The effect of student aid reform on graduation times. A duration analysis

Iida Häkkinen and Roope Uusitalo

February 25, 2002
Preliminary draft

Abstract

In this paper we estimate the effect of the change from a loan-based to a grant-based student financial aid system on the graduation times. Starting from simple assumptions on how the explanatory variables affect study intensity, we derive the duration distribution for the times-to-degree that turns out to follow a two parameter Gamma distribution. We then show that the effect of the time-varying covariates can be modelled as in an accelerated failure time model, but that the covariates have a cumulative impact on graduation times. Our empirical results indicate that the student aid reform slightly shortened the graduation times. Other results show that married students graduate faster and female students with children considerably slower. Also university resources and labor market conditions have a significant effect on graduation times. After controlling for various covariates, there are still substantial differences across different fields of study.

1 Motivation

The average graduation age of tertiary level graduates in Finland is the second highest among the OECD countries. On average, the Finnish university students complete their Master's degree at the age of 27, after studying 6.5 years at the university (Figure 1).

Long study times have been considered to be a problem for two main reasons. First, long duration of studies increases the number of enrolled students in the universities. Hence, there are less resources available for each student. Libraries and computing facilities get overcrowded and the workload of thesis advisors grows.

The second argument that lies even more in the center of the current debate is that long graduation times postpone the transition to the labor market. In addition to the private costs for the student, this reduces labor supply and decreases tax revenue. As in other European countries, the Finnish population is aging rapidly. Early retirement further reduces labor supply of the older cohorts, so that not only the fraction of population in working age but also the employment ratio of the working age population has been on a decline. Numerous attempts to raise the employment ratio of the older cohorts have produced only modest improvements. It is quite possible that it would be easier to increase the labor supply of the younger cohorts. Reducing investments in education by reducing the fraction of the age cohort entering universities would be bad economic policy. In contrast, making university studies more effective by shortening graduation times could bring considerable gains to the economy.

Shorter graduation times have been on the government agenda for a long time. When the student support system was first created in 1969, the main goal of the government was to enable full-time studies for the university students and reduce student employment in order to shorten the average graduation times and reduce dropping out (Blomster, 2000). The goal has not changed much in 30 years. The program of the current government states that "The prerequisites for the basic and the postgraduate degree courses will be improved in order to reduce the amount of time required to complete these courses" and that "Student grants will be developed so that they better support graduation". More specific goals are stated in the government action plan for higher education where the target is set to 75 percent of students completing their Master's degree in five years.

A number of possible factors may explain the long duration of the university studies. Some of the most common explanations have to do with inefficient organization of courses and examinations, insufficient support at the thesis phase, lack of resources (e.g. libraries), additional courses compared to minimum graduation requirements, good employment prospects (no need to graduate) and poor employment prospects (no use to graduate). In addition, a number of factors related to the personal or family situation of individual students matter for graduation times (Raivola et. al 2000).

The government has made an attempt to reduce the graduation times by creating incentives to the universities and by reforming the student loan system. The universities now receive a part of their funding based on their results i.e. the number of completed degrees. However, probably the most significant change occurred in 1992, when the student financial aid system was profoundly reformed. The old loan-based system was replaced by a grant-based system. At the same time, the maximum duration of student aid was cut from seven years to 55 months.

In this study, we examine how the changes in the student financial aid, university funding, and the changes in the labour market - especially in the unemployment rates - have affected the duration of university studies. We start by giving a short description of the Finnish university education and the student aid system in Chapter 2. That is followed by a short survey of the previous studies in Chapter 3. In Chapter 4 we derive a duration model that can be used to model graduation times. Chapter 5 describes the data sets, Chapter 6 presents the results, and Chapter 7 concludes.

2 Main features of the Finnish university system

The Finnish education system can be divided into three stages: basic education (compulsory, 9 years), upper secondary education (2-3 years) and higher education (at least 3 years). The Finnish education system chart is shown in Figure 2. A three-year vocational qualification or graduation from senior secondary school with a matriculation examination yields eligibility for all forms of higher education. Higher education is given in universities and polytechnics. Universities are characterized by scientific research and the highest education is based thereon. Polytechnics are more oriented towards working life and the requirements are largely set by the working life. The polytechnic system was built during the 1990s and it has expanded in the recent years. The system is still developing. In this study, we focus only on universities and the university students.

