D. (1999), "The Causal Effect of Education on Earnings", Ch. 30 in Ashenfelter and Card, eds., *Handbook of Labor Economics*, North-Holland.

[17] Christiansen, C., J. S. Joensen and H. S. Nielsen (2007), "The Risk-Return Trade-off in Human Capital Investment", *Labour Economics*, 14(6): 971-986.

[18] Comay, Y., A. Melnik, and M. A. Pollatschek (1973), "The Option Value of Education and the Optimal Path for Investment in Human Capital", *International Economic Review*, 14(2): 421-435.

[19] Cunha, F., J. J. Heckman, L. J. Lochner, and D. V. Masterov (2006), "Interpreting the Evidence on Life Cycle Skill Formation", in E. A. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education*, North-Holland, Amsterdam: 697-812.

[20] Dixit, A. K. and R. S. Pindyck (1994), "Investment Under Uncertainty", Princeton University, Princeton, New Jersey.

[21] Eckstein, Z. and K. I. Wolpin (1989), "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models: A Survey", *Journal of Human Resources*, 24(4): 562-598.

[22] Eckstein, Z. and K. I. Wolpin (1999), "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities", *Econometrica*, 67(6): 1295-1339.

[23] Ehrenberg, R. G. and D. R. Sherman (1987), "Employment While in College, Academic Achievement and Post College Outcomes", *Journal of Human Resources*, 23(1): 1-23.

[24] Garibaldi, P., F. Giavazzi, A. Ichino and E. Rettore (2007), "College Cost and Time to Obtain a Degree: Evidence from Tuition Discontinuities. NBER WP 12863.

[25] Heckman, J. J. (1976), "A Life-cycle Model of Earnings, Learning and Consumption", *Journal of Political Economy*, 84(4): 11-44.

[26] Heckman, J. J. (1981): "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time Discrete Data Stochastic Process", in *Structural Analysis of Discrete Data with Econometric Applications*, C. F. Manski and D. McFadden (eds.), The MIT Press, Cambridge, MA: 179-197.

[27] Heckman, J. J., L. J. Lochner and P. E. Todd (2006), "Earnings Equations and Rates of Return: The Mincer Equation and Beyond", in E. A. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education*, North-Holland, Amsterdam.

[28] Heckman, J. and B. Singer (1984); "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data", *Econometrica*, 52 (2): 271-300.

[29] Hogan, V. and I. Walker (2007), "Education Choice under Uncertainty: Implications for Public Policy", *Labour Economics*, 14(6): 894-912.

[30] Hotz, V. J. and R. A. Miller (1993), "Conditional Choice Probabilities and Estimation of Dynamic Models", *Review of Economic Studies*, 60(3): 497-529.

[31] Hotz, V. J., R. A. Miller, S. Sanders and J. Smith (1994), "A Simulation Estimator for Dynamic Models of Discrete Choice", *Review of Economic Studies*, 61(2): 265-289.

[32] Hotz, V. J., L. C. Xu, M. Tienda and A. Ahituv (2002), "Are there Returns to the Wages of Young Men from Working while in School", *Review of Economics and Statistics*, 84(2): 343-375.

[33] Hungerford, T. and G. Solon, "Sheepskin Effects in the Returns to Education", *Review of Economics and Statistics*, 69(1): 175-177.

[34] Häkkinen, I. (2006), "Working while Enrolled in a University: Does it Pay?", *Labour Economics*, 13(2): 167-189.

[35] Jaeger, D. A. and M. E. Page (1996), "Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education", *Review of Economics and Statistics*, 78(4): 733-740.

[36] Joensen, J. S. and H. S. Nielsen (2008), "Is there a Causal Effect of High School Math on Labor Market Outcomes?", forthcoming in the *Journal of Human Resources*.

[37] Keane, M. and K. I. Wolpin (1997), "The Career Decisions of Young Men", *Journal of Political Economy*, 105(3): 473-522.

[38] Light, A. (2001), "In-School Work Experience and the Return to Schooling", *Journal of Labor Economics*, 19(1): 65-93.

[39] McFadden, D. (1978), "Modeling the Choice of Residential Location", in A. Karlqvist, L. Lundqvist, F. Snickars, and J. W. Weibull (eds.), *Spatial Interaction Theory and Planning Models*, North-Holland, Amsterdam: 75-96.

[40] McFadden, D. (1981), "Econometric models of probabilistic choice", in *Structural Analysis of Discrete Data with Econometric Applications*, eds C. F. Manski and D. McFadden (eds.), The MIT Press, Cambridge, MA: 198-272.

[41] Mincer, J. A. (1974), "Schooling, Experience, and Earnings", Columbia UP: New York.

[42] OECD *Education at a Glance* (2005).

[43] Ruhm, C. (1997), "Is High School Employment Consumption or Investment?", *Journal of Labor Economics*, 15(4): 735-776.

[44] Rust, J. (1987), "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher", *Econometrica*, 55(5): 999-1033.

[45] Rust, J. (1994), "Structural Estimation of Markov Decision Processes", in R. Engle and D. McFadden (eds.), *Handbook of Econometrics*, North-Holland, Amsterdam.

[46] Ryan, P. (2001), "The School-to-Work Transition: A Cross-National Perspective", *Journal of Economic Literature*, 39(1): 34-92.

[47] Staiger, D. and J. H. Stock (1997), "Instrumental Variables Regression with Weak Instruments", *Econometrica*, 65(3): 557-586.

[48] Stinebrickner, R. and T. R. Stinebrickner (2003), "Working During School and Academic Performance", *Journal of Labor Economics*, 21(2): 473-491.

[49] Stokey, N. L. and R. E. Lucas (1989), "Recursive Methods in Economic Dynamics", Harvard University Press, Cambridge, MA.

[50] Topel, R. H. and M. P. Ward (1992), "Job Mobility and the Career of Young Men", *The Quarterly Journal of Economics*, 57(2): 439-479.

[51] Velfærdskommisionen (2005), "Fremtidens velfærd - vores valg", December 2005. *www.velfaerd.dk*

[52] Weisbrod, B. A. (1962), "Education and Investment in Human Capital", *Journal of Political Economy*, 70, 106-123.

[53] Willis, R. J. and S. Rosen (1979), "Education and Self-Selection", *Journal of Political Economy*, 87(5): 7-36.

A Appendix

This Appendix clarifies how the timing of educational events and course credit accumulation are coded in the data.

