Timing and Incentives: Impacts of Student Aid on Academic Achievement

Juanna Schrøter Joensen*

August 7, 2011

Abstract

This paper models the university-to-work transition in a stochastic dynamic environment, where students may study and work simultaneously. The structural model is estimated using a unique panel data set with exogenous variation from changing threshold levels for maximum student grants. Estimates reveal that uniformly increasing student aid increases enrollment time. Policy simulations show that because of the non-linear effect of student working hours on academic achievement, however, tilting student aid towards those who work fewer hours increases graduation rates by 5 percentage points, but is ineffective in shortening time-to-graduation. A combination of tilting student aid and improving student abilities earlier in the education production process both increases graduation rates and lowers time-to-graduation. Including incentives into the student aid package with merit aid or timely graduation bonuses also tend to be effective policy devises to amend these academic outcomes.

JEL Classification: I21, I28, J24

Keywords: Student Aid, Student Employment, Investment in Education, Grade Level Progression.

---

*Department of Economics, Stockholm School of Economics. Juanna.Joensen@hhs.se
Financial support from the Jan Wallanders and Tom Hedelius stiftelse is gratefully acknowledged. Thanks to seminar participants at SOFI, Royal Holloway, IFAU, Tinbergen Institute, DIW Berlin, Tor Vergata, Statistics Norway, and at the workshops on 'Skill Formation' in Freiburg, 'Labor Markets' in Stavanger, ESPE (Sevilla 2009) , EALE (Tallinn 2009), SED (Montréal 2010), ESWC (Shanghai 2010), and SOLE (Vancouver 2011) conferences for valuable discussions and comments. The usual disclaimers apply.

1 Introduction

Increasing graduation rates and the speed at which individuals obtain higher education are declared social objectives in many countries, as a highly educated labor force is seen as key to sustain economic development and growth. Despite the considerable amount of public debate on these issues, not much is known about the impacts of potential policy interventions. This paper provides an economic model capable of evaluating how academic outcomes can be amended by public policies; particularly, changing student aid packages to include incentives through merit aid and timely graduation bonuses.

This paper models the university-to-work transition in a stochastic dynamic environment, where students may study and work simultaneously. Each year students decide whether to stay enrolled at the university or to enter the labor market and work full-time given wages that depend on academic degrees and labor market experience. Enrolled students also decide how many hours to work part-time. These decisions are all conditional on prior academic achievement, labor market opportunities, and expectations about the future. The model thus explicitly accounts for the sequential nature of education and employment decisions, and simultaneously models labor market opportunities, the accumulation of academic capital in terms of grade level progression, and the accumulation of labor market experience through employment decisions while attending university. Graduation probabilities, times-to-graduation, and accumulation of work experience while enrolled are thus endogenously determined. The main advantage of this approach is the possibility of constructing counterfactuals enabling evaluation of the impacts of changing the direct costs of university enrollment on academic outcomes.

The most important estimation issue is self-selection. Selection bias arises if students choose the amount of employment based on unobservable characteristics correlated with academic and labor market ability. To account for unobserved heterogeneity, I assume students are drawn from a finite mixture distribution: each student belongs 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. Student type is unobserved to the econometrician, but 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 (2008) and to relate unobserved student type to observed parental background. Potential self-selection biases are eliminated if students make decisions according to the model conditional on observed characteristics and type.

The dynamics of the model are important for three reasons. First, they separate the effect of the educational environment from the effect of the labor market. Second, the dynamics allow individuals to learn about their academic 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.

The model parameters are estimated using an extensive register-based panel data, which covers a random 10% sample of the Danish population. These data include detailed educational event histories and labor market histories; including actual labor market experience, approximate working hours, unemployment degree, labor income, and wages. Furthermore, data on accumulated university course credits has been collected and merged for the particular purpose of this study. The data also comprises parental background variables, courses taken in high school, and high school grade point average (GPA). University admission is almost exclusively conditional on GPA and Math level from high school. In Denmark, there are no tuition fees for university education, the enrolment period is not limited, and all admitted students are eligible for a study grant that suffices for necessary costs of living. Three important differences between the US and Denmark minimize self-selection issues and make it easier to accurately model students' potential funding opportunities: First, US grants are mainly provided by the states and colleges, while Danish grants are predominantly provided by the government. Second, US grants depend on parental income - typically grants are a decreasing function of parental income, because of need-based grants (e.g. the Pell grant), whereas Danish grants are largely independent of parental income. Third, Danish grants depend on students' earnings, making it pivotal to jointly model students' employment choices and their impacts on academic achievement. Furthermore, an exogenous increase in the study grant threshold of maximum allowable earnings is exploited to identify the effect of grants on student employment choices.

The estimated 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 job search costs. Conditional on observed student 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 or motivation and higher consumption value of university attendance tend to have lower dropout rates, higher Master's graduation rates, but also a higher probability of spending excess time to graduation. These academic outcomes are not easily amendable by tilting financial aid towards students who work less; however, there seem to be some promise from changing incentive and timing aspects of financial aid packages.

Quantifying the effects of student aid packages is pivotal to inform policy makers how to spend scarce public subsidies more effectively. The effects depend on how strongly student behavior responds to changes in the direct/opportunity cost of university attendance and various implicit financial aid incentives. Besides quantifying the effects of pecuniary incentives on student behavior, we also want to understand the mechanisms by which they operate. Higher non-graduate wages increase the opportunity cost of university attendance. Higher study grants (lower tuition) lower the direct cost of university. For students working part-time, a study grant scheme like the Danish one works like an earned income tax credit (or implicit negative income tax) by making it more costly to increase work hours. This paper quantifies the impacts of study grants in an environment where students can work to self-finance university attendance. Working lowers current opportunity costs, but increases future opportunity costs of education through increased labor market experience. Working can also increase the direct costs by lowering study grants, as well as decrease the consumption value and future opportunity costs of education to the extent that there are adverse effects on academic achievement. These effects must be quantified in order to fully capture the impact of policies designed to amend educational outcomes. The main mechanism underlying policy makers' presumption is that adverse student outcomes (such as dropping out of university) are due to credit contraints rather than being an outcome of an informed choice in an unconstrained environment. The desirability of a study grant is much greater if credit constraints are binding, because it would improve efficiency.[1] Ignoring students ability to self-finance their education would overestimate the opportunity cost of education and thus also the impacts of credit constraints. This paper explores these mechanisms by comparing the impacts of study grants and student employment for

[1] There is a large literature on the impacts of credit constraints on education outcomes; e.g. Cameron and Heckman (1998), Eckstein and Wolpin (1999), Keane and Wolpin (2001), Carneiro and Heckman (2002), Cameron and Taber (2004), Nielsen, Sørensen and Taber (2010), Stinebrickner and Stinebrickner (2010). The main finding is that credit constraints seem to be present, but relaxing them does not improve education outcomes considerably.

students whose parents have high and low education, income and wealth, respectively. Parents with very high wealth are unlikely to be credit constrained.[2] There do not seem to be significant behavioral differences between constrained and unconstrained students, hence there is not much evidence that the estimated impacts are due to credit constraints instead of unconstrained cost effects. Since parental information is not used in the estimation, my model's ability to capture these differences across parental background also provide a valuable out-of-sample fit corroborating the credibility of policy simulations based on the model.

The rest of the paper is organized as follows: Section 2 discuses this paper's contribution to the literature. Section 3 lays out the dynamic model of education and student employment choices, and the econometric techniques used to estimate the model. The data and several empirical regularities are presented in Section 4. Section 5 discusses the empirical results, model fit, and policy simulations. Section 6 concludes.

2 Background

Despite the tremendous public subsidies to students attending higher education, little is known about the impacts of student aid on academic achievement. Dynarski (2008) concludes that financial aid policies can play a welfare enhancing role in increasing college graduation rates. However, dropout rates are high even with free tuition, suggesting that the direct costs of college are not the only impediments to college completion. This suggests that more than tuition reductions is needed in order to substantially increase college graduation rates; for example, funding that extends beyond direct costs to opportunity costs. In this sense, the Danish case is interesting because it is one of the most generous in the world - with free tuition and large public study grants to which all enrolled students are eligible.[3] The model in this paper allows students to further relax potential credit constraints by working part-time while enrolled in education. The effects of financial incentives are thus estimated in an environment with

[2]Whether they allocate more resources to their childrens' education is another issue not addressed here. Keane and Wolpin (2001) find that parents with higher education provide higher conditional education transfers.

[3]Subsidies directly to students make up more than 30% of total public expenditures on higher education, which is the equivalent of 0.85% of GDP, cf. OECD Education at a Glance (2005). The average public cost per student per year in higher education in Denmark was $125,000$ DKK in 2004 and has been fairly stable over the past decades, cf. Statistics Denmark, www.statistikbanken.dk

very low opportunity costs of college attendance, hence very low pecuniary barriers to college graduation.

The literature provides ambiguous evidence on the impacts of financial aid on academic achievement.[4] However, most of this literature does not control adequately for confounding unobservable factors. Quasi-experimental studies find that financial aid has a negative impact on college drop-out and retention, while it has a positive impact on completion; see e.g. Bound and Turner (2002), Dynarski (2003, 2008), Bettinger (2004), Scott-Clayton (2011) and Arendt (2008).[5] Garibaldi, Giavazzia, Ichino and Ettore (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, relative to an average late graduation probability of 80%. This increase in tuition is equivalent to a reduction in study grants in my model. On the other hand, exploiting exogenous variation from a reform of the Danish study grant scheme, Arendt (2008) does not find significant impacts on times-to-graduation - only on drop-out rates.[6] Financial aid can not only relax credit constraints, but also induce student effort. Angrist et al. (2009), DesJardins and McCall (2009), and Scott-Clayton (2011) demonstrate the potential effectiveness of providing incentives related to merit and timing in financial aid packages. Scott-Clayton (2011) examines the PROMISE scholarship in West Virginia, which provides free tuition to college students maintaining a minimum GPA and course load. She exploits discontinuities in eligibility and timing of implementation to identify program effects, and finds significant impacts on academic outcomes, including an increase in timely completion rates. She concludes that the program works by establishing clear academic goals and incentives to meet them, rather than simply reducing the cost of college. The effects of the PROMISE and other similar widespread scolarships can be simulated in my model. This paper thus contributes to the literature by providing estimates of both the direct effects of study grants on academic achievement and the indirect effects operating through student employment choices. This enables a coherent and unified framework for interpreting existing evidence through evaluation of

[4] There is a related and rapidly growing literature on how financial incentives affect education; see e.g. Angrist, Bettinger, Bloom, King, and Kremer (2002), Angrist, Lang, and Oreopoulos (2009), and Leuven, Oosterbeek, and van der Klaauw (2010) for recent contributions.

[5] The quasi-experimental literature also finds positive effects of financial aid on college enrollment; see Dynarski (2002) for a review.

[6] Nielsen, Sørensen and Taber (2008) also use this reform to show that higher study grants marginally increase college enrollment.

the effects of various public policy interventions on important academic outcomes such as graduation rates and times-to-graduation. Hereby also providing a validation of the model by comparing ex ante predictions to ex post out-of-sample evaluations.

Another novelty of this paper is that it explicitly accounts for uncertainty in academic success. Despite the fact that dropout rates are high, traditional human capital 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.[7] However, most students spend excess time-to-graduation and dropout rates are high. 30% dropped out of higher education in the average OECD country in 2003.[8] 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. Bound, Lovenheim and Turner (2007) compare academic outcomes in NLS72 and the National Educational Longitudinal Study (NELS:88) eight years after high school graduation. They find that college graduation rates fall from 60% to 57% at public 4-year colleges, and rise from 68% to 78% at private 4-year colleges. Hence, the graduation rate of 77% at Danish universities is slightly above the OECD average of 70% and very close to the NELS:88 private 4-year college graduation rate. Bound et al. (2007) also document that dropout rates and times-to-graduation have increased at US colleges over the last decades, as has student employment, which is common among university students. In 2003, an average OECD country 15-29 year old student was employed the equivalent of 27% of full-time employment while enrolled in education. The average US student works the equivalent of 39% of full-time employment while enrolled in education.[9] Students at Danish universities work approximately as much as students at 4-year colleges in the US, where Bound et al. (2007) document that around 40% of students are employed, and 10% of students work more than 20 hours a week. There are several reasons why students choose to work part-time. Students may be credit constrained and depend on the extra income;[10] but employment may also be an investment in enhancing labor market skills. There might, however, also be negative effects of student employment.

---

[7] The human capital literature testifying to the importance of education to numerous economic outcomes has grown out of the pioneering work by Mincer (1974) and Becker (1964). See Belzil (2007) for an excellent survey of the evolution of this literature with particular focus on stochastic dynamic programming models.

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

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

[10] Leslie (1984) reports that US youth self-finance around 20% of college expenses.

Obviously, there is a time-use trade-off between working and studying. Working during the semester may interfere with learning and academic performance if it crowds out study time. Previous studies found that student employment 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, student employment is found to lower academic achievement by increasing the probability of dropping out and time-to-graduation; see e.g. Ehrenberg and Sherman (1987), and Stinebrickner and Stinebrickner (2003). Joensen (2009) corroborates this evidence and disentangles the mechanisms through which student employment affects academic and labor market success. Since the enrolment period at the university is not restricted, working may also lead to longer times-to-graduation. Long times-to-graduation can be costly by decreasing both the private and social returns to university education. It is possible that the excess study time reflects increased human capital acquisition, however empirical evidence deems this unlikely. Bound et al. (2007) and Garibaldi et al. (2007) provide evidence on excess times-to-graduation in the US. Brunello and Winter-Ebner (2003) study expected times-to-graduation for Economics and Business 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. In my sample of Danish university students, the average time-to-graduation with a Master's degree is 20 months 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. Not much is known about the optimal length of the period for learning the required skills to obtain any given degree. Garibaldi et al. (2007) argue that with market distortions, like public subsidies to education, private student incentives do not lead to socially optimal times-to-graduation. Long times-to-graduation provide private monetary costs to individuals by shortening their careers after graduation. Velfærdskommisionen (2005) points out that long times-to-graduation are considered a waste of high skilled labor. Brodaty, Gary-Bobo and Prieto (2006) provide evidence that French individuals with longer than average time-to-graduation have significantly lower wages and employment rates in their early career. All in all, the structural model in this paper is more consistent with the data, as well as capable of jointly testing these hypotheses.