There are twenty universities in Finland: ten multidisciplinary universities, three universities of technology, three schools of economics and business administration, and four art academies. The network of universities covers the different parts of the country. All universities are operated by the state. Universities select their own students independently. Competition for student places in higher education is fierce, and an annual intake quota applies to all fields of study. Various entrance examinations form a central part of the selection process. Competition for student places can be largely attributed to the expansion of general senior secondary education, preparing for studies at the tertiary level. Also, the extension of eligibility for higher education to those with an initial vocational qualification has increased competition for student places. The total intake of the universities guarantees a student place for about a third of the relevant age group. The annual number of university applicants is 66,000, of whom 23,000 are admitted. (Havén 1999). Because of highly competitive selection process, it may take a few years for a senior secondary school graduate to be accepted in the desired field.

The number of university students has increased by some 40 % during the last ten years. In total, there are about 147,000 students in the universities, of whom 19,000 are new entrants. Slightly more than half of the university students are female (52 % in 1997-98). The largest fields of study are engineering, humanities, and natural sciences. (Havén 1999).

Most of the students selected for the universities are admitted for courses leading to a higher academic degree (equivalent to the US Master's degree). The Master's degree, which consists of 160-180 credits, is designed to take five years to complete. In most fields of study students are not tied to a predetermined schedule, but they are able to choose the pace of their studies and the minor subjects. Thus, the average duration for completing a Master's degree is about 6.5 years. Students may also work for a lower academic degree (Bachelor's degree). The scope of the Bachelor's degree is 120 credits and it can usually be completed in three years. After the lower degree student can continue for a higher degree or move to the labour market.

2.1 The 1992 student financial aid reform

The student financial aid system dates back to 1969. At first the system was mainly based on student loans with a state subsidized interest rates. In the late 1970's the loan-based student aid system faced a crisis as banks reduced the number of loans granted claiming that the loans were unprofitable. The state was forced to compensate the cuts in student loans by upgrading the direct financial support, the study grant. The share of direct aid by the state grew annually until the year 1992, but still the student loans were the cornerstone of the system. Between 1985 and 1992, the average maximum study grant was 70 euro per month, the maximum housing supplement was 97 euro per month, and the maximum student loan was 262 euro per month. The total student aid in 1985-1992 ranged from 351 euro to 542 euro per month (nominal prices). (Blomster 2000). Student financial aid was granted for up to 7 academic years. The aid was generally not means tested but there were income limits for student earnings in order to ensure that the aid was used to finance full-time studies. Student was allowed to earn 336 - 504 euro per month before the student aid was cut. However, many students worked during their studies to avoid incurring of a debt.

In summer 1992, the financial aid system was profoundly reformed. The reform was made because the interest rates had risen rapidly and the banks once again saw study loans as unprofitable. In addition, the system with interest subsidized loans became expensive for the state as the interest rate rose. Government subsidy for interest rate was ended, but government still provided a guarantee for a loan up to 202 euro. The amount of study grant was more than doubled to 264 euro month. At the same time the study grant became subject to income tax. (Blomster 2000). However, students under 20 years were granted only half of the study grant and students living with their parents received even less. In 1997 also the 19-year-old students not living with their parents were granted the full study grant and in 1998 this was extended to include the 18-year-olds. (Raivola et al. 2000).

In addition to the study grant and the loan, single students living on their own are eligible to a housing supplement. The maximum amount of housing supplement was at first 149 euro per month, but after the first year it was increased to 161 euro per month. The amount one receives depends on the rent. Students living with their parents are not eligible to any housing supplement, and students living with their spouses or cohabiting can apply for a general housing allowance.

In autumn 1995 the study grant was cut by two per cent to 259 euro per month and it has been on the same level since that. Also the maximum amount of housing supplement was reduced to 144 euro per month. However, at the same time the maximum amount of student loan was increased to 219 euro per month. In 1995-1999 the maximum amount of student aid was 621 euro per month. (KELA 2001).

In the 1992 reform also the period of eligibility for the student financial aid was cut from seven years to 55 months[1]. If a student gets her degree in less than the time for which financial aid is available, she can use the months left over for post-graduate or other academic studies. Students, who received financial aid before July 1992, can have student financial aid for up to 7 years - however they are not granted for more than 70 months of the "new", post-1992, student aid.