A.1 Timing of Educational Events

Educational event histories are observed montly and accumulated course credits are observed yearly. Likewise are the other individual characteristics and outcomes observed yearly. Hence, I have to take a stance on how to code the timing of educational events. Throughout, it is assumed that individuals make choices at the beginning of the year, whereas outcomes are observed at the end of the year. If the individual is enrolled at the university most of the year then $D_t = 1$, otherwise, $D_t = 0$. Figure 1 illustrates how the educational event histories are coded. Particularly:

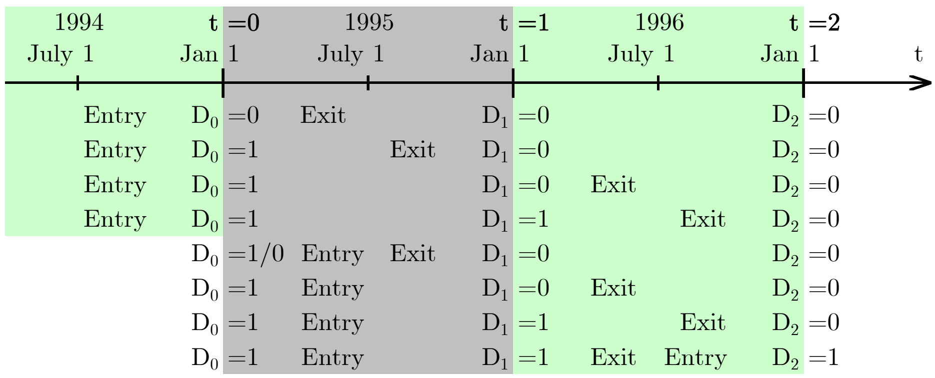
- If university *entry* occurs prior to July 1 in the year starting on time t and ending on time $t+1$, the educational spell is coded to start on time t, i.e. at the beginning of the year,
- whereas it is coded to start at time $t+1$ if entry occurs after July 1, i.e. at the beginning of the following year.
- If university *exit* occurs prior to July 1 in the year starting on time t and ending on time $t+1$, the educational spell is coded to end at time t,
- whereas it is coded to end at time $t+1$ if exit occurs after July 1 in the year.

Two exceptions are made to this rule:

- If both entry and exit occur in the year between time t and $t+1$, then the individual is coded to be in education, $D_t = 1$, if the individual is enrolled at the university for more than 6 months of the year. Otherwise, the individual is coded to be working full-time on the labor market in the year, $D_t = 0$. This exception corresponds to the fifth row in Figure 1.
- If an individual is coded to graduates with a Bachelor degree in June and to start a Master education on September 1 in the same year then I assume that this is

a continued educational spell, hence, $D_t = D_{t+1} = 1$. This exception corresponds to the last row in Figure 1.

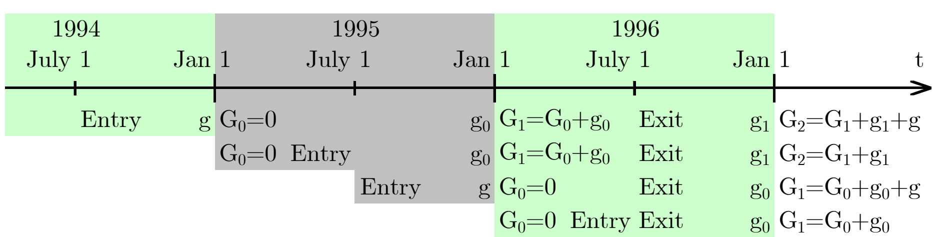
Figure 1: Timing of Educational Events



A.2 Course Credit Accumulation

Accumulated course credits are coded accordingly. Figure 2 illustrates how accumulated course credits in the year are coded based on when the individual enters and exits the university.

Figure 2: Course Credit Accumulation



Corresponding to the first and third row in Figure 2, 92% of individuals enter the University in September, hence are coded to start full-time education from the beginning of the following year. Final exams are usually in January and June, however there are some rare December exams. It is assumed that there are equally many December exams each year of any given university study spell. Hence, the g course credits accumulated in the December exams in the year prior to $t = 0$ are transfered to the course credits accumulated in the following year and further g course credits are transferred to the next year for all $t \in \{1, ..., \tilde{t} - 1\}$, where $\tilde{t}$ is the time of university exit. Hence, for example $G_1 = G_0 + g_0 + g - g = G_0 + g_0$ if university attendance is continued after the first time period, $\tilde{t} > 1$, since g course credits are transfered from the previous period and g course credits are also transfered to the following period. In practice, this means that accumulated course credits at the beginning of all university attendance periods are given by: $G_{t+1} = G_t + g_t$, $\forall t < \tilde{t} - 1$, and then g is added to accumulated course credits at exit: $G_{\tilde{t}} = G_{\tilde{t}-1} + g_{\tilde{t}-1} + g$. In the first two rows of the illustrated course credit accumulation in Figure 2 time to university exit is $\tilde{t} = 2$, and in the last two rows it is $\tilde{t} = 1$. After university exit no course credits are accumulated, and consequently: $G_{t+1} = G_t$, $\forall t > \tilde{t} - 1$.

B Appendix

TAlternative specific flow utilities are specified in their detail in this Appendix. Throughout the Appendix the notation of the underlying paper is adopted. See Section 4 for details.

B.1 Preferences

The flow utility of working full-time on the labor market, $(D_t, h_t) = (0, 1)$, is given by:

$$U_t^0\left(S_t\right) = W_t + c\mathbf{1}\left[h_{t-1} = 0\right] + \varepsilon_t^0,$$

and for attending the university full-time and not work, $(D_t, h_t) = (1, 0)$:

$$U_t^1\left(S_t\right) = \beta_0 + \beta_1 A + \beta_2 K + \beta_3 t + \beta_7 \overline{b} + \varepsilon_t^1.$$

The flow utilities for the three combined university attendance and work alternatives are in their entirety given by:

$$\begin{aligned} U_t^2\left(S_t\right) &= \frac{1}{4}W_t + \frac{1}{4}c\mathbf{1}\left[h_{t-1} = 0\right] \\ &+\beta_0 + \left(\beta_1 + \beta_1^2\right)A + \left(\beta_2 + \beta_2^2\right)K + \left(\beta_3 + \beta_3^2\right)t \\ &+\beta_4^2\mathbf{1}\left[h_{t-1} = \frac{1}{4}\right] + \beta_5^2\mathbf{1}\left[h_{t-1} = \frac{1}{2}\right] + \beta_6^2\mathbf{1}\left[h_{t-1} = \frac{3}{4}\right] \\ &+\beta_7\overline{b} + \varepsilon_t^2, \end{aligned}$$

if the student works the equivalent of 10 hours per week, $(D_t, h_t) = \left(1, \frac{1}{4}\right)$.

$$\begin{aligned} U_t^3\left(S_t\right) &= \frac{1}{2}W_t + \frac{1}{2}c\mathbf{1}\left[h_{t-1} = 0\right] \\ &+\beta_0 + \left(\beta_1 + \beta_1^3\right)A + \left(\beta_2 + \beta_2^3\right)K + \left(\beta_3 + \beta_3^3\right)t \\ &+\beta_4^3\mathbf{1}\left[h_{t-1} = \frac{1}{4}\right] + \beta_5^3\mathbf{1}\left[h_{t-1} = \frac{1}{2}\right] + \beta_6^3\mathbf{1}\left[h_{t-1} = \frac{3}{4}\right] \\ &+\beta_7\underline{b} + \varepsilon_t^3 \end{aligned}$$

if the student works the equivalent of 20 hours per week, $(D_t, h_t) = \left(1, \frac{1}{2}\right)$.

$$\begin{aligned} U_t^4(S_t) &= \frac{3}{4}W_t + \frac{3}{4}c\mathbf{1}\left[h_{t-1} = 0\right] \\ &+\beta_0 + \left(\beta_1 + \beta_1^4\right)A + \left(\beta_2 + \beta_2^4\right)K + \left(\beta_3 + \beta_3^4\right)t \\ &+\beta_4^4\mathbf{1}\left[h_{t-1} = \frac{1}{4}\right] + \beta_5^4\mathbf{1}\left[h_{t-1} = \frac{1}{2}\right] + \beta_6^4\mathbf{1}\left[h_{t-1} = \frac{3}{4}\right] \\ &+\beta_7\underline{\underline{b}} + \varepsilon_t^4 \end{aligned}$$

if the student works the equivalent of 30 hours or more per week during the year, $(D_t, h_t) = \left(1, \frac{3}{4}\right)$. The three levels of received study grant in the year are given by: $\overline{b} = 50,090$, $\underline{b} = 41,741$, and $\underline{\underline{b}} = 0$.