3 Model Setup

This section presents the structural model specifying the educational environment in terms of grade level progression, choice sets, and graduation requirements, as well as the labor market environment in terms of wage opportunities. Although all decisions are taken by individuals, the subscripts i are suppressed for notational ease.

Education and work choices and outcomes are modeled from initial university entry until exit. Individuals 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 labor market experience H_0. In each period $t = 1, ..., T$ individuals have the option to stay enrolled, $D_t = 1$ and work $h_t \in \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}\right\}$ hours, or to drop out or graduate and start working full-time, $D_t = 0$ and $h_t = 1$, receiving wages that depend on education level and accumulated labor market experience. Each year students receive a wage offer and an idiosyncratic labor market productivity shock.

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 student abilities, skills, prior academic achievement, the degree of participation in the labor market, and time since university entry.[11] Enrolled students accumulate course credits following the law of motion: $G_{t+1} = G_t + g\left(A, K, G_t, E_t, t, h_t, \varepsilon_t^g\right)$, where ε_t^g is an idiosyncratic academic achievement shock, possibly from changes in motivation due to, say, an interesting lecturer or sudden health issues. All stochastic components are revealed at the beginning of the decision period, for example, students know ε_t^g when making their decisions at time $t+1$, but not before. The econometrician does not observe ε_t^g at any time.

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.

[11] 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, study time, 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.

3.1 Individuals' Optimization Problem

Enrolled individuals face five mutually exclusive and exhaustive alternatives: $(D_t, h_t) \in \left\{(0,1), (1,0), \left(1, \frac{1}{4}\right), \left(1, \frac{1}{2}\right), \left(1, \frac{3}{4}\right)\right\}$. The discrete choices can be represented by $d_t = (d_t^0, d_t^1, d_t^2, d_t^3, d_t^4)$, where $d_t^k = \mathbf{1}\left[\text{alternative } k \text{ chosen at } t\right]$ and $\mathbf{1}\left[\cdot\right]$ is an indicator function equal to unity when the argument is true. An individual makes the sequence of choices $\{d_t\}_{t=1}^T$ to maximize the expected present value of utility:

$$\max_{d_t^k} E\left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} \sum_{k=0}^{4} d_\tau^k U_\tau^k\left(S_\tau\right)\right]. \tag{1}$$

The state variables $S_t = (X_t, \varepsilon_t)$ include all the information known to the individual at time t, where $X_t = (A, K, G_t, E_t, H_t)$ are the state variables also observed by the econometrician, but $\varepsilon_t = (\varepsilon_t^w, \varepsilon_t^g, \varepsilon_t^0, \varepsilon_t^1, \varepsilon_t^2, \varepsilon_t^3, \varepsilon_t^4)$ are known only to the individual.[12] The current utility of an individual with state variable S_t from choosing alternative k is assumed additively separable in X_t and ε_t^k, i.e. $U_t^k\left(S_t\right) = U_t^k\left(X_t\right) + \varepsilon_t^k$. The value of university attendance consists of both the current consumption value of education and its investment effect on future wages. Maximization of (1) is achieved by choosing the optimal sequence of feasible control variables d_t. Since the shocks in period t are revealed when 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.

The optimization problem (1) can be rewritten as a dynamic programming problem via Bellman's principle of optimality. The value function, $V_t\left(S_t\right)$, is defined to be the maximal expected present value at time t, given state, S_t, and discount factor, δ:[13]

$$V_t\left(S_t\right) = U_t^k\left(S_t\right) + \delta E\left[V_{t+1}\left(S_{t+1}\right)\right] \tag{2}$$

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 $V_t\left(S_t\right) \equiv \max_k V_t^k\left(S_t\right)$, where $V_t^k\left(S_t\right)$ denotes the alternative specific

[12] 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 student choices and outcomes, but does so as a surjective function of G_t. Hence, I only have to keep track of the laws of motion for H_t and G_t, respectively.

[13] This approach dates back to Bellman (1957). See e.g. Adda and Cooper (2003), Stokey and Lucas (1989), or Ljungqvist and Sargent (2004) for excellent presentations of dynamic programming.

value functions:

$$V_t^k(S_t) = U_t^k(S_t) + \delta E\left[V_{t+1}(S_{t+1})\,|S_t, d_t^k = 1\right]. \tag{3}$$

In order to estimate the model, it is necessary to adopt explicit forms of the wage equation, the grade level progression function, and the consumption 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. In the estimation Section 2.2 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 Wages are assumed to depend on highest acquired degree, E_t, accumulated work experience, $H_t = H_{t-1} + h_{t-1}$, and an idiosyncratic labor market productivity shock, ε_t^w. Log wages are given by:

$$\ln W_t = \alpha_0 + \sum_{j=1}^{2} \alpha_j \mathbf{1}\left[E_t = j\right] + \alpha_3 H_t + \alpha_4 H_t^2 + \alpha_5 \mathbf{1}\left[h_t < 1\right] + \varepsilon_t^w \tag{4}$$

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. This approach thus takes the nonlinearities in rates of returns to education into account.[14] Note that the parameter α_5 shifts the intercept in the student employment wage equation relative to that for full-time wages.[15]

Students receive a wage offer with probability $p_w = 1$ each period and then decide

[14] See e.g. Hungerford and Solon (1987), Kane and Rouse (1995), Jaeger and Page (1996), and Park (1999) for evidence on sheepskin effects, and Heckman, Lochner and Todd (2006) for documentation of the importance of nonlinearities in the returns to education.

[15] Alternatively, I let individual wages also depend on the local labor market wage level, $\bar{W}_{lt}$, assumed to reflect the local price of human capital which provides the (sufficient, but not necessary) exclusion restriction for identifying the wage effect on the university-work choices. The key identifying assumption is that $\bar{W}_{lt}$ varies across localities because of different labor market conditions and as such is exogenous to individual university-work choices. However, in the present version I choose earnings as numeraire.

their degree of labor market participation.[16]

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: $h_t \in \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\}$.[17] 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: $h_t = 1$.[18] The labor market is assumed to be an absorbing state, hence, labor market opportunities are only explicitly specified during university enrolment. It is reasonable to treat full-time labor market work as an absorbing state, since transitions from full-time work back to full-time university education are very infrequent. Table 2 shows that 97% of those who work full-time in period t also work full-time in period $t + 1$.

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.[19] A level j degree is acquired if $G_t \geq \overline{G}^j$, i.e. highest

[16] Alternatively, I estimated p_w assumed to depend on the same variables as wage offers, whether the student is already employed, an i.i.d. logistically distributed idiosyncratic labor market opportunity shock, ε_t^p, and the youth unemployment rate in the local labor market, Z_{lt}, proxying local labor market opportunities assumed to affect individual university-work choices and outcomes only through the probability of receiving a wage offer. Since p_w was estimated to be 0.97 and it did not change any of the other results significantly, I assume $p_w = 1$ for parsimony.

[17] This is equivalent to the average number of hours worked per week taking on values in the intervals: $h_t \in \{0, (0; 10], (10; 19], (19; 28], (28; 37]\}$.

[18] This is a reasonable approximation, since less than 3% of university graduates have an unemployment spell of more than two months following graduation.

[19] Less than 2% of university students enroll in a PhD program. All PhD students at Danish universities get scolarships corresponding to the starting wage of a well-qualified state employed Master graduate. Furthermore, the PhD study entails a certain amount of predetermined TA 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 university-work choices during this period.

completed degree corresponds to:

$$E_t = \begin{cases} 0 & , \ if \ G_t < 18 \\ 1 & , \ if \ 18 \leq G_t < 30 \\ 2 & , \ if \ 30 \leq G_t \end{cases} . \tag{5}$$

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,[20] as well as highest acquired university degree, 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 course credits attained in later periods.[21] If the student fails a course, it is implicitly assumed that no incremental academic capital is produced. 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 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:

$$\begin{aligned} g_t^* &= \gamma_1 A + \gamma_2 K + \gamma_3 \mathbf{1}\left[E_t = 1\right] + \gamma_4 G_t + \gamma_5 t \\ &\quad + \gamma_6 \mathbf{1}\left[h_t \geq \frac{1}{4}\right] + \gamma_7 \mathbf{1}\left[h_t \geq \frac{1}{2}\right] + \gamma_8 \mathbf{1}\left[h_t \geq \frac{3}{4}\right] + \varepsilon_t^g \\ &= g\left(X_t\right) + \varepsilon_t^g \end{aligned} \tag{6}$$

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: $g \in \{0, 1, 2, 3, 4, 5, 6, 7\}$.[22] Therefore the grade level progression is

[20] High school GPA is used as a proxy for initial 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 (2009) 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.

[21] 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.

[22] I recognize the potential importance of student grades and 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, I believe that I capture the most essential features of grade level progression,

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.[23]

Preferences Individuals choose to divide each year between units of time devoted to work and non-work. 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, $\bar{b}$. 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.[24] Each period's consumption is assumed to be the sum of earnings and

because of the substantial degree wage premium and because attainment of university degrees only requires accumulation of a requisite number of course credits. Consult Eckstein and Wolpin (1999) for a similar model with a more detailed modeling of the entire grade distribution.

[23]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 19 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 student 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 5.4.

[24]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 wealth and academic and labor market choices and outcomes, respectively, in the data. The same holds true for parental wealth.

grants. Hence, the (non-linear) budget constraint assumed to be satisfied each period is given by:

$$C_t = Y_t + \bar{b}\left(Y_t, E_t, t\right) D_t$$

where $Y_t = 1739 W_t h_t$ denotes earned income.[25] All admitted students are eligible to receive a study grant specified by:

$$\bar{b}\left(Y_t, E_t, t\right) = \beta_7 \mathbf{1}\left[Y_t \in y_s\right] b_s \cdot \left(\mathbf{1}\left[t \leq 4\right] \mathbf{1}\left[E_t = 0\right] + \mathbf{1}\left[t \leq 6\right] \mathbf{1}\left[E_t = 1\right]\right). \tag{7}$$

The grant can be received 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. It also depends on the amount of student employment in the year, since it is a decreasing function of student earnings. Figure 5 displays the relationship between study grant intervals, b_s, and the corresponding allowable earnings intervals, y_s. Note that the parameter β_7 is a multiplier converting study grant DKK into a monetary equivalent consumption value. Hence, β_7 reflects the fact that there may be a difference in salience between direct costs (forgone study grants) and indirect costs (forgone earnings).[26]

Utility in period t is assumed to be linear and additive in consumption, job finding cost, and value of university attendance.[27] The alternative specific flow utility can be compactly written as:

$$U_t^k = C_t + s\left(1 - D_t\right)\left(1 - h_{t-1}\right) + b_t D_t + \varepsilon_t^k, \tag{8}$$

$$b_t = b_t^1 + b_t^k d_t^k \mathbf{1}\left[h_t > 0\right], \tag{9}$$

$$b_t^1 = \beta_0 + \beta_1 A + \beta_2 K + \beta_3 t, \tag{10}$$

$$b_t^k = \beta_1^k A + \beta_2^k K + \beta_3^k t + \sum_{j=1}^{3} \beta_{3+j}^k \mathbf{1}\left[h_{t-1} = \frac{j}{4}\right], \tag{11}$$

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 s. This cost is linear in the amount of work in the last period. That is, if the student is not working, the cost of entering the labor market and finding a full-time job is twice what the cost would be if the student had a half-time job.

[25] The implicit assumption is that a full-time job comprises $47 * 37 = 1739$ hours per year.

[26] The model is also estimated constraining no difference in salience, $\beta_7 = 1$.

[27] This makes earnings risk and timing of consumption irrelevant for student choices.

The utility from attending education in any year also depends on the value attached to effort and learning. Time devoted to education is valued at b_t DKK per unit of time (relative to working) and specified in equations (9), (10) and (11). 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. (9) shows that b_t consists of a basic component common to all education alternatives, b_t^1, and an additional component if the student works, b_t^k, $k \in \{2, 3, 4\}$. b_t^1 is allowed to depend on ability, skills, and time since university entry as specified in (10).[28] Working may 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 may 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) as specified in (11).

The alternative specific preference shocks, ε_t^k, capture the fact that new information about alternative specific tastes is revealed to students each period. These taste shocks may affect alternative specific utilities, and are treated as state variables that are unobserved to the econometrician but revealed to students just before they make their choices at time t.

3.1.1 Economics of the Model

Being enrolled at the university provides 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 accumulated labor 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

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

producing academic capital (course credits leading to degrees) - both of which enhance future wages.

In the basic model, students have heterogeneous initial characteristics, A and K, that affect their consumption value of education, b_t, as well as their academic achievement, G_t, and accordingly their labor market opportunities through acquired degrees. Furthermore, they also have heterogeneous initial experience, 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 heterogeneous in their consumption value of education, β_0, and academic abilities, γ_0, that affect their university attendance and academic success and thereby their wages, as well as earnings ability, α_0, that directly affects their wages and their outside opportunities while enrolled at the university.

The model provides a number of explanations as to why student employment increases monotonically with time since enrollment, cf. Figure 2. *(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 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 enrollment.