Student financial aid is available for full-time studies. The full-time requirement is considered to be met if the purpose of the studies is to obtain an academic degree. However, there are limitations on how much a student can earn before the student aid is cut. In 1992-1995 the maximum monthly income that did not reduce the student aid was 285 euro. In 1995-1997 the income limit of study grant was raised by 17 euro. In aid-free months, in the summer for example, there was no income limit. In 1998 the income limit system changed totally. In stead of checking monthly incomes, all taxable income during the calendar year is now taken into account. For each month during which student receives study grant or housing supplement, the exempt income amount is 505 euro, and for each aid-free month 1515 euro. Assuming that a student receives aid for nine months (the length of academic year), she is allowed to have up to 9090 euro in other income during the calendar year. The timing of the income doesn't matter, but the annual exempt income must not be exceeded. The development of the Finnish student financial aid system is described in Figure 3. The inflation has naturally affected the real value of student financial aid. Thus in Figure 3, real values in 2000 euro prices are presented.

The main aim of the student aid reform was to make studying more effective by distributing more financial aid at a faster rate. However, the system has not functioned as planned. The changes in the loan system made loans unpopular. Before 1992 roughly half of the students took out the loan, after the reform only less than a third (Raivola et al. 2000, 100). Rather than taking out the loan, students have preferred to add to their income by working. According to a recent study on University of Helsinki, 75 % of students worked during the semester 2000-2001 (Härkönen 2001). Most of the students are working part-time.

3 Previous research on schooling duration

The duration of education is commonly measured as the number of years of schooling obtained. Schooling attainment has been subject to intensive theoretical and empirical research usually motivated by the human capital theory, and optimal investment on education. Some of the early work formalizing the ideas include Ben-Porath (1967) and Mincer (1974). In the optimal schooling choice

[1] In the case of certain degrees or studies at a foreign university, the aid period can be longer. For specific reasons, the aid period can also be extended on application by a maximum of nine months.

model the individuals continue their education until the marginal rate of return on additional investment falls below the marginal cost. The optimal schooling choice model has also served as a way of controlling endogeneity of schooling choices in the vast literature that tries to estimate the return to education. Card (1999) surveys much of that literature in his chapter in the Handbook of Labor Economics. However, it is hard to argue that the students who spend six years obtaining their Bachelor's degree obtain more human capital than the students who graduate in four years. Although longer study duration may involve substantial private and social costs, there are very few studies that try to model the time students spend obtaining their degrees. The few examples surveyed here include van Ophem and Jonker (1999), Löfgren and Ohlsson (1999), Siegfried and Stock (2001) and Ehrenberg and Mavros (1995).

van Ophem and Jonker (1999) model the duration of four-year college education in the U.S. They draw a sample of 721 individuals who started college between 1977 and 1988 from the National Longitudinal Survey of Labor Market Experiences. They assume that at each point in time the student calculates the optimal additional amount of schooling. If the optimal amount of schooling is less than the time needed to graduate the student drops out. For the students who graduate the optimal schooling duration exceeds the time needed to graduate. Using the observed durations of graduation and drop-out times they then develop a dependent competing risks model. Their empirical estimates indicate that there is considerable variation in graduation times. Afro-American and Hispanic students need more time to graduate. Younger students graduate faster. Also higher intellectual ability measured by the AFQT-test scores and recipiency of a scholarship shorten graduation times. They also find that the more time the student spends on the labor market the longer are the graduation times, but note that hours of work is an endogenous variable and that modeling the effect of working while in college on graduation times would require a model for the work decision also.

Löfgren and Ohlsson (1999) study when the undergraduates complete their thesis required for the Bachelor's and the Master's degree. They use data from 181 students who registered on undergraduate thesis courses in economics at Uppsala and Umeå universities during 1993. They show that only half of the students complete their thesis within the expected time. One out of five students have not finished seven semesters after they first register for the thesis course. They model the hazard of completing the thesis with a logit model. Due to small number of observations, the authors find few significant effects, main result is that coauthoring increases the probability on completing. The results also show that there is negative duration dependence, failing to complete a thesis in first two years, reduces the hazard of completing. Interestingly, there seems to be no significant gender effects nor any effect of success in prior economics courses.