C Appendix

C.1 Evolution of Endogenous State Variables

This Appendix provides some details on the evolution of the two endogenous state variables: accumulated course credits, G_t, and accumulated labor market experience, H_t. The state transition matrix for these two variables is given by:

	$(G_t, H_t) = (G, H)$				
	$d_t^0 = 1$	$d_t^1 = 1$	$d_t^2 = 1$	$d_t^3 = 1$	$d_t^4 = 1$
(G_{t+1}, H_{t+1})	$h_t = 1$	$h_t = 0$	$h_t = \frac{1}{4}$	$h_t = \frac{1}{2}$	$h_t = \frac{3}{4}$
$(G, H + h_t)$	1	P_0^1	P_0^2	P_0^3	P_0^4
$(G + 1, H + h_t)$	0	P_1^1	P_1^2	P_1^3	P_1^4
$(G + 2, H + h_t)$	0	P_2^1	P_2^2	P_2^3	P_2^4
$(G + 3, H + h_t)$	0	P_3^1	P_3^2	P_3^3	P_3^4
$(G + 4, H + h_t)$	0	P_4^1	P_4^2	P_4^3	P_4^4
$(G + 5, H + h_t)$	0	P_5^1	P_5^2	P_5^3	P_5^4
$(G + 6, H + h_t)$	0	P_6^1	P_6^2	P_6^3	P_6^4
$(G + 7, H + h_t)$	0	P_7^1	P_7^2	P_7^3	P_7^4

where the columns refer to current period's state, (G_t, H_t), and choice, d_t, and rows refer to next period's state, (G_{t+1}, H_{t+1}). $P_g^k = p\left(X_{t+1} | X_t, d_t^k = 1\right) = P\left(g_t = g | X_t, d_t^k = 1\right)$ denotes the probability that an individual that faces state variable X_t and chooses alternative k at time t will produce g course credits between time t and $t + 1$. Note that given the current state variable, X_t, next period's state variable, X_{t+1}, can take $(1 + 4 \cdot 8 =)$ 33 possible values. Hence, conditional on period t's choice, all uncertainty regarding period $t + 1$'s state is concerning how many course credits, $g \in \{0, 1, 2, 3, 4, 5, 6, 7\}$, are produced given that one of the educational alternatives, $k \in \{1, 2, 3, 4\}$, is chosen. Consider, for example, an individual i that has been enrolled at the university for four years, has passed all courses within the required amount of time, and accumulated the equivalent of one year of full-time labor market experience, $(G_4, H_4) = (24, 1)$. If i at $t = 4$ chooses to continue the university education and work half-time, $(D_4, h_4) = \left(1, \frac{1}{2}\right)$, the probability that i will graduate with a Master's degree at $t = 5$ is given by $P_6^3 = P\left(g_4 = 6 | X_4, d_4^3 = 1\right)$. Equivalently, $P_6^3 = P\left(\left(G_5, H_5\right) = \left(30, 1\frac{1}{2}\right) | \left(G_4, H_4\right) = (24, 1), d_4^3 = 1\right)$ abstracting from the exogenous state variables.

Table 1: Descriptive Statistics of University Graduates and Dropouts.

	Mean (Standard Deviation)		
	Acquired University Degree		
Individual Characteristics	**Dropout**	**Bachelor**	**Master**
At University Entry:			
High school GPA	8.64 (0.89)	8.89 (0.80)	9.13 (0.72)
High school Math level	1.91 (0.91)	1.90 (0.92)	2.28 (0.86)
High school Science level	1.73 (0.83)	1.72 (0.80)	2.04 (0.84)
High school Social Science level	1.80 (0.89)	1.89 (0.88)	1.77 (0.86)
High school Language level	2.09 (0.80)	2.29 (0.75)	2.21 (0.68)
Female	0.40	0.53	0.50
Had sabbatical after high school	0.63	0.68	0.73
Duration of sabbatical (months)	12.80 (8.89)	14.29 (9.28)	14.46 (8.59)
Accumulated work experience	0.75 (0.68)	0.59 (0.63)	0.55 (0.54)
Age	21.22 (0.77)	21.17 (0.79)	21.14 (0.76)
During University Enrolment:			
Target duration from entry until exit (months)	0	34	58
Duration from entry until exit (months)	71.83 (25.42)	83.08 (32.19)	78.19 (14.22)
Accumulated course credits per year	2.32 (2.49)	3.92 (2.41)	4.70 (2.08)
Accumulated work experience per year (years)	0.30 (0.36)	0.18 (0.25)	0.15 (0.21)
At University Exit:			
Requisite course credits to acquire degree	0	18	30
Accumulated course credits	7.40 (8.87)	25.94 (6.88)	31.18 (5.22)
Accumulated work experience (years)	1.24 (1.23)	1.26 (1.21)	1.26 (0.99)
After University Exit:			
Hourly wages (real 2000 DKK)	147.97 (68.31)	153.51 (77.84)	193.12 (72.38)
Yearly earnings (real 2000 DKK)	138,036 (118,555)	146,533 (120,261)	269,138 (122,284)
Number of Individuals	487	473	1,169
Fraction of total sample	0.23	0.22	0.55
Number of Observations	4,400	4,249	10,700
Total number of Individuals		**2,129**	
Total number of Observations		**19,349**	

Notes to Table 1: The table shows average characteristics of university graduates and dropouts (standard deviation in parenthesis). For indicator variables the fraction of the sample is reported. The descriptive statistics are displayed separately by highest degree acquired post initial university enrolment. The exchange rate on December 31, 2000 was 8.0205 DKK/USD and 7.4631 DKK/Euro.

Table 2: State Transitions.

	# observations (% relative to last periods' choice)					
	C_t					
C_{t-1}	Working full-time	Education h=0	Education 0< h =¼	Education ¼< h =½	Education h >½	All
Working full-time	5,927 (96.96)	62 (1.01)	55 (0.90)	31 (0.51)	38 (0.62)	6,113
Education, h=0	806 (18.00)	2,404 (53.70)	908 (20.28)	246 (5.49)	113 (2.52)	4,477
Education, 0< h =¼	510 (13.00)	878 (22.38)	1,768 (45.07)	597 (15.22)	170 (4.33)	3,923
Education, ¼< h =½	342 (20.85)	239 (14.57)	367 (22.38)	448 (27.32)	244 (14.88)	1,640
Education, h >½	331 (31.02)	176 (16.49)	82 (7.69)	86 (8.06)	392 (36.74)	1,067
All	7,916	3,759	3,180	1,408	957	17,220

Notes to Table 2: The table displays state transitions. Rows concern last periods alternatives, d^j_{t-1}, and columns concern current periods alternatives, d^k_t, for $j, k \in \{0, 1, 2, 3\}$. Hence, each cell refers to the number of individuals choosing alternative j in period $t-1$ and alternative k in period t. The fraction of individuals who choose alternative j in period $t-1$ and then choose alternative k in period t, relative to those who choosing alternative j in period $t-1$, is displayed in parantheses (row % in parentheses).