Individuals will stay enrolled at the university as long as their expected returns are larger than 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 their 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 pecuniary 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 expectations about future academic achievement, which is the key uncertain component in the state 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 probability of receiving higher wages, hence they also affect the extent of student employment which in turn affects grade level progression. Hence there is a 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 grade level progression, 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.

3.2 Solution and Estimation

The dynamic programming problem (2) can be solved by forward recursion.[29] 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. The problem becomes trivial after university exit, since the labor market is an absorbing state, $D_{\tilde{t}+\tau} = 0$, $\forall \tau \geq 0$. Hence, the conditional value function for the full-time working alternative becomes particularly simple:

$$V_t^0\left(S_t\right) = U_t^0(S_t) + \sum_{\tau=t+1}^{\infty} \delta^{\tau-t} E\left[W_\tau\left(X_\tau^w, \varepsilon_\tau^w\right)\right] \tag{12}$$

where $X_\tau^w = (E_t, H_{t+\tau})$. Hence, the only state variable that evolves is accumulated labor market experience and it does so deterministically: $H_{t+\tau} = H_t + \tau - t$, $\forall \tau \geq 1$. When calculating the expected value of lifetime earnings, wages are assumed to be constant after 25 years of labor market experience. Given that wages are log-normally distributed, but enter in levels in the optimization problem, the wage expectation includes the variance of the idiosyncratic labor market productivity shock: $\exp(\alpha X_\tau^w + \sigma_w^2)$. Note that year effects cannot be included in the calculation, since these cannot be forecasted beyond the sample period.

[29]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:

$$V_t^k(S_t) = U_t^k(S_t) + \delta \sum_{g=0}^{7} P\left(g_t = g | S_t, d_t^k = 1\right) \cdot \max_\kappa V_{t+1}^\kappa \left(S_{t+1}\right) \tag{13}$$

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 $t+1$.

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 that the choice specific components of utility not observed by students before they make their university-work choices at time t, ε_t^k, are also the utility components not observed by the econometrician (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, $F\left(\varepsilon_t^k\right) = \exp\left(-e^{-\varepsilon_t^k}\right)$. One great simplification provided by this assumption (together with the assumptions of additive separable utility and conditional independence) is that the $E\max$ in (3) becomes a closed form expression:

$$\begin{aligned} E\left[V_{t+1}\left(S_{t+1}\right) | S_t, d_t^k = 1\right] &\equiv E\left[\max_\kappa V_{t+1}^\kappa \left(S_{t+1}\right) | S_t, d_t^k = 1\right] \\ &= \gamma + E\left[\ln\left(\sum_\kappa \exp\left(V_{t+1}^\kappa \left(X_{t+1}\right)\right)\right) | X_t, d_t^k = 1\right] \end{aligned} \tag{14}$$

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

Not only do these admittedly restrictive assumptions result in a substantial computational gain in terms of solving the dynamic programming problem (2), they also provide a simple analytical form for the conditional choice probabilities. The conditional probability of choosing alternative k is given by:

$$P\left(d_t^k = 1 | X_t\right) = \frac{\exp\left(U_t^k\left(X_t\right) + \delta E\left[V_{t+1}\left(X_{t+1}\right) | X_t, d_t^k = 1\right]\right)}{\sum_{\kappa=0}^{4} \exp\left(U_t^\kappa\left(X_t\right) + \delta E\left[V_{t+1}\left(X_{t+1}\right) | X_t, d_t^\kappa = 1\right]\right)} \tag{15}$$

Note that this is the same form as the upper part of a multilevel nested logit.[30] The difficulty in calculating the current choice probabilities arises because they depend on future expected utilities; hence the computation requires that 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 (CCPs) 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 (14) can be used in the calculation of the probability statements in (15), which are necessary to evaluate the likelihood function. The CCPs are then treated as nuisance parameters in the estimation. Hotz and Miller (1993) show that Bellman's equation (3) can always be written as a function of current utilities and future CCPs. Note that given the model assumptions, the term inside the $\ln(\cdot)$ in (14) is the denominator of the CCP 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 (14) can be written as:

$$
\begin{aligned}
& E\left[V_{t+1}(X_{t+1})|X_t, d_t^k = 1\right] && (16) \\
= \ & \gamma + \sum_{g=0}^{7} P\left(g_t = g|X_t, d_t^k = 1\right) \ln\left(\sum_\kappa \exp\left(V_{t+1}^\kappa(X_{t+1})\right)\right) \\
= \ & \gamma + \sum_{g=0}^{7} P\left(g_t = g|X_t, d_t^k = 1\right)\left(V_{t+1}^0(X_{t+1}) - \ln(\xi_0(X_{t+1}))\right)
\end{aligned}
$$

Consequently, the value function 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. $\xi_0(X_{t+1})$ can be thought of as the discrete hazard function for exiting the university. This hazard function represents the probability that a student 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 enrollment up until then).

The grade level progression probability, $P\left(g_t = g|S_t, d_t^k = 1\right)$, controls student ex-

[30] Consult McFadden (1978, 1981) for details and a derivation of the results in (14) and (15).

pectation 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 (12). Finally, this implies that the conditional probability of choosing alternative k that enters the likelihood function is given by substituting (12) and (16) into (15).

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 (2008), 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.

3.2.1 Estimation

The parameters of the structural model are estimated by a maximum likelihood based procedure. The model basically requires two types of parameters to be estimated: utility function (preference) parameters: β's (and α's), and transition parameters: γ's and α's.[31] The transition parameters are used in forming expectations about uncertain future events; these include the γ parameters in the course credit generating process (6) through which students learn about their academic abilities, as well as the α parameters of the wage process (4) that form student expectations about future wages. Note that one important feature of the model is that the wage equation (4) is both part of the law of motion and an important part of utility.[32]

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

[31] The basic model has 24 parameters to estimate: $\beta = (\beta_0, \beta_1, ..., \beta_7)$, $\gamma = (\gamma_1, ..., \gamma_8)$, $\alpha = (\alpha_0, \alpha_1, ..., \alpha_5)$, σ_w, and s.

[32] 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.

conditional on the model parameters. Because of the additive separability and conditional independence assumptions, the individual likelihood contribution, $\mathcal{L}_i(\theta)$, 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 factors into:

$$\begin{aligned}
&\mathcal{L}_w(\alpha) - \text{ the likelihood contribution of wages}\\
&\mathcal{L}_g(\gamma) - \text{ the likelihood contribution of grade level progression}\\
&\mathcal{L}_d(\theta) - \text{ the likelihood contribution of utility (university-work choices)}
\end{aligned}$$

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

$$\begin{aligned}
\ln \mathcal{L}(\theta) &= \ln \prod_{i=1}^{N} \left(\mathcal{L}_{gi}(\gamma) \times \mathcal{L}_{wi}(\alpha) \times \mathcal{L}_{di}(\theta) \right) \qquad (17)\\
&= \sum_{i=1}^{N} \left(l_{gi}(\gamma) + l_{wi}(\alpha) + l_{di}(\theta) \right),
\end{aligned}$$

where $l_i(\theta) = \ln \mathcal{L}_i(\theta)$. 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 - α through the wage equation and γ 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 γ and α parameter vectors by maximizing $\mathcal{L}_g$ and $\mathcal{L}_w$ separately. Then using the $\hat{\gamma}$ and $\hat{\alpha}$ parameter estimates, I can consistently estimate the CCPs, $\hat{\xi}$, and preference parameter vectors β by maximizing $\mathcal{L}_d\left(\hat{\gamma}, \hat{\alpha}, \beta, \hat{\xi}\right)$. 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 $\delta = 0.95$ 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, the discounted one-period ahead expected value of entering the labor market, and the conditional probability of this event.

3.2.2 Unobserved Heterogeneity

Given the diversity of individual characteristics at university entry, it is unlikely that individuals have the same preferences for education, and unobserved abilities (with respect to education and work). Hence, it seems important to account for persistent unobserved 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). I assume 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 and who spends excess time in university education? How do these individuals differ from timely graduates in terms of unobserved persistent initial traits and how are those traits related to observed family background characteristics? Secondly, which initial traits are important in explaining the propensity to drop out or the excess time-to-graduation?

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 (4). Persistent academic abilities 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 (6). Likewise, the consumption value of attending university in equation (9) is allowed to differ by unobserved type, $\beta_0 = \sum_{m=1}^{M} \beta_{0m} \mathbf{1}\left[type = m\right]$.[33] 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 or motivation, and their preferences for university attendance.

To conserve on parameters and avoid identification issues, I consistently only allow

[33] Hence, the model has $21 + 3M$ parameters to estimate.

for first-order heterogeneity effects. This approach is common in the literature; see e.g. Eckstein and Wolpin (1999). However, there may be type-specific effects of student employment if it is more valuable for types who are more likely to drop out.

3.2.3 CCP Estimation with Unobserved Heterogeneity

The model with unobserved heterogeneity is estimated using the strategy of Arcidiacono and Miller (2008). 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 ability, academic ability, or consumption 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 $(X_{it}, type_i = m)$, 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 sequence $X_i = (X_{i1}, ..., X_{iT_i})$ and unobserved type m, is obtained by forming the product over the T period likelihoods. The sample log likelihood is thus given by:

$$\ln \mathcal{L}\left(\Theta\right) = \sum_{i=1}^{N} \ln \left(\sum_{m=1}^{M} \pi_m \prod_{t=1}^{T_i} \mathcal{L}_{imt}\left(O_{it}|X_{it}, type_i = m; \theta, \pi, \xi\right) \right). \tag{18}$$

where $\Theta = (\theta, \pi)$. 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 (18) directly is to iteratively maximize the expected log likelihood function, where the n^{th} iteration involves maximizing:

$$\ln \mathcal{L}\left(\Theta\right) = \sum_{i=1}^{N} \sum_{m=1}^{M} \sum_{t=1}^{T_i} q_{im}^{(n)} \ln \mathcal{L}_{imt} \left(O_{it}|X_{it}, type_i = m; \theta, \pi^{(n)}, \xi^{(n)}\right) \tag{19}$$

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. $q_{im} \equiv q_{im}\left(O_i, X_i; \theta^{(n)}, \pi^{(n)}, \xi^{(n)}\right)$ denotes the probability that individual i is of type m given 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 type probabilities, q_{im}, are taken as given and 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 (22).

The estimation algorithm is triggered by setting initial values for the CCPs, $\xi^{(1)}$, the sample proportion of each unobserved type, $\pi^{(1)} = \frac{1}{m}$, 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 2.2.1 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 $q_{im}^{(n+1)}$ 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:

$$q_{im}^{(n+1)} = \frac{\pi_m^{(n)} \prod_{t=1}^{T} \mathcal{L}_{imt}\left(O_{it}|X_{it}, type_i = m; \theta^{(n)}, \pi^{(n)}, \xi^{(n)}\right)}{\sum_{m=1}^{M} \pi_m^{(n)} \prod_{t=1}^{T} \mathcal{L}_{imt}\left(O_{it}|X_{it}, type_i = m; \theta^{(n)}, \pi^{(n)}, \xi^{(n)}\right)}. \tag{20}$$

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 is the type-specific equivalent.

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

$$\pi_m^{(n+1)} = \frac{1}{N} \sum_{\iota=1}^{N} q_{im}^{(n+1)}. \tag{21}$$

Step 3 Maximize the expected log likelihood function (19) to obtain parameter estimates $\theta^{(n+1)}$; given $q_{im}^{(n+1)}$, $\left(\pi^{(n+1)}, \xi^{(n)}\right)$, and (O_i, X_i).

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

$$\xi_{0Xm}^{(n+1)} = P\left(d_{it}^{0} = 1 | X, m; \theta^{(n+1)}, \xi^{(n)}\right) = \mathcal{L}_{d^0}\left(X, m; \theta^{(n+1)}, \xi^{(n)}\right) \tag{22}$$

Arcidiacono and Miller (2008) show that this algorithm converges to a fixed point and is computationally feasible for many problems with the finite time dependence property. The great 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 (17).

3.3 Identification

The endogenous variables in the model include current university enrolment and employment, as well as accumulated course credits and labor market experience. I control for endogeneity using all the restrictions implied by the economic model, by modeling the entire university enrolment period as a sequence of endogenous choices that drive subsequent outcomes. Furthermore, identification is achieved through an exogenous change in study grant rules.

Identification of the wage offer and grade level progression functions rests on variation in wages, work 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 wages, 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 parameters, γ_{0m}, are 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 university-work choices and 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. The identification of the study grant effect on university-work choices is further provided by a change in the threshold levels for maximum allowable amounts of grants and earnings, cf. details in the following sections.

4 Data

The model is estimated using a very rich register-based panel data set comprising a random 10% sample of the Danish population. 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 actual labor market experience, labor income and wages for the period 1984-2004.[34] Complete detailed educational event histories are observed for each individual from 1978 to 2005, indicating level of education, field of study, institution, and the dates of entry and exit, along with 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

[34] All incomes are observed at year-end and deflated to real values measured in 2000 DKK using the average consumer price index, $PRIS8$, from Statistics Denmark.

collected for the particular purpose of this study.[35] Since educational event histories are available on a monthly basis but accumulated course credits, income, and the other socioeconomic background variables are available only on a yearly basis, I have to take a stance on the timing of the educational events. I assume that an individual is in full-time education if the individual is enrolled at the university more that half of the calendar year.

The data set also contains course choices in high school and high school GPA.[36] The GPA is a weighted average of grades at final exams of each course. All high school courses can be obtained on three different levels according to difficulty: low, medium, and high. Course quality and GPA are comparable across high schools because of centralized administration and examination procedures.