Siegfried and Stock (2001) analyze the elapsed time to degree of the economics PhD graduates in US universities. They use data from a survey of 455 individuals who earned a PhD in economics during the 1996-97 academic year. They estimate a Weibull-model and explain the effect of individual characteristics on the elapsed time until the completion of PhD. They find that students who received fellowship support, student who had a prior Master's degree and foreign students graduated faster. Presence of children and years employed prior to PhD increased graduation times.[2]

Also Ehrenberg and Mavros (1995) study duration of Ph.D. programs. They have data on all graduate students who entered Ph.D programs in Economics, English, Mathematics or Physics in Cornell University during the 1962-86 period. So in contrast to Siegfried and Stock they sample from the inflow of students and can, therefore, also analyze the drop-outs. They formulate a competing risks model of graduation and drop-out behavior, where each exit probability follows a proportional hazard model with a piecewise constant baseline hazard. Their results show that students with a prior Master's degree are less likely to drop out, gender does not matter and that foreign students graduate faster. They also find that students' ability measured with GRE scores does not affect completion nor drop-out hazards. Also they find mostly no effect of labor market conditions (academic salaries or proportion of graduates seeking academic employment). Their main interest is in the effect of financial support which they model as a fraction of time a student has received one of the four possible forms of support (teaching or research assistantship, fellowships or tuition waivers). They find that students receiving fellowships or research assistantships have higher completion rates and shorter times-to-degree than students who receive teaching assistantships or tuition waivers or who are totally self-supporting. A major finding is that financial aid patterns have much higher impact on completion rates than mean durations.

4 How to model the duration of higher education

The standard duration model is specified in terms of hazard functions. The hazard function gives the conditional probability that spells that have lasted until time t end at time t,

$$\lambda(t) = \frac{f(t)}{S(t)}$$

where $\lambda(t)$ is the hazard, $f(t)$ the density, and $S(t) = 1 - F(t)$ the survival function. If the hazard of ending the spell is constant, the expected duration of the spell is then $E(T) = 1/\lambda$.

The hazard function approach offers a number of advantages over modeling the expected durations directly. In particular, the "right censored observations" i.e. the spells that have not ended by the end of the observation period can be handled in a straightforward manner, and time-variant covariates can be introduced without conceptual problems. A simple way of adding explanatory

[2] As the sampling in the study is from the outflow of students (graduation, instead of entering the program) the study risks biasing the coefficients and must ignore the drop-outs. However, in time homogenous environment (the number of PhD's has been relatively constant) with to time variant covariates this should not be a major problem.

variables in the model is the proportional hazard specification

$$\lambda(t, x, \beta, \lambda_0) = \phi(x, \beta)\lambda_0(t)$$

where the baseline hazard $\lambda_0(t)$ is multiplied by a factor ϕ that depends on explanatory variables x and parameters β but not on duration t. The proportional hazard model can be easily estimated with maximum likelihood, since the log likelihood function can be written as

$$\ln L = \sum d_i \ln \lambda(t, x, \beta, \lambda_0) + \sum \ln S(t, x, \beta, \lambda_0)$$

where d_i indicates uncensored observations and rest of the notation is as above. With simplest parametric assumptions on the form of the baseline hazard function estimation is easy. For example, with constant baseline hazard the log-likelihood function is globally concave and numerical maximization straightforward[3]. For more general baseline hazard functions, the Cox partial likelihood approach can be used to estimate the coefficients of the explanatory variables without specifying the baseline hazard. (Kiefer 1988, Lancaster 1990)

The hazard approach is a natural way of modeling events that can occur at any point in time. Typical economic applications include the duration of unemployment, the duration of strikes, time to the adoption of new technologies, and duration of marriages. A common feature of these applications is that a large enough change in an explanatory variable may increase the probability of ending the spell instantly. For example, the exhaustion of the unemployment benefits is shown to increase the hazard of exiting unemployment (e.g. Mayer 1990, Katz and Mayer 1990). In his much cited study, Kennan (1985) shows that an unexpected increase in industrial production increases the cost of strikes and increases the probability of ending the strike.

Modeling the effect of exogenous variables on the duration of higher education is more complicated since the effects are likely to involve considerable lags. For example, improving incentives provided by the student aid system may induce a first year student to study more intensively, but this does not show in graduation rates until several years later. Similarly, more lucrative work opportunities at the boom of a business cycle may draw students to the labor market and delay graduation. The effect on graduation times cannot be seen until the student actually graduate, which could be already during an economic downturn. However, the intensity of studies during the first years at university influences the graduation hazards later. The key conclusion appears to be that the relevant explanatory variables include the entire time path of these variables, not only their current values.

Including the entire time path of several covariates creates a high dimensional estimation problem[4]. To reduce dimensionality without losing the essential

[3] When the number of possible exit points is limited as in annual data, a flexible parametrization involves just the full set of year dummies and a constant hazard.