Table 3: Parameter Estimates.

	One type (M=1)	Two types (M=2)
	w_t	
α_0	4.517 (0.006) ***	
α_{01}		4.509 (0.012) ***
α_{02}		4.521 (0.009) ***
α_1 (Bachelor degree)	0.158 (0.008) ***	0.159 (0.012) ***
α_2 (Master degree)	0.404 (0.010) ***	0.405 (0.015) ***
α_3 (Experience)	0.094 (0.005) ***	0.093 (0.008) ***
α_4 (Experience^2)	-0.002 (0.001) ***	-0.002 (0.001)
	g^*	
γ_{01}		0.176 (0.045) ***
γ_1 (High school GPA=9)	0.216 (0.029) ***	0.212 (0.043) ***
γ_2 (High level Math)	0.276 (0.029) ***	0.273 (0.042) ***
γ_3 (Bachelor degree, $\mathbf{1}[E_t=1]$)	0.208 (0.037) ***	0.207 (0.054) ***
γ_4 (Accumulated course credits, G_t)	0.084 (0.002) ***	0.082 (0.004) ***
γ_5 (Time since enrolment)	-0.618 (0.012) ***	-0.612 (0.017) ***
γ_6 (Student employment, h_t=¼)	0.301 (0.035) ***	0.296 (0.051) ***
γ_7 (Student employment, h_t=½)	-0.215 (0.045) ***	-0.214 (0.066) ***
γ_8 (Student employment, h_t=¾)	-1.590 (0.051) ***	-1.572 (0.075) ***
	$P(d_t^k=1)$	
c (Not employed previous period, h_{t-1}=0)	0.818 (0.077) ***	5.026 (0.000) ***
β_0	-19.939 (0.092) ***	
β_{01}		-22.859 (0.310) ***
β_{02}		-53.73 (380.67)
β_1 (High school GPA=9)	0.603 (0.068) ***	0.323 (0.199) *
β_2 (High level Math)	0.921 (0.068) ***	0.299 (0.195)
β_3 (Time since enrolment)	-3.760 (0.016) ***	-3.357 (0.047) ***
$\beta_1+\beta_1^2$ (High school GPA=9)	-0.054 (0.069)	-0.215 (0.201)
$\beta_2+\beta_2^2$ (High level Math)	0.234 (0.069) ***	-0.309 (0.197)
$\beta_3+\beta_3^2$ (Time since enrolment)	-3.832 (0.017) ***	-3.578 (0.049) ***
β_4^2 (Student empl. previous period, h_{t-1}=¼)	2.764 (0.058) ***	2.136 (0.090) ***
β_5^2 (Student empl. previous period, h_{t-1}=½)	2.389 (0.090) ***	1.934 (0.150) ***
β_6^2 (Student empl. previous period, h_{t-1}=¾)	1.615 (0.124) ***	1.360 (0.221) ***
$\beta_1+\beta_1^3$ (High school GPA=9)	-0.651 (0.074) ***	-0.828 (0.207) ***
$\beta_2+\beta_2^3$ (High level Math)	-0.372 (0.075) ***	-0.967 (0.205) ***
$\beta_3+\beta_3^3$ (Time since enrolment)	-3.837 (0.018) ***	-3.606 (0.051) ***
β_4^3 (Student empl. previous period, h_{t-1}=¼)	2.282 (0.069) ***	1.640 (0.111) ***
β_5^3 (Student empl. previous period, h_{t-1}=½)	3.114 (0.093) ***	2.656 (0.158) **
β_6^3 (Student empl. previous period, h_{t-1}=¾)	2.045 (0.133) ***	1.691 (0.247) ***
$\beta_1+\beta_1^4$ (High school GPA=9)	-0.363 (0.091) ***	-0.077 (0.234)
$\beta_2+\beta_2^4$ (High level Math)	-0.096 (0.093)	-0.307 (0.231)
$\beta_3+\beta_3^4$ (Time since enrolment)	-3.858 (0.019) ***	-3.537 (0.052) ***
β_4^4 (Student empl. previous period, h_{t-1}=¼)	3.406 (0.113) ***	2.745 (0.206) ***
β_5^4 (Student empl. previous period, h_{t-1}=½)	4.216 (0.119) ***	3.686 (0.222) ***
β_6^4 (Student empl. previous period, h_{t-1}=¾)	4.873 (0.121) ***	4.274 (0.243) ***
β_7 (Study grant)	-6.687 (0.194) ***	-8.168 (0.369) ***
Log Likelihood	-83026	-24876

Notes to Table 3: The table displays the estimates of the basic model parameters. (Standard errors in parentheses). ***, **, * indicates parameter significance at the 1%, 5%, 10% level of significance, respectively.

Table 4: Determinants of Academic Achievement.

	Marginal Effect on Probability							
	All University Entrants				Master Graduates			
	Drop out		Master		Excess time-to-graduation		Excess time-to-graduation (> 1 year)	
Individual characteristics:								
Type 1 probability		**-2.44**		**0.83**		**2.21**		**10.66**
Parental education:								
Mother high school	-0.03	0.01	0.04	0.03	0.03	-0.02	0.08	-0.08
Mother vocational training	-0.02	0.00	0.04	0.03	0.03	0.02	0.01	-0.02
Mother short higher education	-0.05	0.00	0.05	0.03	-0.01	-0.03	0.09	0.02
Mother medium higher education	-0.02	0.01	0.04	0.04	0.00	-0.01	0.07	0.03
Mother long higher education	-0.03	0.02	0.02	0.01	-0.01	-0.02	0.03	0.00
Father high school	-0.06	-0.04	-0.02	-0.02	0.03	0.01	0.15	0.13
Father vocational training	-0.03	-0.04	0.03	0.03	-0.02	-0.01	0.00	0.04
Father short higher education	-0.07	-0.06	0.09	0.09	0.02	0.03	0.00	0.03
Father medium higher education	-0.03	-0.02	0.02	0.02	0.00	0.01	-0.01	0.04
Father long higher education	**-0.06**	-0.04	**0.08**	0.08	-0.04	-0.04	-0.02	-0.02
Parental finances (real 2000 DKK):								
Mother's gross income/1000000	-0.07	-0.03	0.09	0.09	0.00	-0.03	0.17	0.13
Mother's wealth/1000000	-0.01	-0.01	0.03	0.03	-0.01	0.00	0.00	0.00
Father's gross income/1000000	**-0.09**	**-0.09**	0.07	0.07	**0.10**	**0.09**	0.01	0.01
Father's wealth/1000000	-0.04	-0.04	0.00	0.00	0.00	0.00	0.00	0.01
Pseudo R^2	0.02	0.14	0.01	0.02	0.01	0.05	0.01	0.29
Log likelihood	-1047	-920	-1365	-1356	-1314	-1255	-717	-510
Wald test of joint significance of parental background variables (p-value)	0.00	0.04	0.08	0.13	0.30	0.42	0.49	0.73

Notes to Table 4: The table displays marginal effects of parental background characteristics and the type 1 probability from logit estimations on the probabilities of dropping out of the university, acquiring a Master's degree, spending excess time to Master graduation, and spending more than one excess year to Master graduation. The comparison groups are mothers and fathers with no more than elementary school, respectively. Bold indicates statistically significant marginal effects on a 5% level of significance. The last row displays p-values from Wald tests of the joint significance of parental background characteristisc.