As Danish university entrants are solely screened on high school GPA and course choices, and admission requirements are openly available, educational choice sets for potential students can be determined precisely. 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; to obtain a Master's degree, 300 ECTS have to be accumulated. Most programs are designed to graduate in five years.[37] Although students can stay enrolled as long as they wish.

There are no tuition fees in Denmark, and students over 18 receive a study grant from the government.[38] 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.[39] All enrolled students are eligible to collect the study grant for a maximum of 70 months. The grant is reduced if income from student employment exceeds a certain threshold: the threshold level for permissible student earnings was raised from $47,000$ DKK to $60,000$ DKK in 1996, while the maximal study grant remained unchanged at $47,000$ DKK per

[35] I am working on further augmenting the data with information on grades from university courses.

[36] 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.

[37] Medicine is an exeption, requiring six and a half years, since the last year and a half consists of mandatory vocational training.

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

[39] All university students in the sample are above 18 years old, and 99% do not live with their parents.

year; cf. Figures 5 and 6. This threshold change may increase student incentive to work more and is incorporated into the structural estimation strategy.

4.1 Sample Selection

Among high school graduates eligible to enter university, I select those not older than 22 when initially enrolling between September 1994 and 1996, as course credit data is available only from 1995 onwards. These entrants are observed until the end of 2004. The sample comprises $2,129$ individuals, with $19,349$ observations over time.

4.2 Descriptive Statistics

The average individual in the sample enters the university in 1995 and is 21 years old at the time of initial enrolment. The prerequisites from 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% drop out, 22% acquire a Bachelor's degree, while 55% acquire a Master's degree as their highest completed university degree. The top part of Table 1 presents student characteristics at university entry. Master graduates have significantly higher GPA, Math and Science level from high school.[40] 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. The fact that Bachelor graduates at university exit have accumulated 8 course credits on average more than required to obtain the Bachelor's degree could indicate that many of them are dropouts from the Master program. Figure 4 further reveals that 8% are still enrolled 10 years after uni-

[40] 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).

versity entry leading to some right censored 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. Dropouts tend to work more during their studies, accumulating around twice as much work experience each year of university enrolment. On average, students accumulate the equivalent of one year of full-time labor market experience through student employment.

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. This suggests that the primary pecuniary value of a Bachelor's degree is the option value associated with pursuing further university education.[41]

Annual labor market experience while enrolled in full-time education is shown in Figure 2. 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. Figure 2 and Figure 3 further illustrate 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. Figure 3 shows the average course credit accumulation over time since university entry for those still enrolled at the university. Students who work less than 10 hours a week do not seem to perform worse academically - if anything they perform better. However, students who work more than 19 hours a week have very low academic achievement. The figure 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. Figure 2 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 employ-

[41] 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).

ment 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 or motivation rather than the acquisition of skills that are complementary to academic achievement. The estimated model will take this selection into account.

Figure 4 displays the university-to-work transition separately for individuals choosing each of the five possible university-work alternatives. 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. The amount of student employment at Danish universities is similar to US 4-year colleges, where Bound et al. (2007) document that around 40% of students are employed, and 10% of students work more than 20 hours a week.

Figure 4 also reveals that more than 70% of students are still enrolled at the university five years after initial enrolment, although only 55% obtain a Master's degree. Although students do not produce course credits in the excess enrollment time, they may acquire other skills, e.g. through student employment. Particularly, dropouts both have the highest excess enrollment time and student employment. This could be because of high technological uncertainty associated with university education investment, which makes it valuable 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.[42]

Figure 4 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 university-work alternatives. This is crucial for the identification of the parameters of the structural model.

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

[42] 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), and Stange (2011) demontrates the empirical importance of the option value of college enrolment.

5 Empirical 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 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 structural model parameters 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 parameters from the model without unobserved heterogeneity are presented in column one, and column two (three) presents the parameters from the model with unobserved heterogeneity through a mixture of two (three) types. In what follows, the focus will be on the model with two types.

Wage Parameters 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 13 and 30 percentages, 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 percentages, which is very high, but not surprising given the selective sample of relatively high skilled individuals in their early career. Part-time jobs pay 20 percent lower wages than full-time jobs.

Grade Level Progression Parameters 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. Regarding student employment, it is found that working the equivalent of 10 hours a week significantly increases academic achievement, while working additional hours tends to be 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, and it increases the probability of passing all the courses in the year by 3 percentage points. Working additional 9 hours per week significantly increases the probability of accumulating zero course credits by 4 percentage points, and reduces the probability of producing six course credits by 2 percentage points. Hence, working the equivalent of 19 hour per week reduces the probability of not passing any course by 3 percentage points and increases the probability of passing all courses by 1 percentage point. Working more than 19 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.

Preference Parameters The estimates of the parameters in the wage equation reveal that student employment increases wages through acquired experience. The estimate of the job finding cost, s, 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.

The model allows for a direct utility (or disutility) flow from university attendance. This consumption value of university attendance is found to be increasing in initial ability and Math skills, while it decreases with time since initial enrolment. High ability students value university attendance at $460DKK$ more, and Math level students at $556DKK$ more, in terms of yearly consumption. 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 so much for students who have been enrolled at the university for longer. 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 $k \in \{2, 3, 4\}$. This further corroborates that there is evidence of substantial state dependence.

Finally, a higher level of study grant tends to increase the likelihood that students will stay enrolled at the university.

The model that allows for unobserved heterogeneity through a mixture of two unobserved types reveals that type 2 students have significantly higher (unskilled) labor market ability, $\alpha_{01} < \alpha_{02}$, higher unobserved academic ability or motivation, $\gamma_{01} > 0$, and also tend to 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 most significant change is that the job finding cost and the study grant effect diminish substantially.[43]

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 in equation (20). 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 preferences for university education.

Type 2 individuals who seem to be academic types comprise 88% of the sample. Table 4 presents observed differences between the two types. The first part of the table concerns own characteristics, and the last part concerns parental characteristics. Type 2 individuals have less work experience (although same age) at university entry, have higher high school GPA and Math level, and are much more likely to be in Health Sciences (the most lucrative field). The most pronounced difference between the two types is that 91% of type 1 individuals drop out, while 61% of type 2 are Master graduates. Furthermore, type 2 individuals are more likely to have financially well-off parents with a higher education.

Table 5 provides further evidence on initial conditions. It shows that parental education is an important determinant of high school outcomes, since individuals with higher GPA and Math level are more likely to have parents with higher education. Likewise, parental background has some effect on unobserved type probability, but conditional on high school outcomes, parental background is insignificant.

[43] Preliminary results from introducing a third type reveal that there are two academic types that only differ significantly in (unskilled) labor market ability, since: $\alpha_{01} > \alpha_{02} > \alpha_{03}$, $\gamma_{02} < \gamma_{01} = \gamma_{03} = 0$, and $\beta_{02} < \beta_{01} = \beta_{03}$.

Unobserved types play a significant role in explaining dropout and graduation behavior (as seen from Table 4), but how do they compare to other initial conditions. Table 6 answers this question for three measures of academic achievement. All estimations are first performed only with parental background characteristics as controls and then the type 1 probability is added as an additional control variable. The three academic outcomes are indicator variables for: *(i)* dropping out of the university, *(ii)* acquiring a Master's degree, and finally *(iii)* excess months to Master graduation. Table 6 reveals that individuals with higher type 1 probability have significantly higher probability of dropping out, significantly lower probability of Master graduation, and are also more prone to spend excess time-to-graduation. This corroborates the fact that type 1 individuals are low academic achievers. 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 and Master graduation probability, but when controlling for the type 1 probability, the parental background effect is jointly insignificant. Apparently, the unobserved type embodies 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. Parental background only predicts 2% (1%) of the variance in dropout (Master graduation) behavior, while it jointly with the type 1 probability predicts 36% (15%) of the variation. All in all, initial conditions - particularly unobserved heterogeneity - is important for explaining variation in dropout and graduation behavior. However, initial conditions do not predict much of excess-time-to-graduation behavior, which to a larger extent is explained by other factors, such as field of university study.

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. Furthermore, this comparison is done by demographic characteristics which are not explicitly incorporated in the model.

Table 7 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 8 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 it overestimates this probability in the later enrolment years. The 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 prescribed years for Bachelor and Master graduation. However, when unobserved heterogeneity is introduced the model predicts the timing of university exit very precisely.

Table 9 compares observed and predicted choices and outcomes for students with parental wealth below the 10^{th} and above the 90^{th} percentile, respectively. The students with wealthiest parents are less likely to be credit constrained. Table 9 shows that the wealthy students are more likely not to work, work less during university enrollment, and produce slightly more course credits. These differences are not very significant, hence it is unlikely that the observed patterns are due to credit constraints. Despite not using parental background in the estimation, the predictions from the structural model qualitatively captures all the observed differences over parental wealth. This out-of-sample model fit gives more confidence in policy simulations based on the model also capturing other important distributional differences.

5.4 Robustness

The basic model is estimated on samples stratified by field of university study in order to assess the robustness of its predictions for subsamples of students who differ substantially in terms of educational preferences, abilities, skills, dropout rates, times-to-graduation, wages, and the amount and type of student employment.

Table 10 describes some of the student traits that differ across fields. The fields are ranked in terms of wages - with the fields furthest to the right being the most lucrative. Table 10 shows that average yearly earnings vary from around 150,000 DKK (Humanities / Arts / Education) to almost 270,000 DKK (Health Sciences).[44] 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. The reasons for dropping out, the required skills, as well as employment prospects may also differ across fields. The fact that times-to-graduation tend to decrease with post-graduation earnings suggests the importance of option values. The value of staying enrolled is higher if there is higher post-graduation employment uncertainty, because of the value of waiting for more information on labor market opportunities.[45] Student employment might be an effective device to resolve this uncertainty.

Business students stand out regarding student employment. They work the equivalent of 27% of full-time employment each enrollment year, which is around twice as much as average students in 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.[46] However, fewer of them work in public services and in on-campus jobs. On the other hand, Health Sciences and Natural Science / Engineering students are to a larger extent employed in high-skilled student jobs on campus.

Table 11 presents the parameter estimates from the basic model estimated separately on each of the six fields of university study. 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,

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

[45] 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.

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

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 for grade level progression. Overall, there do not seem to be differential effects of student employment over fields, and basically all the conclusions remain. This deems it unlikely that the estimates are confounded by unobservable student traits or unconventional behavior, such as students enrolling just to hang out and collect grants or students staying enrolled to avoid unemployment after graduation. This section thus corroborates that the parameter estimates are not artifacts of confounding factors.

5.5 Policy Effects

Having assessed model fit and robustness, it is of obvious interest to evaluate the effects of potential public policies. Although it has been the subject of numerous studies and is still much debated, not much is known about the impact of potential policy interventions. Changing the study grant system is the most obvious policy instrument in order to increase university graduation rates and decrease times-to-graduation. With the structural parameter estimates at hand, it is possible to determine the extent to which restrictions on student employment would affect dropout rates and times-to-graduation - either directly by putting restrictions on hours worked while attending university or indirectly by changing study grants to induce less student employment. The fact that study grants tend to increase the probability of attending the university, suggests that policies aimed at reducing study grants (or increase tuition costs) could reduce times-to-graduation. However, the overall effect of tilting study grants to punish student employment harder is ambiguous, because of the non-linear effect of student employment on grade level progression.

As a first step, the data allows for an evaluation of an actual policy implementation that increased maximum allowable student earnings threshold. Before 1996 students could earn up to $47,000$ DKK a year while receiving full benefits. In 1996 this threshold was raised to $60,000$ DKK a year, giving students an incentive to work more while enrolled at the university. The amount of benefits received was unchanged at around $47,000$ DKK per year during the whole period, cf. Figure 1.[47] Using

[47]Smoothing the study grant function of income around the "notches", the study grant can be seen

$Z_t = \mathbf{1}[year \geq 1996]$ as an instrument for the amount of student employment would identify the effect of student employment on accumulated course credits (and wages) for those students who were induced to work more (or less) because of the higher threshold. Comparing first-year work choices and course credits for the cohorts who entered university in September 1994 and 1995, respectively, the change in benefit rules had no significant effect on mean student labor supply.[48] However, as is evident from Figure 5, the change in study grant rules has an ambigious effect on student labor supply. Figure 5 displays total student income as a function of working hours for a student with an hourly wage of 140 DKK. It is seen that there is no effect on the extensive margin of student labor supply, as students working 339 hours (i.e. $1739 - 1400$) a year or less are not affected. Students working slightly more than 1039 hours (i.e. $1739 - 700$) a year, who were far from being eligible before, but become closer to eligibility after the change, may reduce working hours. Students working an intermediate amount of hours may work more or less, depending on whether the substitution effect dominates the income effect. Furthermore, students with lower wages are more responsive to the change, since the hour span over which they are potentially affected is larger.

Having estimated the structural model allows for a more elaborate evaluation of the effects of changing the benefit system. Table 12 present the simulated impacts of changing different aspects of the educational environment on four academic outcomes: dropout rates, Master's graduation rates, and the fraction of students with excess-time-to-graduation exeeding 0 and 1 years, respectively.[49] The first column of the table represents the benchmark values. Each policy simulation changes one aspect of the educational environment, holding all other variables at their benchmark values. The simulated study grant systems are approximately cost neutral in order to avoid dealing with issues of public finances. Simulations 1-3 represent the effects of gradually tilting the study grant system towards students who work fewer hours. Overall, tilting the benefits towards those who devote less time to working does not alter academic outcomes significantly. The only outcome that is significantly improved is the Master's

as an Earned Income Tax Credit (EITC) plus a Negative Income Tax (NIT); see e.g. Moffitt (2003). Hence, the policy change can be seen as an increase of the guaranteed study grant by 28% together with a reduction in the implied NIT rate by 17 percentage points (from -73% to -56%)

[48] This may suggest that identification of the study grant effect on university-work choices, β_7, is mainly driven by variation in study grant because of academic timing and wage variation for a given alternative.