[4] In addition these covariates are endogenous since the time path of a covariate $X(t)$ depends on whether the spell has lasted until the period t. However, Lancaster (1990) shows that though density and survival functions cannot be properly defined with such endogenous covariates, one can still estimate the parameters of the hazard fuction.

feature that the entire path of covariates during the time at the university may matter for graduation times we first model the instantaneous effect of covariates $x(t)$ on study intensity $\eta(t)$.

$$\eta(t) = \eta(x(t), \gamma)$$

In annual data, the study intensity could be interpreted for example as credits, or passed courses, per year. The study intensity at time t is influenced by the covariates measured at time t. A positive estimate for a parameter implies that the variable increases intensity or accelerates the pace of studies. The cumulative number of credits at time t is then

$$C(t) = \int_{h=0}^{t} \eta(h)dh = \int_{h=0}^{t} \eta(x(h), \gamma)dh.$$

The graduation requirement is typically defined as a minimum number of credits. If the credit requirement is $\underline{c}$, the student graduates when

$$\begin{aligned} C(t) &\geq \underline{c} \\ \int_{h=0}^{t} \eta(x(h), \gamma)dh &\geq \underline{c} \end{aligned}$$

In practice, the estimation problem is that one would need to estimate the effect of the covariates on the study intensity, but the study intensity is not observed. Only the graduation times, or the time when the cumulative study intensity exceeds the graduation requirement is observed in data. On the other hand, the cumulative effect of the covariates is the question of interest. If one assumes that, conditional on covariates, the study intensities in the different periods are independent and normally distributed, the problem could be formulated in terms of stochastic difference equations. Cumulative sum of credits is

$$C(t+1) = C(t) + \widehat{\eta}(t) + u(t)\sigma,$$

where $u(t)$ are iid N(0,1) variates and $C(0) = 0$. This is a Wiener process with drift $\widehat{\eta}(t)$. The time when such a process hits a boundary $\underline{c}$ has an inverse Gaussian distribution with density and survival functions given in Lancaster, 1990, p. 119.

Although the process has some very reasonable distributional properties[5], there are also conceptual problems. In particular, the normal error distribution may produce also negative shocks, cumulative credits could also decrease. Some other error distribution restricted to positive values could solve the problem, but would create very complicated expressions for the cumulative sum of credits. Also the hazard function of the inverse Gaussian distribution is a complicated expression and introducing unobserved heterogeneity would lead to analytically untractable expressions.

[5] For example $lim_{t\rightarrow\infty}\ S(t) = 0$, $E(T) = \frac{\underline{c}}{\widehat{\eta(t)}}$ and $Var(T) = \frac{\underline{c}\sigma^2}{\widehat{\eta(t)^3}}$.The graduation hazard rises from zero to a maximum at time t_m and then falls but is in the limit positive$\frac{\widehat{\eta(t)^2}}{2\sigma^2}$.

To keep the estimation problem simpler, we made slightly different, but no stronger assumptions. To specify the baseline hazard, we assume that the credits accumulate according to a Poisson process with parameter $\eta(t)$. In a Poisson-process the time until the next occurence is independent of the time elapsed. The waiting time between the consecutive credits follows an exponential distribution with expected waiting time equal to $\frac{1}{\eta(t)}$. The expected graduation time with constant study intensity, i.e. the expected time until $\underline{c}$ credits is simply the sum of waiting times $\frac{\underline{c}}{\eta}$ i.e. the graduation requirement divided by the study intensity.

To estimate the model we also need to derive the distribution of graduation times. A standard result with Poisson process with constant intensity is that the number (j) of events in time (t) follows a Poisson distribution

$$p_j(t) = \frac{e^{-\lambda t}(\eta t)^j}{j!}$$

It can also be verified that the result generalizes to time - variant intensities (Lancaster 1990, p. 88)

$$p_j(t) = \frac{\exp\left(-\int_0^t \eta(u)du\right)\left(\int_0^t \eta(u)du\right)^j}{j!}$$

Now the survival function for graduation times can be derived noting that

$$\begin{aligned} S(t,\eta_t) &= P(T>t) = P(j<\underline{c}) \\ &= \sum_{j=0}^{\underline{c}-1} p_j(t) = \sum_{j=0}^{\underline{c}-1} \frac{\exp\left(-\int_0^t \eta(u)du\right)\left(\int_0^t \eta(u)du\right)^j}{j!} \end{aligned}$$

where $\underline{c}$ is the required number of credits for graduation. The distribution function is then $F = 1-S$. Differentiating the distribution function with respect to t produces the density function for the graduation times.