Table 5: Observed and Predicted Academic and Labor Market Outcomes.

	Observed	Predicted	
		One type	Two types
Accumulated course credits			
$P(g_t=0)$	0.42	0.43	0.42
$P(g_t=1)$	0.01	0.02	0.02
$P(g_t=2)$	0.05	0.05	0.05
$P(g_t=3)$	0.09	0.10	0.10
$P(g_t=4)$	0.03	0.03	0.03
$P(g_t=5)$	0.10	0.10	0.10
$P(g_t=6)$	0.20	0.19	0.19
$P(g_t=7)$	0.09	0.08	0.09
g_t	2.83	2.76	2.81
Across alternatives:			
g_t given $d_t^1=1$	4.31	3.80	3.86
g_t given $d_t^2=1$	4.68	4.30	4.34
g_t given $d_t^3=1$	4.18	3.78	3.83
g_t given $d_t^4=1$	2.25	1.43	1.49
Each time period:			
g_0	4.46	4.54	4.60
g_1	4.24	4.22	4.27
g_2	4.14	3.80	3.86
g_3	3.48	3.46	3.52
g_4	3.26	3.08	3.14
g_5	2.61	2.43	2.48
g_6	1.77	1.70	1.75
g_7	1.12	1.14	1.18
g_8	0.68	0.75	0.78
g_9	0.42	0.46	0.48
Total in last time period:			
G_{10}	25.53	25.03	25.12
Highest acquired degree:			
$E_{10} = 0$	0.23	0.18	0.19
$E_{10} = 1$	0.22	0.23	0.23
$E_{10} = 2$	0.55	0.58	0.57
Wages after University Exit			
W_t	158.79	143.80	143.27
Across highest acquired degree:			
W_t given $E_t=0$	136.64	119.84	119.33
W_t given $E_t=1$	141.55	131.75	131.31
W_t given $E_t=2$	183.43	167.82	167.25

Notes to Table 5: The table displays observed and predicted measures of academic achievement in terms of accumulated course credits and acquired university degrees and labor market achievement in terms of wages.

Table 6: Observed and Predicted Choices and Transitions.

	Observed	Predicted	
		One type	Two types
Distribution over alternatives			
$P(d_t^0=1)$	0.42	0.42	0.42
$P(d_t^1=1)$	0.23	0.20	0.21
$P(d_t^2=1)$	0.21	0.21	0.21
$P(d_t^3=1)$	0.09	0.11	0.10
$P(d_t^4=1)$	0.06	0.07	0.05
State transitions over time			
Full-time work:			
d_0^0	0.06	0.00	0.03
d_1^0	0.13	0.01	0.07
d_2^0	0.19	0.10	0.13
d_3^0	0.24	0.30	0.24
d_4^0	0.28	0.56	0.40
d_5^0	0.44	0.67	0.57
d_6^0	0.66	0.65	0.69
d_7^0	0.81	0.63	0.76
d_8^0	0.88	0.75	0.85
d_9^0	0.92	0.89	0.93
Education, $h_t=0$:			
d_0^1	0.43	0.38	0.32
d_1^1	0.34	0.30	0.28
d_2^1	0.30	0.26	0.28
d_3^1	0.29	0.19	0.27
d_4^1	0.26	0.11	0.23
d_5^1	0.21	0.08	0.17
d_6^1	0.12	0.13	0.14
d_7^1	0.07	0.20	0.13
d_8^1	0.05	0.15	0.09
d_9^1	0.03	0.07	0.05
Education, $0< h =¼$:			
d_0^2	0.35	0.36	0.39
d_1^2	0.36	0.41	0.40
d_2^2	0.31	0.38	0.37
d_3^2	0.28	0.28	0.29
d_4^2	0.26	0.17	0.21
d_5^2	0.17	0.11	0.13
d_6^2	0.09	0.09	0.07
d_7^2	0.04	0.07	0.04
d_8^2	0.02	0.04	0.02
d_9^2	0.01	0.02	0.01
Education, $¼< h =½$:			
d_0^3	0.11	0.21	0.21
d_1^3	0.11	0.21	0.20
d_2^3	0.13	0.19	0.17
d_3^3	0.11	0.15	0.14
d_4^3	0.13	0.09	0.10
d_5^3	0.11	0.06	0.06
d_6^3	0.06	0.05	0.03
d_7^3	0.03	0.04	0.02
d_8^3	0.02	0.02	0.01
d_9^3	0.02	0.01	0.00
Education, $h >½$:			
d_0^4	0.05	0.06	0.05
d_1^4	0.06	0.07	0.05
d_2^4	0.07	0.08	0.05
d_3^4	0.08	0.08	0.06
d_4^4	0.07	0.06	0.05
d_5^4	0.08	0.08	0.07
d_6^4	0.07	0.08	0.07
d_7^4	0.06	0.06	0.05
d_8^4	0.04	0.04	0.03
d_9^4	0.03	0.01	0.01

Notes to Table 6: The table displays observed and predicted choices and transitions over time.

Table 7: Decsriptive Statistics, by Field of University Education.

Individual Characteristics	Humanities / Art / Education	Life Sciences	Social Science	Business	Natural Science / Engineering	Health Sciences
	Mean (Standard Deviation)					
	Field of University Education					
At University Entry:						
High school GPA	8.94 (0.77)	8.88 (0.81)	9.23 (0.74)	8.49 (0.83)	8.95 (0.81)	9.26 (0.60)
High school Math level	1.55 (0.79)	2.67 (0.61)	2.13 (0.87)	1.80 (0.94)	2.73 (0.54)	2.70 (0.54)
Female	0.69	0.43	0.49	0.36	0.24	0.60
Accumulated work experience	0.60 (0.60)	0.53 (0.56)	0.66 (0.58)	0.70 (0.66)	0.47 (0.55)	0.65 (0.57)
Age	21.32 (0.76)	21.23 (0.73)	21.18 (0.78)	21.10 (0.74)	20.93 (0.77)	21.13 (0.77)
During University Enrolment:						
Accumulated course credits per year	3.99 (2.44)	3.97 (2.42)	4.30 (2.24)	4.08 (2.41)	4.26 (2.44)	4.88 (1.94)
Accumulated work experience per year (years)	0.16 (0.22)	0.13 (0.22)	0.18 (0.22)	0.27 (0.32)	0.16 (0.25)	0.14 (0.16)
Type of student employment (job description):						
High skilled jobs	0.12	0.10	0.12	0.09	0.17	0.17
Medium skilled jobs	0.03	0.02	0.03	0.10	0.04	0.03
Office work	0.11	0.12	0.15	0.18	0.08	0.07
Care, sales and service jobs	0.18	0.14	0.16	0.11	0.09	0.13
Cleaner and other low skilled jobs	0.51	0.54	0.49	0.45	0.53	0.57
Type of student employment (sector of industry):						
Shops, hotels, and restaurants	0.18	0.11	0.17	0.22	0.12	0.11
Finance-, telecom-, and transport companies	0.06	0.03	0.07	0.13	0.05	0.03
Business and consultancy services	0.12	0.16	0.24	0.25	0.16	0.38
Unions, associations, societies, and outfits	0.11	0.05	0.08	0.05	0.04	0.03
Public and personal services (incl. jobs at universities and other educational institutions)	0.20	0.20	0.15	0.09	0.18	0.19
At University Exit:						
Highest acquired degree:						
Dropout	0.25	0.34	0.16	0.28	0.21	0.09
Bachelor	0.35	0.18	0.23	0.25	0.08	0.03
Master	0.40	0.48	0.61	0.46	0.70	0.88
Excess time-to-graduation (Bachelor)	0.42	0.36	0.51	0.32	0.43	0.50
Excess time-to-graduation (Master)	0.94	0.80	0.95	0.88	0.75	0.84
Excess time-to-graduation > 1 year (Master)	0.81	0.69	0.68	0.49	0.51	0.65
After University Exit:						
Hourly wages (real 2000 DKK)	154.90 (59.31)	162.81 (46.82)	176.17 (70.93)	178.29 (79.83)	183.71 (57.01)	206.23 (87.91)
Yearly earnings (real 2000 DKK)	159,786 (111,764)	189,898 (123,479)	219,134 (138,667)	237,730 (130,085)	221,700 (131,816)	266,481 (128,499)
Number of Individuals	589	222	439	370	361	148
Fraction of total sample	0.28	0.10	0.21	0.17	0.17	0.07
Number of Observations	5343	2017	3964	3377	3282	1366