[49] Each number in the table is based on a simulation of 2000 individuals of each type, weighted by their estimated type proportion.

graduation rate that increases from 0.55 to 0.60 if students who do not work receive a study grant of $70,000$ DKK a year, students who work $1-10$ hours a week receive a study grant of $35,000$ DKK a year, and students who work more get no study grant.

Overall, a combination of tilting study grants and improving student abilities early on would somewhat increase graduation rates and times-to-graduations, but not nearly enough to level with the typical policy goals.

5.5.1 Timing and Incentives

Given the ineffectiveness of tilting study grants, changing the study grant scheme towards merit aid or change the timing of grants towards a front- or backloaded schemes could be more effective policy devises. This subsection provides evidence on such very much coveted policy effects on a more general level.[50]

First, the impacts of including merit aid by conditioning study grants on accumulating at least $g \in \{3, 4, 5\}$ course credits a year are evaluated. Merit aid based study grants are becoming more common as most US states fund merid aid programs, but are still not widespread in Europe. Simulations 4-6 in Table 12 represent the effects of gradually increasing course credit requirements. Overall, linking study grants to the speed of course credit accumulation substantially increases academic achievement; e.g. conditioning eligibility on passing at least two thirds of yearly courses, $g \geq 4$, decreases dropout rates by 3 percentage points, increases MSc graduation rates by 4 percentage points, decreases the number of late MSc graduates by 6 percentage points.

Second, the impacts of education policies with a temporal dimension are evaluated. Frontloading study grants explicitly alters the grant gradient by reducing the cost of the first years of university; e.g. community colleges. Backloading study grants directly alters financial gains to final years of university attendance, but not prior years. Simulations 7-8 in Table 12 represent the effects of increased backloading. Providing timely graduation bonuses also tends to substantially speed up graduation, increase graduation rates, and reduce dropout. Doubling the study grant in the final year $t = 5$ if acquiring a timely MSc degree, increases MSc graduation rates by 3 percentage points and decreases the number of late MSc graduates by 5 percentage points.

[50] Angrist et al. (2009), DesJardins and McCall (2009), and Scott-Clayton (2011) demonstrate the potential effectiveness of providing incentives related to merit and timing in particular financial aid packages at the University of Toronto, University of Michigan, and the PROMISE scholarship in West Virginia, respectively.

Preliminary policy simulations show that the main impact of frontloading is that students drop out later, but without acquiring a higher degree. On the other hand, backloading tends to induce students to graduate sooner. Likewise, merit aid schemes that condition on accumulating 4 or 5 course credits a year tend to increase overall academic achievement.

6 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 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 goals.

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 cognitive skills at younger ages.

Excessive student employment of more than 19 hours a week is found to be very detrimental to academic achievement. However, student employment of a moderate number of hours significantly increases academic achievement and future labor market outcomes. Therefore tilting study grants towards students who work fewer hours is not a very effective devise to attain the social goals.

This paper sheds more light on how public study grants can be spent more productively by redistribution across time and across students. Targeting study grants towards students who meet particular academic standards tends to be an effective policy devise. Particularly, timely graduation bonuses and merit aid may not only relax students' credit constraints, but also provide incentives to meet palpable academic goals.

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] Angrist, J. D., Bettinger, E., Bloom, E., King, E., and Kremer, M. (2002) "Vouchers for Private Schooling in Columbia: Evidence from a Randomized Natural Experiment", *American Economic Review*, 92(5): 1535-1558.

[5] Angrist, J. D., D. Lang and P. Oreopoulos (2009) "Incentives and Services for College Achievement: Evidence from a Randomized Trial", *American Economic Journal: Applied Economics*, 1(1): 136-163.

[6] Arendt, J. N. (2008), "The Impact of Public Student Grants on Drop-out and Completion of Higher Education - Evidence from a Student Grant Reform", HE paper 2008:10, University of Southern Denmark.

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

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

[9] Arcidiacono, P. and R. A. Miller (2008), "CCP Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity", manuscript, Duke University.

[10] 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.

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

[12] 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.

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

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

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

[16] Bound, J., M. Lovenheim and S. Turner (2007), "Understanding the Decrease in College Completion Rates and the Increased Time to the Baccalaureate Degree", PSC Research Report No. 07-626.

[17] Bound, J. and S. Turner (2002), "Going to War and Going to College: Did World-War II and the G.I. Bill Increase Educational Attainment for Returning Veterans?", *Journal of Labor Economics*, 20(4): 784-815.

[18] 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.

[19] 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.

[20] 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.

[21] Cameron, S. and C. Taber (2004), "Estimation of Educational Borrowing Constraints Using Returns to Schooling", *Journal of Political Economy*, 112(1): 132-.

[22] Carneiro, P. and J. J. Heckman (2002), "The Evidence on Credit Constraints in Post-Secondary Schooling", *Economic Journal*, 112(482): 705-734.

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

[24] 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.

[25] 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.

[26] 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.

[27] DesJardins, S. L. and B. P. McCall (2009), "Simulating the Effects of Financial Aid Packages on College Student Stopout and Reenrollment Spells and Their Subsequent Graduation Chances", manuscript, University of Michigan.

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

[29] Dynarski, S. (2002), "The Behavioral and Distributional Implications of Aid for College", *American Economic Review*, 92(2): 279-285.

[30] Dynarski, S. (2003), "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion", *American Economic Review*, 93(1): 279-288.

[31] Dynarski, S. (2008), "Building the Stock of College-Educated Labor", *Journal of Human Resources*, 43(3): 576-610.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

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

[37] 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.

[38] 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.

[39] Heckman, J. 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.

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

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

[42] 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.

[43] 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.

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

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

[46] 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.

[47] Joensen, J. S. (2009), "Academic and Labor Market Success: The Impact of Student Employment, Abilities, and Preferences", manuscript, Stockholm School of Economics.

[48] Joensen, J. S. and H. S. Nielsen (2009), "Is there a Causal Effect of High School Math on Labor Market Outcomes?", *Journal of Human Resources*, 44(1): 171-198.

[49] Kane, T. and C. Rouse (1995), "Labor Market Returns to Two- and Four-Year College", *American Economic Review*, 85(3): 600-614.

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

[51] Leslie, L. L. (1984), "Changing Patterns in Student Financing of Higher Education", *Journal of Higher Education*, 55: 313-346.

[52] Leuven, E., H. Oosterbeek and B. van der Klaauw (2010) "The Effect of Financial Rewards on Students' Achievement: Evidence from a Randomized Experiment", *Journal of the European Economic Association*, 8(6): 1243-126.

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

[54] Ljungqvist, L. and T. J. Sargent (2004), "Recursive Macroeconomic Theory", 2nd Ed., The MIT Press, Cambridge, MA.

[55] 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.

[56] 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.

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

[58] Moffitt, R. A. (2003), "The Negative Income Tax and the Evolution of U.S. Welfare Policy", *Journal of Economic Perspectives*, 17(3): 119-140.

[59] Nielsen, H. S., T. Sørensen and C. Taber (2008), "Estimating the Effect of Student Aid on College Enrollment: Evidence from a Government Grant Policy Reform", Economics WP 2008-14, University of Aarhus.

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

[61] Park, (1999), "Estimation of Sheepskin Effects Using the Old and the New Measures of Educational Attainment in the Current Population Survey", *Economics Letters*, 62: 237-240.

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

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

[64] Scott-Clayton, J. (2011), "On Money and Motivation: A Quasi-Experimental Analysis of Financial Incentives for College Achievement", *Journal of Human Resources*, 46(3): 614-646.

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

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

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

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

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

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

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	8.89	9.13
	(0.89)	(0.80)	(0.72)
High school Math level	1.91	1.90	2.28
	(0.91)	(0.92)	(0.86)
High school Science level	1.73	1.72	2.04
	(0.83)	(0.80)	(0.84)
High school Social Science level	1.80	1.89	1.77
	(0.89)	(0.88)	(0.86)
High school Language level	2.09	2.29	2.21
	(0.80)	(0.75)	(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	14.29	14.46
	(8.89)	(9.28)	(8.59)
Accumulated work experience	0.75	0.59	0.55
	(0.68)	(0.63)	(0.54)
Age	21.22	21.17	21.14
	(0.77)	(0.79)	(0.76)
During University Enrolment:			
Target duration from entry until exit (months)	0	34	58
Duration from entry until exit (months)	71.83	83.08	78.19
	(25.42)	(32.19)	(14.22)
Accumulated course credits per year	2.32	3.92	4.70
	(2.49)	(2.41)	(2.08)
Accumulated work experience per year (years)	0.30	0.18	0.15
	(0.36)	(0.25)	(0.21)
At University Exit:			
Requisite course credits to acquire degree	0	18	30
Accumulated course credits	7.40	25.94	31.18
	(8.87)	(6.88)	(5.22)
Accumulated work experience (years)	1.24	1.26	1.26
	(1.23)	(1.21)	(0.99)
After University Exit:			
Hourly wages (real 2000 DKK)	147.97	153.51	193.12
	(68.31)	(77.84)	(72.38)
Yearly earnings (real 2000 DKK)	138,036	146,533	269,138
	(118,555)	(120,261)	(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)					
	$\mathbf{C}_{\mathbf{t}}$					
$\mathbf{C}_{\mathbf{t-1}}$	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)	Three types (M=3)
		w	
α_0	4.615 (0.012) ***		
α_{01}		4.533 (0.015) ***	4.703 (0.010) ***
α_{02}		4.659 (0.013) ***	4.596 (0.012) ***
α_{03}			4.382 (0.012) ***
α_1 (Bachelor degree)	0.157 (0.010) ***	0.133 (0.010) ***	0.150 (0.008) ***
α_2 (Master degree)	0.345 (0.013) ***	0.295 (0.014) ***	0.308 (0.011) ***
α_3 (Experience)	0.082 (0.007) ***	0.086 (0.007) ***	0.083 (0.005) ***
α_4 (Experience^2)	-0.002 (0.001) *	-0.002 (0.001) **	-0.002 (0.001) ***
α_5 (Part time work)	-0.160 (0.010) ***	-0.196 (0.011) ***	-0.151 (0.008) ***
		g^*	
γ_{02}		0.421 (0.207) ***	-2.726 (0.084) ***
γ_{03}			0.018 (0.039)
γ_1 (High school GPA=9)	0.216 (0.029) ***	0.218 (0.029) ***	0.171 (0.029) ***
γ_2 (High level Math)	0.276 (0.029) ***	0.277 (0.029) ***	0.245 (0.029) ***
γ_3 (Bachelor degree, $\mathbf{1}[E_t=1]$)	0.208 (0.037) ***	0.208 (0.037) ***	0.171 (0.037) ***
γ_4 (Accumulated course credits, G_t)	0.084 (0.002) ***	0.084 (0.002) ***	0.058 (0.003) ***
γ_5 (Time since enrolment)	-0.618 (0.012) ***	-0.618 (0.012) ***	-0.554 (0.012) ***
γ_6 (Student employment, h_t=¼)	0.301 (0.035) ***	0.301 (0.035) ***	0.264 (0.035) ***
γ_7 (Student employment, h_t=½)	-0.215 (0.045) ***	-0.216 (0.045) ***	-0.156 (0.046) ***
γ_8 (Student employment, h_t=¾)	-1.590 (0.051) ***	-1.589 (0.051) ***	-1.594 (0.052) ***
		$P(d^k=1)$	
s (Full-time job finding cost)	-5.088 (0.063) ***	-2.841 (0.061) ***	-2.806 (0.062) ***
β_0	-0.097 (0.065)		
β_{01}		-2.828 (0.090) ***	-1.041 (0.073) ***
β_{02}		-1.103 (0.072) ***	-3.207 (0.099) ***
β_{03}			-1.002 (0.090) ***
β_1 (High school GPA=9)	0.490 (0.054) ***	0.460 (0.053) ***	0.451 (0.054) ***
β_2 (High level Math)	0.601 (0.054) ***	0.556 (0.053) ***	0.544 (0.053) ***
β_3 (Time since enrolment)	-0.611 (0.012) ***	-0.390 (0.012) ***	-0.401 (0.012) ***
$\beta_1+\beta_1^2$ (High school GPA=9)	0.191 (0.055) ***	0.200 (0.055) ***	0.195 (0.055) ***
$\beta_2+\beta_2^2$ (High level Math)	0.234 (0.055) ***	0.224 (0.055) ***	0.213 (0.055) ***
$\beta_3+\beta_3^2$ (Time since enrolment)	-0.755 (0.013) ***	-0.536 (0.013) ***	-0.548 (0.013) ***
β_4^2 (Stud.empl. previous period, h_{t-1}=¼)	1.663 (0.048) ***	1.561 (0.048) ***	1.552 (0.048) ***
β_5^2 (Stud.empl. previous period, h_{t-1}=½)	1.762 (0.075) ***	1.439 (0.073) ***	1.427 (0.073) ***
β_6^2 (Stud.empl. previous period, h_{t-1}=¾)	1.775 (0.111) ***	1.218 (0.107) ***	1.222 (0.107) ***
$\beta_1+\beta_1^3$ (High school GPA=9)	-0.394 (0.061) ***	-0.388 (0.061) ***	-0.391 (0.062) ***
$\beta_2+\beta_2^3$ (High level Math)	-0.366 (0.062) ***	-0.375 (0.063) ***	-0.385 (0.063) ***
$\beta_3+\beta_3^3$ (Time since enrolment)	-0.753 (0.015) ***	-0.533 (0.015) ***	-0.545 (0.015) ***
β_4^3 (Stud.empl. previous period, h_{t-1}=¼)	1.152 (0.060) ***	1.053 (0.060) ***	1.044 (0.060) ***
β_5^3 (Stud.empl. previous period, h_{t-1}=½)	2.456 (0.078) ***	2.135 (0.077) ***	2.123 (0.077) ***
β_6^3 (Stud.empl. previous period, h_{t-1}=¾)	2.172 (0.120) ***	1.656 (0.116) ***	1.659 (0.116) ***
$\beta_1+\beta_1^4$ (High school GPA=9)	0.094 (0.078)	-0.368 (0.068) ***	-0.363 (0.069) ***
$\beta_2+\beta_2^4$ (High level Math)	0.269 (0.080) ***	-0.311 (0.070) ***	-0.313 (0.070) ***
$\beta_3+\beta_3^4$ (Time since enrolment)	-0.822 (0.017) ***	-0.482 (0.015) ***	-0.495 (0.015) ***
β_4^4 (Stud.empl. previous period, h_{t-1}=¼)	1.605 (0.102) ***	0.539 (0.089) ***	0.539 (0.089) ***
β_5^4 (Stud.empl. previous period, h_{t-1}=½)	2.932 (0.105) ***	2.127 (0.090) ***	2.122 (0.090) ***
β_6^4 (Stud.empl. previous period, h_{t-1}=¾)	4.213 (0.107) ***	3.487 (0.088) ***	3.499 (0.088) ***
β_7 (Study grant)	-8.145 (0.232) ***	-2.169 (0.119) ***	-2.208 (0.119) ***
π_1	1.00	0.12	0.74
π_2		0.88	0.10
π_3			0.15
Log Likelihood	-65536	-61480	-56035

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: Behavioral and Background Differences Between Unobserved Types.