Notes to Table 7: The table shows average characteristics of university graduates and dropouts (standard deviation in parentheses). For indicator variables the fraction of the sample is reported. The descriptive statistics are displayed separately by field of university education. The exchange rate on December 31, 2000 was 8.0205 DKK/USD and 7.4631 DKK/Euro.

Table 8: Parameter Estimates, by field of University Education.

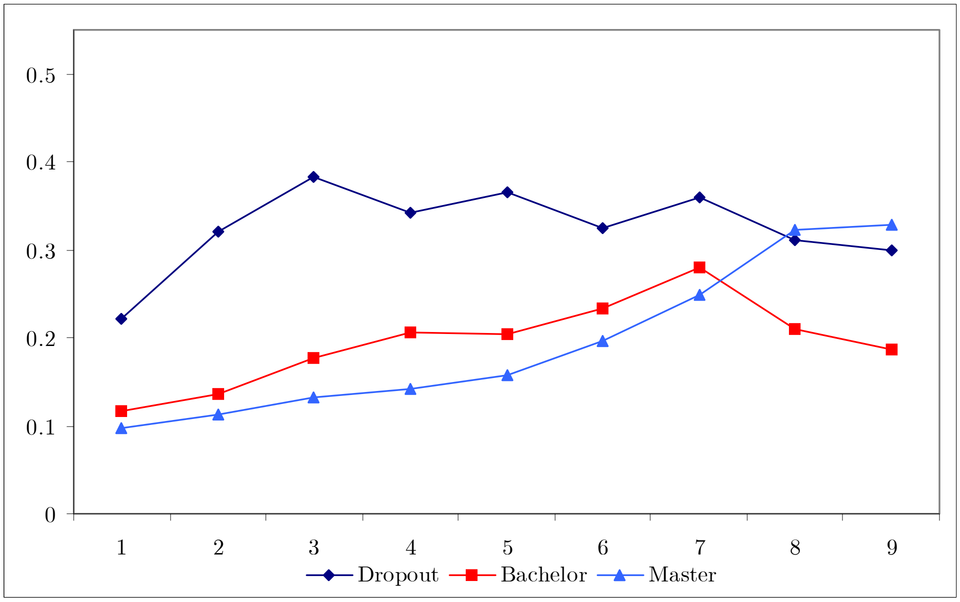
	Humanities / Art / Education	Life Sciences	Social Science
		w_t	
α_0	4.53 (0.01) ***	4.48 (0.02) ***	4.53 (0.01) ***
α_1 (Bachelor degree)	0.09 (0.01) ***	0.11 (0.03) ***	0.16 (0.02) ***
α_2 (Master degree)	0.31 (0.02) ***	0.39 (0.03) ***	0.36 (0.02) ***
α_3 (Experience)	0.09 (0.01) ***	0.09 (0.02) ***	0.08 (0.01) ***
α_4 (Experience^2)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
		g^*	
γ_1 (High school GPA=9)	0.23 (0.05) ***	0.28 (0.09) ***	0.16 (0.06) **
γ_2 (High level Math)	0.08 (0.06)	0.10 (0.10)	0.06 (0.06)
γ_3 (Bachelor degree, $\mathbf{1}[E_t=1]$)	0.07 (0.07)	0.31 (0.12) **	0.59 (0.08) ***
γ_4 (Accumulated course credits, G_t)	0.09 (0.00) ***	0.06 (0.01) ***	0.06 (0.01) ***
γ_5 (Time since enrolment)	-0.61 (0.02) ***	-0.48 (0.04) ***	-0.62 (0.03) ***
γ_6 (Student employment, h_t=¼)	0.26 (0.07) ***	0.32 (0.11) ***	0.38 (0.08) ***
γ_7 (Student employment, h_t=½)	-0.46 (0.09) ***	-0.40 (0.17) **	0.03 (0.09)
γ_8 (Student employment, h_t=¾)	-1.14 (0.10) ***	-1.15 (0.19) ***	-1.37 (0.11) ***
		$P(d_t^k=1)$	
c (Not employed previous period, h_{t-1}=0)	-0.07 (0.13)	0.33 (0.23)	2.18 (0.23) ***
β_0	-23.08 (0.12) ***	-15.71 (0.33) ***	-2.22 (0.70) ***
β_1 (High school GPA=9)	0.74 (0.11) ***	0.94 (0.21)	8.97 (0.44) ***
β_2 (High level Math)	0.93 (0.13) ***	1.19 (0.23)	4.76 (0.43) ***
β_3 (Time since enrolment)	-2.07 (0.02) ***	-3.16 (0.05) ***	-24.64 (0.11) ***
$\beta_1+\beta_1^2$ (High school GPA=9)	0.21 (0.11) **	0.29 (0.22)	7.80 (0.44) ***
$\beta_2+\beta_2^2$ (High level Math)	0.48 (0.13) ***	0.22 (0.23)	3.74 (0.44) ***
$\beta_3+\beta_3^2$ (Time since enrolment)	-2.19 (0.03) ***	-3.17 (0.05) ***	-24.61 (0.11) ***
β_4^2 (Student empl. previous period, h_{t-1}=¼)	2.30 (0.10) ***	2.52 (0.17) ***	5.31 (9.00)
β_5^2 (Student empl. previous period, h_{t-1}=½)	1.72 (0.16) ***	2.58 (0.30) ***	4.17 (0.33) ***
β_6^2 (Student empl. previous period, h_{t-1}=¾)	1.57 (0.23) ***	1.20 (0.39) ***	1.92 (0.26) ***
$\beta_1+\beta_1^3$ (High school GPA=9)	-0.32 (0.12) **	-0.56 (0.26) *	6.81 (0.45) ***
$\beta_2+\beta_2^3$ (High level Math)	0.15 (0.16)	-0.39 (0.26)	2.93 (0.44) ***
$\beta_3+\beta_3^3$ (Time since enrolment)	-2.23 (0.03) ***	-3.22 (0.06) ***	-24.62 (0.11) ***
β_4^3 (Student empl. previous period, h_{t-1}=¼)	1.60 (0.12) ***	2.22 (0.23) ***	618.72 (9.00)
β_5^3 (Student empl. previous period, h_{t-1}=½)	2.14 (0.17) ***	3.02 (0.35) ***	5.58 (0.34) ***
β_6^3 (Student empl. previous period, h_{t-1}=¾)	1.85 (0.24) ***	1.67 (0.48) ***	2.62 (0.29) ***
$\beta_1+\beta_1^4$ (High school GPA=9)	-0.52 (0.16) ***	0.19 (0.30)	7.46 (0.45) ***
$\beta_2+\beta_2^4$ (High level Math)	0.30 (0.19)	0.00 (0.31)	3.44 (0.45) ***
$\beta_3+\beta_3^4$ (Time since enrolment)	-2.25 (0.03) ***	-3.31 (0.06) ***	-24.71 (0.11) ***
β_4^4 (Student empl. previous period, h_{t-1}=¼)	2.18 (0.20) ***	2.45 (0.34) ***	22.92 (9.00)
β_5^4 (Student empl. previous period, h_{t-1}=½)	3.08 (0.20) ***	3.50 (0.41) ***	5.81 (0.37) ***
β_6^4 (Student empl. previous period, h_{t-1}=¾)	4.04 (0.23) ***	3.50 (0.40) ***	4.54 (0.28) ***
β_7 (Study grant)	-4.86 (0.33) ***	-4.47 (0.56) ***	-4.91 (0.43) ***
Log Likelihood	-20511	-7303	-8400

Table 8 *continued*

	Business	Natural Science / Engineering	Health Sciences
		w_t	
α_0	4.49 (0.01) ***	4.60 (0.01) ***	4.35 (0.03) ***
α_1 (Bachelor degree)	0.24 (0.02) ***	0.17 (0.03) ***	0.06 (0.07)
α_2 (Master degree)	0.43 (0.02) ***	0.37 (0.02) ***	0.55 (0.04) ***
α_3 (Experience)	0.09 (0.01) ***	0.09 (0.01) ***	0.21 (0.03) ***
α_4 (Experience^2)	0.00 (0.00)	0.00 (0.00)	-0.02 (0.00) ***
		g^*	
γ_1 (High school GPA=9)	0.16 (0.08) *	0.22 (0.08) ***	0.07 (0.10)
γ_2 (High level Math)	0.45 (0.08) ***	0.30 (0.09) ***	0.48 (0.11) ***
γ_3 (Bachelor degree, $\mathbf{1}[E_t=1]$)	0.21 (0.09) **	0.59 (0.15) ***	0.40 (0.24) **
γ_4 (Accumulated course credits, G_t)	0.08 (0.01) ***	0.09 (0.01) ***	0.13 (0.01) ***
γ_5 (Time since enrolment)	-0.66 (0.03) ***	-0.67 (0.03) ***	-0.64 (0.05) ***
γ_6 (Student employment, h_t=¼)	0.33 (0.09) ***	0.25 (0.09) ***	0.17 (0.12)
γ_7 (Student employment, h_t=½)	-0.27 (0.11) **	-0.30 (0.12) **	0.16 (0.15)
γ_8 (Student employment, h_t=¾)	-1.69 (0.12) ***	-1.83 (0.14) ***	-3.49 (0.21) ***
		$P(d_t^k=1)$	
c (Not employed previous period, h_{t-1}=0)	1.32 (0.21) ***	1.01 (0.19) ***	0.47 (0.30)
β_0	-22.17 (0.28) ***	-16.39 (0.26) ***	-7.51 (0.30) ***
β_1 (High school GPA=9)	-1.54 (0.22) ***	0.08 (0.15)	0.76 (0.21) ***
β_2 (High level Math)	-0.41 (0.22) *	0.53 (0.19) ***	0.75 (0.22) ***
β_3 (Time since enrolment)	-6.32 (0.04) ***	-3.12 (0.04) ***	-1.78 (0.05) ***
$\beta_1+\beta_1^2$ (High school GPA=9)	-2.74 (0.21) ***	-0.26 (0.16)	0.76 (0.21) ***
$\beta_2+\beta_2^2$ (High level Math)	-1.60 (0.22) ***	-0.38 (0.19) **	0.29 (0.21)
$\beta_3+\beta_3^2$ (Time since enrolment)	-6.31 (0.04) ***	-3.22 (0.04) ***	-1.93 (0.06) ***
β_4^2 (Student empl. previous period, h_{t-1}=¼)	4.05 (0.19) ***	2.26 (0.14) ***	2.92 (0.18) ***
β_5^2 (Student empl. previous period, h_{t-1}=½)	4.41 (0.29) ***	2.02 (0.23) ***	2.78 (0.30) ***
β_6^2 (Student empl. previous period, h_{t-1}=¾)	0.83 (0.33) **	2.91 (0.36) ***	2.46 (0.56) ***
$\beta_1+\beta_1^3$ (High school GPA=9)	-3.34 (0.22) ***	-0.95 (0.19) ***	0.12 (0.23)
$\beta_2+\beta_2^3$ (High level Math)	-2.24 (0.23) ***	-1.38 (0.21) ***	-0.35 (0.24)
$\beta_3+\beta_3^3$ (Time since enrolment)	-6.27 (0.05) ***	-3.18 (0.05) ***	-1.88 (0.06) ***
β_4^3 (Student empl. previous period, h_{t-1}=¼)	3.74 (0.20) ***	2.12 (0.18) ***	2.01 (0.22) ***
β_5^3 (Student empl. previous period, h_{t-1}=½)	5.16 (0.29) ***	3.35 (0.24) ***	3.24 (0.32) ***
β_6^3 (Student empl. previous period, h_{t-1}=¾)	1.84 (0.29) ***	3.37 (0.41) ***	2.86 (0.60) ***
$\beta_1+\beta_1^4$ (High school GPA=9)	-3.17 (0.23) ***	-1.14 (0.22) ***	-0.15 (0.38)
$\beta_2+\beta_2^4$ (High level Math)	-2.09 (0.24) ***	-1.59 (0.24) ***	-0.26 (0.43)
$\beta_3+\beta_3^4$ (Time since enrolment)	-6.26 (0.05) ***	-3.06 (0.05) ***	-1.99 (0.08) ***
β_4^4 (Student empl. previous period, h_{t-1}=¼)	3.95 (0.25) ***	2.67 (0.29) ***	3.18 (0.53) ***
β_5^4 (Student empl. previous period, h_{t-1}=½)	5.80 (0.31) ***	3.89 (0.30) ***	3.76 (0.53) ***
β_6^4 (Student empl. previous period, h_{t-1}=¾)	4.29 (0.22) ***	6.46 (0.32) ***	5.42 (0.70) ***
β_7 (Study grant)	-3.83 (0.36) ***	-5.25 (0.49) ***	-8.37 (1.17) ***
Log Likelihood	-17476	-11426	-5713