	Mean (Standard Deviation)	
	Unobserved Type	
Individual Characteristics	Type 1	Type 2
At University Entry:		
High school GPA	8.54 (0.91)	9.03 (0.77)
High school Math level	1.92 (0.92)	2.14 (0.90)
Female	0.44	0.49
Accumulated work experience	0.81 (0.71)	0.58 (0.58)
During University Enrolment:		
Accumulated course credits per year	2.46 (2.44)	4.26 (2.33)
Accumulated work experience per year (years)	0.25 (0.31)	0.18 (0.25)
Field of Sudy:		
Humanities / Arts / Education	0.29	0.26
Life Sciences	0.11	0.09
Social Sciences	0.14	0.10
Business	0.19	0.22
Natural Science / Engineering	0.20	0.16
Health Sciences	0.06	0.17
At University Exit:		
Highest acquired degree:		
Dropout	0.91	0.14
Bachelor	0.04	0.25
Master	0.05	0.61
Bachelor Graduates:		
Switching field of study	0.20	0.26
Excess time-to-graduation (months)	58.86 (35.27)	48.78 (32.13)
Excess time-to-graduation	0.60	0.41
Excess time-to-graduation > 1 year	0.60	0.37
Master Graduates:		
Switching field of study	0.08	0.12
Excess time-to-graduation (months)	36.00 (10.73)	20.01 (14.16)
Excess time-to-graduation	0.92	0.87
Excess time-to-graduation > 1 year	0.85	0.64
After University Exit:		
Hourly wages (real 2000 DKK)	132.37 (57.64)	166.48 (69.38)
Yearly earnings (real 2000 DKK)	134,462 (113,064)	199,033 (133,693)
Fraction of Sample	0.12	0.88

Table 4 *continued*

	Mean (Standard Deviation)	
	Unobserved Type	
Parental Background Characteristics	Type 1	Type 2
Parental Education:		
Mother basic school	0.23	0.16
Mother high school	0.03	0.04
Mother vocational training	0.32	0.31
Mother short higher education	0.04	0.06
Mother medium higher education	0.29	0.32
Mother long higher education	0.08	0.12
Father basic school	0.21	0.15
Father high school	0.04	0.05
Father vocational training	0.33	0.31
Father short higher education	0.02	0.04
Father medium higher education	0.21	0.22
Father long higher education	0.19	0.23
Parental Finances (real 2000 DKK):		
Mother's gross income	207,533 (117,553)	228,905 (140,590)
Mother's wealth	109,848 (514,260)	184,500 (1,106,866)
Mother's assets	172,216 (280,445)	237,433 (426,735)
Mother's liabilities	259,001 (565,443)	373,520 (1,297,834)
Mother's housevalue	149,153 (228,089)	189,019 (415,615)
Father's gross income	347,883 (417,285)	406,950 (489,828)
Father's wealth	32,133 (3,530,629)	640,871 (4,783,449)
Father's assets	427,635 (3,379,785)	1,124,081 (5,147,710)
Father's liabilities	395,501 (418,630)	483,210 (1,113,906)
Father's housevalue	517,030 (407,752)	600,828 (651,696)
Fraction of Sample	0.12	0.88

Notes to Table 4: The table displays behavioral and background differences between individuals of each of the unobserved types. The top panel shows differences in individual choices and outcomes, and the bottom panel shows differences in parental background characteristics. (Standard deviation in parenthesis). For indicator variables the fraction of the sample is reported. The exchange rate on December 31, 2000 was 8.0205 DKK/USD and 7.4631 DKK/Euro.

Table 5: Initial Conditions.

	Marginal Effect on Probability			
			High School	
	Type 1 probability		GPA	Math
Individual initial characteristics:				
High school GPA		-0.07 ***		
		(0.01)		
High school Math level		-0.02 **		
		(0.01)		
Parental Education:				
Mother high school	-0.03	-0.02	-0.07	0.25 **
	(0.02)	(0.02)	(0.06)	(0.06)
Mother vocational training	-0.03	-0.01	0.00	0.09
	(0.03)	(0.03)	(0.09)	(0.10)
Mother short higher education	-0.05	-0.01	0.02	0.30 ***
	(0.02)	(0.02)	(0.07)	(0.07)
Mother medium higher education	-0.03	0.00	0.14 **	0.20 ***
	(0.03)	(0.03)	(0.09)	(0.09)
Mother long higher education	-0.05 *	-0.02	0.27 ***	0.11
	(0.03)	(0.04)	(0.10)	(0.11)
Father high school	-0.03	-0.01	-0.03	0.19 *
	(0.02)	(0.02)	(0.07)	(0.07)
Father vocational training	-0.03	-0.03	-0.10	0.07
	(0.04)	(0.04)	(0.11)	(0.11)
Father short higher education	-0.07 **	-0.05	0.03	0.19 *
	(0.02)	(0.02)	(0.07)	(0.07)
Father medium higher education	-0.04	-0.02	0.02	0.20 ***
	(0.02)	(0.03)	(0.07)	(0.08)
Father long higher education	-0.04 *	-0.01	0.15 **	0.27 ***
	(0.00)	(0.00)	(0.00)	(0.00)
Parental Finances (real 2000 DKK):				
Mother's gross income/1000000	-0.05	-0.03	0.25	0.15
	(0.00)	(0.00)	(0.00)	(0.00)
Mother's wealth/1000000	0.00	0.00	0.02	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)
Father's gross income/1000000	-0.01	-0.01	0.00	0.02
	(0.00)	(0.00)	(0.00)	(0.00)
Father's wealth/1000000	-0.03 *	-0.03 *	0.01	0.05
	(0.00)	(0.00)	(0.00)	(0.00)
Pseudo R^2	0.01	0.05	0.05	0.03
Wald test of joint significance of parental background variables (p-value)	0.06 *	0.81	0.00 ***	0.00 ***

Notes to Table 5: The table displays regression coefficients for determinants of initial conditions represented by type 1 probabilities and observed high school outcomes. (Standard errors in parentheses). ***, **, * indicates coefficient significance at the 1%, 5%, 10% level of significance, respectively.

Table 6: Determinants of Academic Achievement.

	Marginal Effects								
	All University Entrants						**Master Graduates**		
	Drop out			**Master Graduation**			**Excess time-to-graduation (months)**		
Individual initial characteristics:									
Type 1 probability		0.96 ***	0.91 ***		-1.33 ***	-1.32 ***		15.70 ***	14.16 ***
High school GPA			-0.03			0.09 ***			1.87 *
High school Math level			-0.10 ***			0.10 ***			1.06
Initial field of study:									
Life Sciences			-0.09 **			0.44 ***			-3.73 ***
Social Sciences			0.14 ***			0.08			0.31
Business			-0.06 **			0.18 ***			-5.02 ***
Natural Science / Engineering			0.00			0.10 **			-7.48 ***
Health Sciences			0.06			0.27 ***			-7.29 ***
Parental Education:									
Mother high school	-0.03	-0.01	0.00	0.04	0.02	0.01	2.06	2.58	1.93
Mother vocational training	-0.02	-0.01	0.00	0.04	0.02	0.02	-0.05	0.20	-0.58
Mother short higher education	-0.05	-0.01	0.00	0.05	0.01	-0.04	-0.42	-0.01	-0.52
Mother medium higher education	-0.02	0.00	0.01	0.04	0.03	0.01	2.04	2.30	1.43
Mother long higher education	-0.03	0.02	0.03	0.02	-0.02	-0.03	3.57 *	4.04 *	3.08 *
Father high school	-0.06	-0.06	-0.05	-0.02	-0.05	-0.06	1.24	1.11	0.61
Father vocational training	-0.03	-0.01	-0.01	0.03	0.00	0.01	-0.19	0.04	0.04
Father short higher education	-0.07 **	-0.03	-0.03	0.09	0.06	0.07	0.68	1.02	0.72
Father medium higher education	-0.03	-0.01	-0.01	0.02	-0.01	-0.01	0.99	1.14	0.65
Father long higher education	-0.06 **	-0.07 *	-0.06	0.08 *	0.05	0.05	0.78	0.79	-0.05
Parental Finances (real 2000 DKK):									
Mother's gross income/1000000	-0.07	-0.02	0.01	0.09	0.09	0.02	-0.56	-0.29	2.04
Mother's wealth/1000000	-0.01	-0.01	-0.02	0.03	0.03	0.02	0.22	0.20	0.20
Father's gross income/1000000	-0.09 **	-0.03	0.00	0.07 *	0.08 **	0.08 **	-0.69	-0.67	-0.28
Father's wealth/1000000	-0.04 ***	-0.04 ***	-0.04 ***	0.00	-0.01 *	-0.01 *	0.24	0.24	0.23
Pseudo R^2	0.02	0.36	0.38	0.01	0.15	0.22	0.01	0.03	0.12
Log likelihood	-1047	-686	-665	-1365	-1165	-1080	-716	-715	-677
Wald test of joint significance of parental background variables (p-value)	0.00 ***	0.41	0.66	0.08 *	0.33	0.50	0.38	0.45	0.43

Notes to Table 6: The table displays marginal effects of parental background characteristics and the type 1 probability on the probabilities of dropping out and acquiring a Master's degree estimated in a logit model, as well as regression coefficients from a regression of excess time (in months) to Master graduation. The comparison groups are mothers and fathers with no more than elementary schooling, respectively. ***, **, * indicates significance at the 1%, 5%, 10% level of significance, respectively. The last row displays p-values from Wald tests of the joint significance of parental background characteristisc.

Table 7: 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 7: 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 8: 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 8: The table displays observed and predicted choices and transitions over time.

Table 9: Observed and Predicted Choices and Outcomes, by Parental Wealth.

	Mother's Wealth				Father's Wealth			
	< 10th percentile		> 90th percentile		< 10th percentile		> 90th percentile	
	Observed	Predicted	Observed	Predicted	Observed	Predicted	Observed	Predicted
Distribution over alternatives								
$P(d_t^0=1)$	0.42	0.42	0.40	0.40	0.43	0.42	0.40	0.40
$P(d_t^1=1)$	0.23	0.18	0.26	0.22	0.20	0.18	0.25	0.21
$P(d_t^2=1)$	0.21	0.20	0.21	0.21	0.20	0.20	0.21	0.21
$P(d_t^3=1)$	0.09	0.11	0.09	0.10	0.09	0.11	0.08	0.10
$P(d_t^4=1)$	0.06	0.08	0.05	0.07	0.07	0.08	0.06	0.07
Accumulated course credits								
$P(g_t=0)$	0.44	0.45	0.38	0.40	0.45	0.45	0.40	0.41
$P(g_t=1)$	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.02
$P(g_t=2)$	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05
$P(g_t=3)$	0.09	0.09	0.10	0.10	0.09	0.09	0.09	0.10
$P(g_t=4)$	0.03	0.03	0.04	0.03	0.03	0.03	0.03	0.03
$P(g_t=5)$	0.09	0.10	0.11	0.10	0.10	0.10	0.10	0.10
$P(g_t=6)$	0.20	0.18	0.21	0.20	0.18	0.18	0.22	0.20
$P(g_t=7)$	0.08	0.08	0.09	0.09	0.08	0.08	0.09	0.09
g_t	2.70	2.67	3.06	2.93	2.67	2.64	2.99	2.86
Total in last time period:								
G_{10}	24.47	24.10	27.19	26.21	24.19	24.06	26.86	25.95
Wages after University Exit								
W_t	162.40	149.08	160.32	146.32	155.87	144.26	159.44	147.30

Notes to Table 9: The table displays observed and predicted choices and outcomes. Observations and predictions are displayed separately by students with parental wealth below the 10^{th} and above the 90^{th} percentile.

Table 10: Decsriptive Statistics, by Field of University Education.