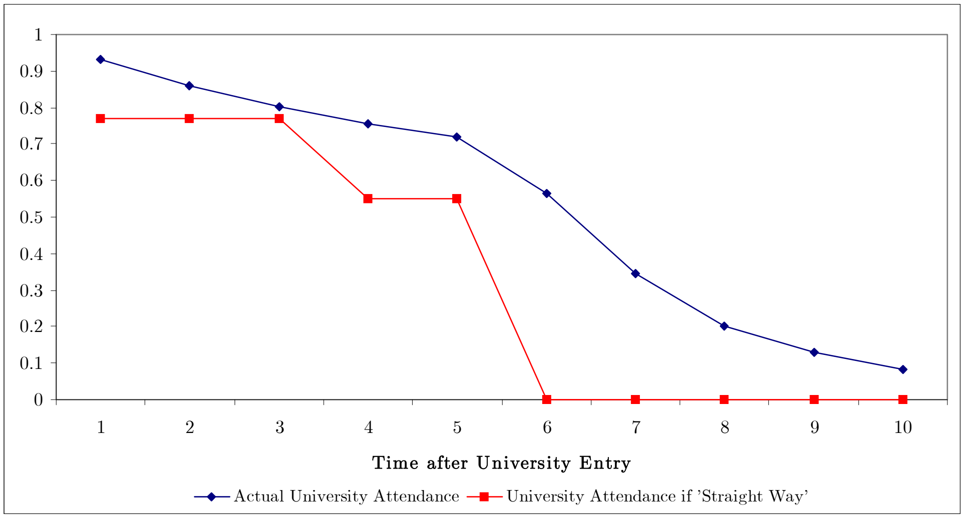
Notes to Table 8: The table displays the estimates of the basic model parameters estimated separately by field of university education. (Standard errors in parentheses). ***, **, * indicates parameter significance at the 1%, 5%, 10% level of significance, respectively.

Figure 1: Annual Student Employment Experience.



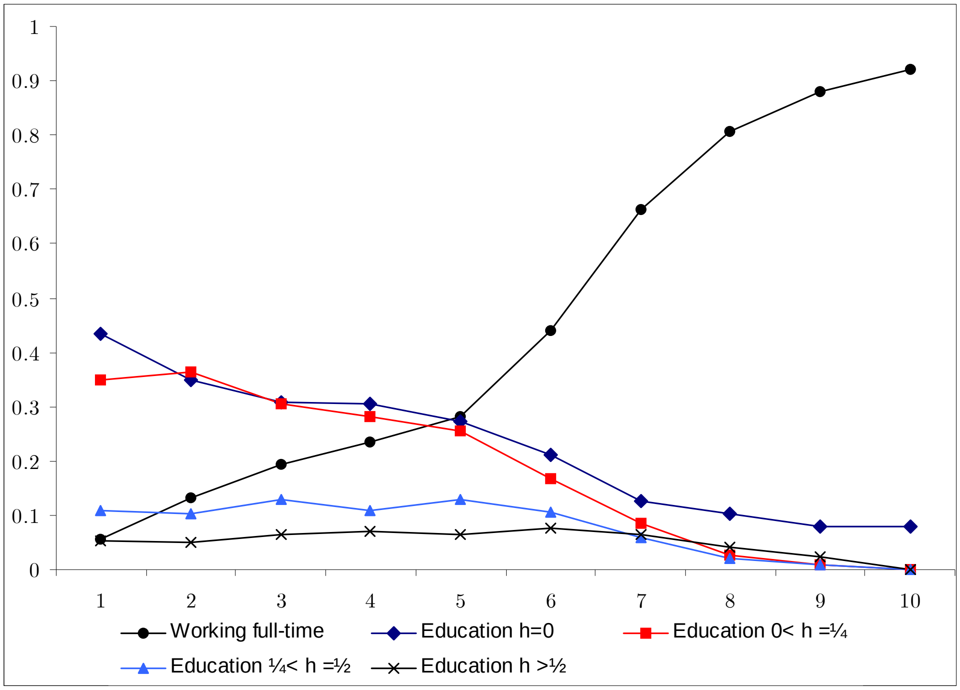
Notes to Figure 1: The figure displays accumulated labor market experience (in years) in the year over time after university entry for full-time university students. The figure displays separate student employment experience profiles for university dropouts and individuals graduating from the university with Bachelor's and Master's degrees, respectively.

Figure 2: Actual and 'Straight Way' University Attendance Rates.



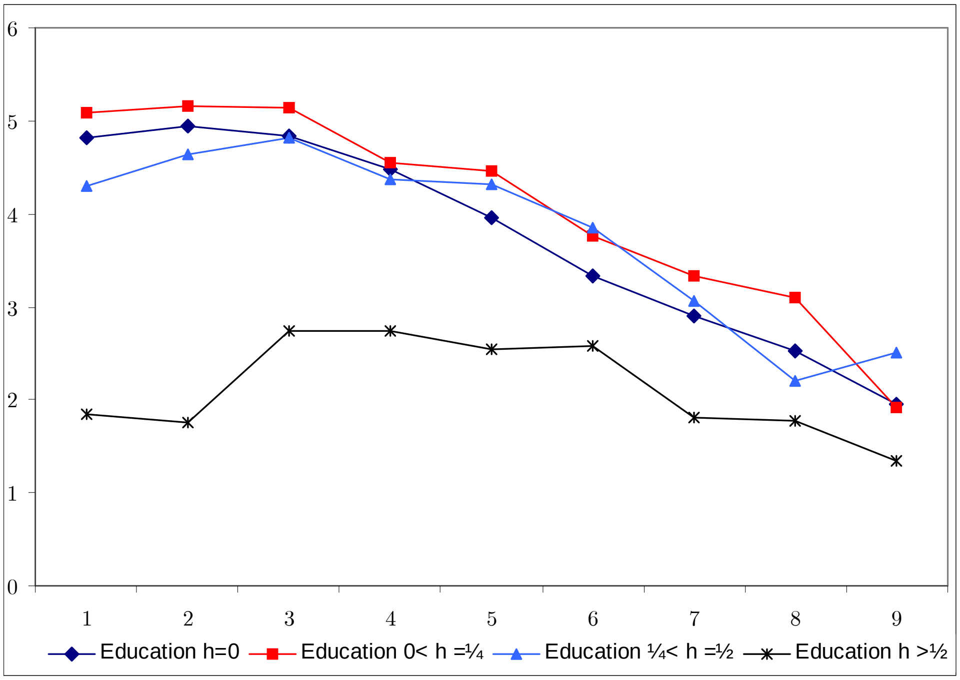
Notes to Figure 2: The figure shows the actual and 'straight way' hypothetical university attendance over time after university entrance. The blue line with diamonds displays the fraction of university entrants still enrolled in full-time education over time (years) after initial university enrolment. The red line with squares displays the fraction of university entrants who would have been enrolled in full-time education over time after initial university enrolment if all of them took the 'straight way' through the educational system, i.e. no dropout and no excess time-to-graduation at any level.

Figure 3: Transition from University Education to Labor Market Work.



Notes to Figure 3: The figure shows the transition from full-time university education to work. It displays the fraction of individuals in each state $k \in \{0, 1, 2, 3\}$ at each point in time after university entry.

Figure 4: Accumulated Course Credits per Year, by Student Employment State.



Notes to Figure 4: The figure displays the average accumulated course credits per year, g_t, for university attendants over time since initial enrolment. The amount of course credits is diplayed separately by amount of labor market work in the year, $k \in \{1, 2, 3, 4\}$. A total of six course credits have to be accumulated in order to successfully pass one year of university study.