	Mean (Standard Deviation)					
	Field of University Education					
Individual Characteristics	**Humanities / Art / Education**	**Life Sciences**	**Social Science**	**Business**	**Natural Science / Engineering**	**Health Sciences**
At University Entry:						
High school GPA	8,94	8,88	9,23	8,49	8,95	9,26
	(0,77)	(0,81)	(0,74)	(0,83)	(0,81)	(0,60)
High school Math level	1,55	2,67	2,13	1,80	2,73	2,70
	(0,79)	(0,61)	(0,87)	(0,94)	(0,54)	(0,54)
Female	0.69	0.43	0.49	0.36	0.24	0.60
Accumulated work experience	0,60	0,53	0,66	0,70	0,47	0,65
	(0,60)	(0,56)	(0,58)	(0,66)	(0,55)	(0,57)
Age	21,32	21,23	21,18	21,10	20,93	21,13
	(0,76)	(0,73)	(0,78)	(0,74)	(0,77)	(0,77)
During University Enrolment:						
Accumulated course credits per year	3,99	3,97	4,30	4,08	4,26	4,88
	(2,44)	(2,42)	(2,24)	(2,41)	(2,44)	(1,94)
Accumulated work experience per year (years)	0,16	0,13	0,18	0,27	0,16	0,14
	(0,22)	(0,22)	(0,22)	(0,32)	(0,25)	(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	162,81	176,17	178,29	183,71	206,23
	(59,31)	(46,82)	(70,93)	(79,83)	(57,01)	(87,91)
Yearly earnings (real 2000 DKK)	159 786	189 898	219 134	237 730	221 700	266 481
	(111 764)	(123 479)	(138 667)	(130 085)	(131 816)	(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 10: 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 11: Parameter Estimates, by Field of University Education.

	Humanities / Art / Education	Life Sciences	Social Science
		w_t	
α_{01}	4.57 (0.03) ***	4.42 (0.06) ***	4.49 (0.05) ***
α_{02}	4.60 (0.03) ***	4.96 (0.07) ***	4.60 (0.05) ***
α_1 (Bachelor degree)	0.08 (0.02) ***	0.12 (0.08)	0.11 (0.04) ***
α_2 (Master degree)	0.27 (0.03) ***	0.11 (0.06) *	0.30 (0.05) ***
α_3 (Experience)	0.08 (0.01) ***	0.12 (0.04) ***	0.09 (0.03) ***
α_4 (Experience^2)	0.00 (0.00)	-0.01 (0.01)	0.00 (0.00)
α_5 (Part time work)	-0.13 (0.02) ***	-0.63 (0.06) ***	-0.17 (0.04) ***
		g^*	
γ_{02}	2.67 (0.13)	3.16 (0.27) ***	2.68 (0.25) ***
γ_1 (High school GPA≥9)	0.16 (0.06) ***	0.05 (0.10)	0.13 (0.09)
γ_2 (High level Math)	0.17 (0.07)	0.37 (0.11) ***	0.06 (0.10)
γ_3 (Bachelor degree, $\mathbf{1}[E_t=1]$)	0.04 (0.07)	0.21 (0.24)	0.37 (0.12) ***
γ_4 (Accumulated course credits, G_t)	0.05 (0.00) ***	0.08 (0.01) ***	0.03 (0.01) ***
γ_5 (Time since enrolment)	-0.52 (0.02) ***	-0.45 (0.05) ***	-0.41 (0.04) ***
γ_6 (Student employment, $h_t \geq ¼$)	0.25 (0.07) ***	0.15 (0.12)	0.23 (0.11) **
γ_7 (Student employment, $h_t \geq ½$)	-0.41 (0.09) ***	0.25 (0.16)	-0.27 (0.17)
γ_8 (Student employment, $h_t \geq ¾$)	-1.19 (0.11) ***	-3.64 (0.22) ***	-1.08 (0.20) ***
		$P(d_t^k=1)$	
s (Full-time job finding cost)	-2.90 (0.12) ***	-3.21 (0.31) ***	-2.81 (0.22) ***
β_{01}	-3.48 (0.16) ***	-4.12 (0.43) ***	-2.99 (0.33) ***
β_{02}	-1.11 (0.13)	0.24 (0.39)	-0.42 (0.29)
β_1 (High school GPA≥9)	0.40 (0.10) **	0.55 (0.22) **	0.41 (0.19) **
β_2 (High level Math)	0.44 (0.12) ***	0.31 (0.23)	0.70 (0.20) ***
β_3 (Time since enrolment)	-0.32 (0.02) ***	-0.65 (0.06) ***	-0.48 (0.04) ***
$\beta_1+\beta_1^2$ (High school GPA≥9)	0.10 (0.11)	0.78 (0.22) ***	-0.02 (0.19)
$\beta_2+\beta_2^2$ (High level Math)	0.22 (0.13)	0.17 (0.23)	0.06 (0.20)
$\beta_3+\beta_3^2$ (Time since enrolment)	-0.49 (0.02) ***	-0.87 (0.06) ***	-0.55 (0.05) ***
β_4^2 (Student empl. previous period, $h_{t-1}=¼$)	1.52 (0.09) ***	1.97 (0.16) ***	1.47 (0.15) ***
β_5^2 (Student empl. previous period, $h_{t-1}=½$)	1.29 (0.14) ***	2.12 (0.26) ***	1.68 (0.25) ***
β_6^2 (Student empl. previous period, $h_{t-1}=¾$)	1.55 (0.21) ***	2.06 (0.52) ***	1.04 (0.37) ***
$\beta_1+\beta_1^3$ (High school GPA≥9)	-0.44 (0.12) ***	0.14 (0.24)	-0.88 (0.24) ***
$\beta_2+\beta_2^3$ (High level Math)	-0.12 (0.15) **	-0.50 (0.25) **	-0.55 (0.23) **
$\beta_3+\beta_3^3$ (Time since enrolment)	-0.52 (0.03) ***	-0.80 (0.07) ***	-0.60 (0.05) ***
β_4^3 (Student empl. previous period, $h_{t-1}=¼$)	0.82 (0.11) ***	0.99 (0.21) ***	1.18 (0.21) ***
β_5^3 (Student empl. previous period, $h_{t-1}=½$)	1.70 (0.15) ***	2.53 (0.29) ***	2.11 (0.31) ***
β_6^3 (Student empl. previous period, $h_{t-1}=¾$)	1.83 (0.22) ***	2.41 (0.55) ***	1.54 (0.46) ***
$\beta_1+\beta_1^4$ (High school GPA≥9)	-0.63 (0.14)	-0.06 (0.37)	-0.34 (0.27)
$\beta_2+\beta_2^4$ (High level Math)	0.08 (0.17) *	-0.43 (0.38)	-0.45 (0.27) *
$\beta_3+\beta_3^4$ (Time since enrolment)	-0.49 (0.03) ***	-0.83 (0.08) ***	-0.63 (0.06) ***
β_4^4 (Student empl. previous period, $h_{t-1}=¼$)	0.59 (0.18) **	0.57 (0.48)	0.69 (0.30) **
β_5^4 (Student empl. previous period, $h_{t-1}=½$)	2.10 (0.18) ***	2.48 (0.47) ***	2.27 (0.35) ***
β_6^4 (Student empl. previous period, $h_{t-1}=¾$)	3.44 (0.19) ***	4.04 (0.54) ***	3.18 (0.33) ***
β_7 (Study grant)	-2.57 (0.23) ***	-4.43 (0.74) ***	-1.92 (0.46) ***
π_1	0.17	0.19	0.10
π_2	0.83	0.81	0.90
Log Likelihood	-20511	-7303	-8400

Table 11 *continued*

	Business	Natural Science / Engineering	Health Sciences
		w_t	
α_{01}	4.59 (0.03) ***	4.43 (0.04) ***	4.54 (0.06) ***
α_{02}	4.64 (0.03) ***	4.65 (0.02) ***	4.80 (0.03) ***
α_1 (Bachelor degree)	0.15 (0.02) ***	0.22 (0.02) ***	0.01 (0.04)
α_2 (Master degree)	0.27 (0.03) ***	0.33 (0.03) ***	0.23 (0.03) ***
α_3 (Experience)	0.08 (0.02) ***	0.07 (0.01) ***	0.08 (0.02) ***
α_4 (Experience^2)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
α_5 (Part time work)	-0.17 (0.03) ***	-0.20 (0.02) ***	-0.27 (0.03) ***
		g^*	
γ_{02}	3.39 (0.16) ***	2.36 (0.26) ***	2.10 (0.29) ***
γ_1 (High school GPA≥9)	0.09 (0.07)	0.12 (0.08)	0.22 (0.08) ***
γ_2 (High level Math)	0.01 (0.06)	0.41 (0.08) ***	0.34 (0.09) ***
γ_3 (Bachelor degree, $\mathbf{1}[E_t=1]$)	0.68 (0.08) ***	0.22 (0.09) **	0.57 (0.15) ***
γ_4 (Accumulated course credits, G_t)	0.01 (0.01)	0.07 (0.01) ***	0.09 (0.01) ***
γ_5 (Time since enrolment)	-0.52 (0.03) ***	-0.65 (0.03) ***	-6.54 (0.31) ***
γ_6 (Student employment, $h_t \geq 1/4$)	0.32 (0.08) ***	0.31 (0.09) ***	0.24 (0.09) ***
γ_7 (Student employment, $h_t \geq 1/2$)	0.20 (0.10) **	-0.21 (0.11) *	-0.26 (0.12) **
γ_8 (Student employment, $h_t \geq 3/4$)	-1.19 (0.11) ***	-1.75 (0.12) ***	-1.87 (0.14) ***
		$P(d_t^k=1)$	
s (Full-time job finding cost)	-2.93 (0.16) ***	-2.51 (0.14) ***	-2.78 (0.15) ***
β_{01}	-3.20 (0.21) ***	-3.30 (0.30) ***	-3.26 (0.40) ***
β_{02}	-0.26 (0.20)	-1.25 (0.17) ***	-1.23 (0.20) ***
β_1 (High school GPA≥9)	0.48 (0.13) **	0.38 (0.14) ***	0.57 (0.13) ***
β_2 (High level Math)	0.44 (0.12) ***	0.66 (0.14) ***	0.92 (0.15) ***
β_3 (Time since enrolment)	-0.53 (0.03) ***	-0.39 (0.03) ***	-0.49 (0.03) ***
$\beta_1+\beta_1^2$ (High school GPA≥9)	0.32 (0.13) **	-0.09 (0.14)	0.41 (0.14) ***
$\beta_2+\beta_2^2$ (High level Math)	0.20 (0.13)	0.27 (0.15) *	0.31 (0.16) *
$\beta_3+\beta_3^2$ (Time since enrolment)	-0.67 (0.03) ***	-0.49 (0.03) ***	-0.63 (0.04) ***
β_4^2 (Student empl. previous period, $h_{t-1}=1/4$)	1.35 (0.10) ***	1.58 (0.13) ***	1.50 (0.12) ***
β_5^2 (Student empl. previous period, $h_{t-1}=1/2$)	1.20 (0.15) ***	1.46 (0.18) ***	1.44 (0.21) ***
β_6^2 (Student empl. previous period, $h_{t-1}=3/4$)	1.48 (0.21) ***	0.17 (0.30) ***	1.29 (0.29) ***
$\beta_1+\beta_1^3$ (High school GPA≥9)	-0.30 (0.14) **	-0.64 (0.15) ***	-0.23 (0.16)
$\beta_2+\beta_2^3$ (High level Math)	-0.35 (0.14) **	-0.31 (0.16) *	-0.62 (0.17) ***
$\beta_3+\beta_3^3$ (Time since enrolment)	-0.66 (0.04) ***	-0.46 (0.03) ***	-0.59 (0.04) ***
β_4^3 (Student empl. previous period, $h_{t-1}=1/4$)	1.03 (0.13) ***	1.19 (0.15) ***	1.28 (0.17) ***
β_5^3 (Student empl. previous period, $h_{t-1}=1/2$)	2.15 (0.15) ***	2.13 (0.18) ***	2.69 (0.21) ***
β_6^3 (Student empl. previous period, $h_{t-1}=3/4$)	1.81 (0.24) ***	1.12 (0.25) ***	1.64 (0.35) ***
$\beta_1+\beta_1^4$ (High school GPA≥9)	-0.19 (0.16)	-0.58 (0.14) ***	-0.24 (0.18)
$\beta_2+\beta_2^4$ (High level Math)	-0.22 (0.16)	-0.32 (0.15) **	-0.60 (0.19) ***
$\beta_3+\beta_3^4$ (Time since enrolment)	-0.64 (0.04) ***	-0.37 (0.03) ***	-0.43 (0.04) ***
β_4^4 (Student empl. previous period, $h_{t-1}=1/4$)	0.45 (0.18) **	0.36 (0.19) *	0.64 (0.24) ***
β_5^4 (Student empl. previous period, $h_{t-1}=1/2$)	1.85 (0.19) ***	1.92 (0.18) ***	2.40 (0.24) ***
β_6^4 (Student empl. previous period, $h_{t-1}=3/4$)	3.30 (0.20) ***	3.10 (0.16) ***	3.86 (0.22) ***
β_7 (Study grant)	-1.90 (0.26) ***	-0.78 (0.24) ***	-2.80 (0.34) ***
π_1	0.14	0.04	0.03
π_2	0.86	0.96	0.97
Log Likelihood	-17476	-11426	-5713

Notes to Table 11: 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.

Table 12: Simulated Effects on Academic Outcomes.

		Tilting Grant			Merit Aid			Backloading	
Academic outcome:	**Benchmark**	**Sim1**	**Sim2**	**Sim3**	**Sim4**	**Sim5**	**Sim6**	**Sim7**	**Sim8**
Dropout rate	0.23	0.01	-0.01	-0.01	-0.02	-0.03	-0.03	-0.01	-0.01
Graduating with MSc	0.55	0.02	0.05	0.04	0.03	0.04	0.05	0.03	0.03
Excess t-t-g, MSc	0.87	0.00	0.00	0.01	-0.05	-0.06	-0.08	-0.05	-0.05
>1y excess t-t-g, MSc	0.65	0.00	0.00	0.00	-0.07	-0.07	-0.09	-0.06	-0.06
Study Grant over Choice:									
k=1	47000	60000	70000	100000	38000	40000	41000	39200	39800
k=2	47000	40000	35000	0	45000	47000	50000	43400	44700
k=3	35000	20000	0	0	38500	40500	41000	39300	40000
k=4	0	0	0	0	14500	12000	9000	13800	13600
Average Grant per Student	23412	23612	23107	22700	22184	22303	23192	22152	22605
Merit Aid:									
$g_t > 3$	$g_t > 0$				+				
$g_t > 4$	$g_t > 0$					+			
$g_t > 5$	$g_t > 0$						+		
Backloading:									
$t = 5$	-							x2	
$t = 5$	-								x3

Notes to Table 12: The table presents the simulated effects on four academic outcomes of changing different aspects of the educational environment. Simulations 1-3 represent the effects of gradually tilting the study grant system towards students who work fewer hours. Simulations 4-6 represent the effects of gradually increasing the minimum course credit grant eligibility requirement. Simulations 7-8 represent the effects of increased backloading. For each simulation, all other aspects of the educational environment are held at their benchmark values presented in the first column of the table.

Figure 1: Study Grant as a function of Student Earnings.

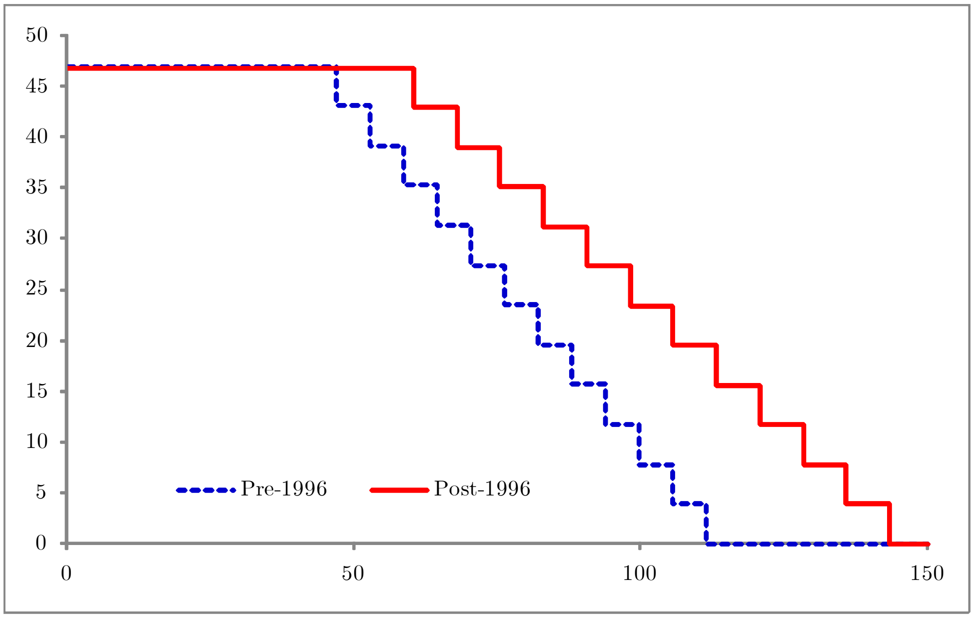
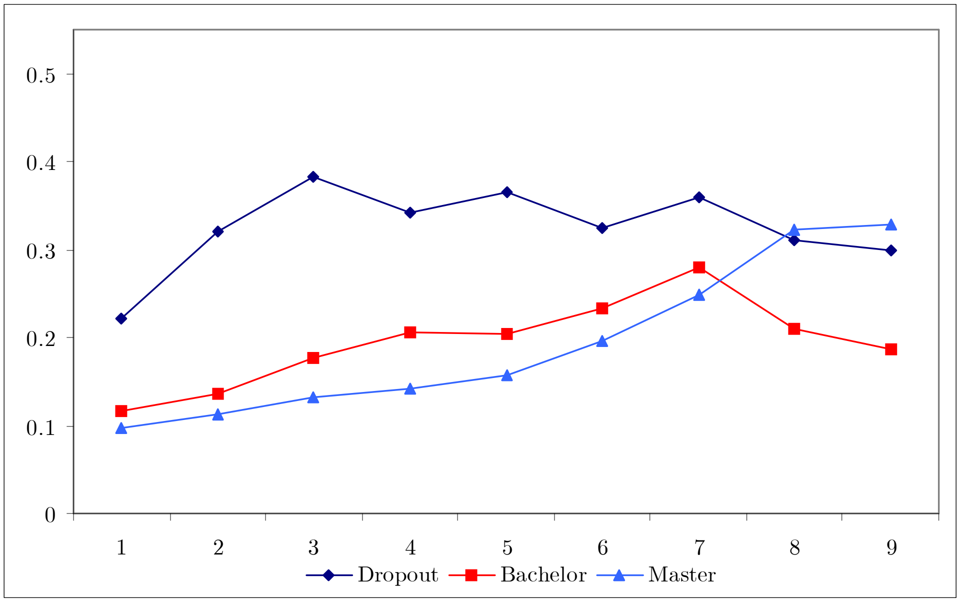


Figure 2:

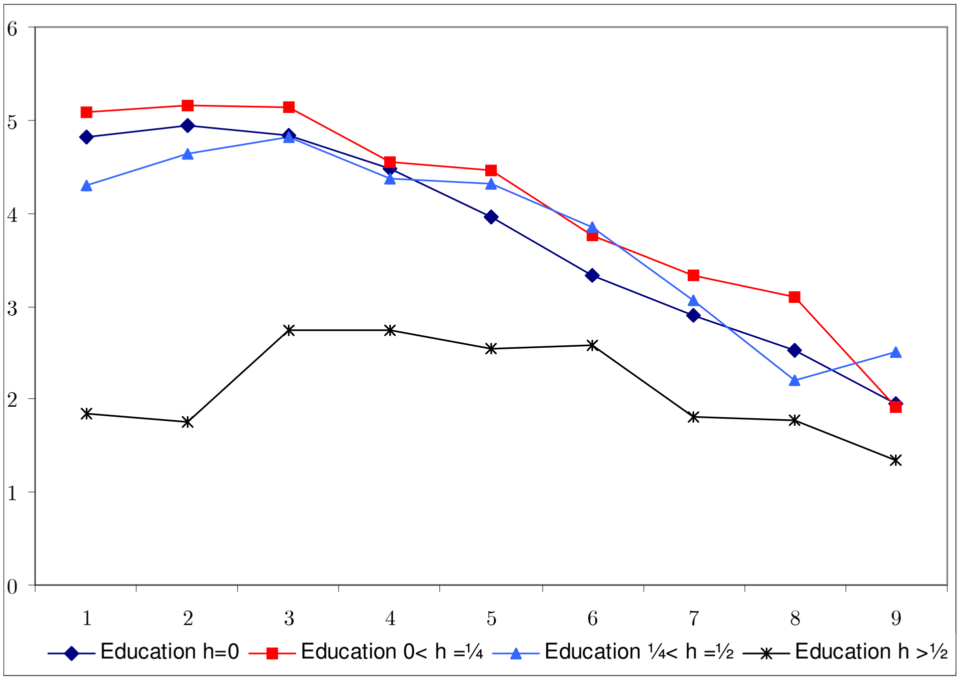
Notes to Figure 1: The figure displays the annual Study Grant as a function of annual Earnings. This dependence of received study grant on student earnings is showed both before and after the increase of the threshold for maximum allowed earnings in 1996. The amounts from 1995 and 1996, respectively, are taken as representative for pre and post threshold change study grant rules. The displayed amounts are given in real 2000 DKK for a university student above 18 years old (100% of students), living away from parents (99% of students), not having children (83% of students), and not having any severe disabilities. The exchange rate on December 31, 2000 was 8.0205 DKK/USD and 7.4631 DKK/Euro.

Figure 3: Annual Student Employment Experience.



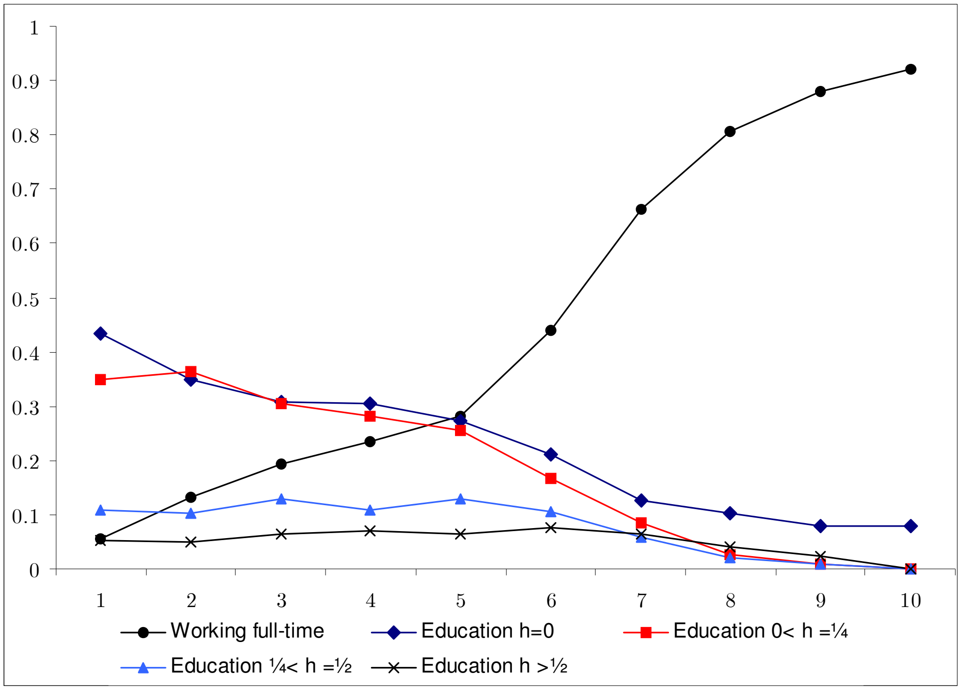
Notes to Figure 2: The figure displays accumulated labor market experience (in years) in the year over time after university entry for full-time university students. The figure displays separate student employment experience profiles for university dropouts and individuals graduating from the university with Bachelor's and Master's degrees, respectively.

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



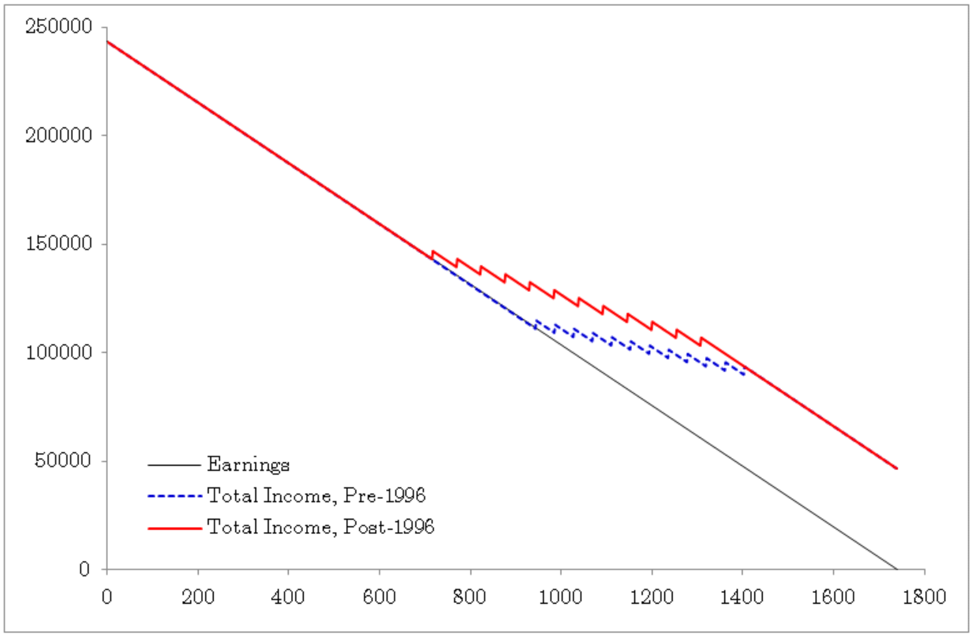
Notes to Figure 3: The figure displays the average accumulated course credits per year, g_t, for university attendants over time since initial enrolment. The amount of course credits is diplayed separately by amount of labor market work in the year, $k \in \{1, 2, 3, 4\}$. A total of six course credits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

Figure 5: Transition from University Education to Labor Market Work.



Notes to Figure 4: 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 6: Student Income and Working Hours.



Notes to Figure 5: The figure displays the annual Student Income as a function of annual Hours Worked. The thin solid black line shows student earnings, i.e. total student income without any study grant. The dashed blue line shows total student income according to the study grant regime pre-1996, while the thick red line displays total student income according to the study grant regime post-1996. The amounts from 1995 and 1996, respectively, are taken as representative for pre and post threshold change study grant rules. The amounts on the vertical axis are given in real 2000 DKK for a university student above 18 years old (100% of students), living away from parents (99% of students), not having children (83% of students), and not having any severe disabilities. The hourly wage rate is assumed to be 140 DKK. The exchange rate on December 31, 2000 was 8.0205 DKK/USD and 7.4631 DKK/Euro. The horizontal axis displays working hours such that 0 corresponds to working full-time year-round. It is assumed that a full-time year-round job requires 1739 annual working hours, i.e. 37 working hours per week multiplied by 47 working weeks per year.