$$\begin{aligned} f(t) &= \frac{\partial}{\partial t}\left[1-\sum_{j=0}^{\underline{c}-1} \frac{\exp\left(-\int_0^t \eta(u)du\right)\left(\int_0^t \eta(u)du\right)^j}{j!}\right] \\ &= -\sum_{j=0}^{\underline{c}-1} \frac{-\eta(t)\exp\left(-\int_0^t \eta(u)du\right)\left(\int_0^t \eta(u)du\right)^j + \exp\left(-\int_0^t \eta(u)du\right)\times j\left(\int_0^t \eta(u)du\right)^{j-1}\eta(t)}{j!} \\ &= \eta(t)\exp\left(-\int_0^t \eta(u)du\right)\times\sum_{j=0}^{\underline{c}-1} \frac{\left(\int_0^t \eta(u)du\right)^j + j\left(\int_0^t \eta(u)du\right)^{j-1}}{j!} \\ &= \eta(t)\exp\left(-\int_0^t \eta(u)du\right)\times\frac{\left(\int_0^t \eta(u)du\right)^{\underline{c}-1}}{(\underline{c}-1)!} \end{aligned}$$

With discrete data, it is natural to replace the integral in the above expression with a sum

$$\int_0^t \eta(u)du = \sum_{u=1}^{t} \eta_u = \overline{\eta}t$$

where $\overline{\eta}$ is average intensity from period one up to time t. The density function then simplifies to

$$\frac{\eta(t)}{\overline{\eta}} \exp\left(-\overline{\eta}t\right) \times \frac{\overline{\eta}^{\underline{c}} t^{\underline{c}-1}}{(\underline{c}-1)!}$$

With constant covariates $\frac{\eta(t)}{\overline{\eta}} = 1$ and the distribution is two parameter Gamma distribution with parameters$\overline{\eta}$ and $\underline{c}$. Even with time-varying covariates Gamma distribution should be a reasonable approximation. The intuitive reason for the appearance of the factor $\frac{\eta(t)}{\overline{\eta}}$ in the density function is that the intensity in the last period affects not only the "average intensity" but also the increase in the graduation hazard due to the current study intensity.

The simplest way of introducing covariates in the model is to assume that they influence study intensity addively. If

$$\eta_t = \gamma x_t$$

then

$$\int_0^t \eta(u)du = \sum_{u=1}^{t} \gamma x_u = \gamma\overline{x}t$$

Hence, the covariates enter into the model as cumulative averages. The formulation is closely related to modeling the effect of average exposure in the epidemiological research. In the model the covariates enter as multipliers of time. The model is, therefore, an accelerated failure-time model. If the baseline survivor function is $S_0(t)$,the survivor function for a person with covariates x is then

$$S(x,t,\gamma) = S_0(t \times \eta(x,\gamma))$$

An individual who would have had a survival time t in the absence of covariates now has survival time t/η.i.e. time is accelerated by the factor η. The idea generalizes to the time-dependent covariates. For an individual with characteristics $x(t)$ at time t^x the time evolves relative to time t^0 for that individual would have been if $x = 0$ according to (Cox and Oakes, 1984)

$$\frac{dt^x}{dt^0} = \frac{1}{\eta(x(t),\gamma)}$$

i.e.

$$t^0 = \int_0^{t^x} \eta(x(h),\gamma)dh$$

We also need to account for the unobserved heterogeneity between individuals. We perform this by estimating parametric frailty models with gamma frailty distribution. We assume that the frailty terms enter into hazard function multiplicatively

$$\lambda(t|\alpha) = \alpha\lambda(t)$$

and that the distribution of frailty term is Gamma(a, b) with density

$$g(x) = \frac{x^{a-1}e^{-x/b}}{\Gamma(a)b^a}$$

Adding Gamma frailty into a Weibull duration distribution produces a mixture distribution known as the Burr distribution. Details on the resulting mixture distibution where Gamma frailty is added to the two-parameter Gamma distribution and the possibilities of adding heterogeneity term into the Gamma distribution directly are to be worked out in the next draft of this paper.

5 Data

Our original data is a random sample of 350 000 individuals aged between 12 and 74 in 1990 from the Employment Statistics (ES). Sample is representative and includes approximately 8 percent of population in the relevant age. The Employment Statistics is a main labor market database of Statistics Finland containing information on individual income, employment, education, household composition etc. The data are based on approximately 30 official registers. For our purposes the two most important source registers are the Student Register that records enrolment in universities, and the Register on Completed Education and Degrees that provides the level and field of education achieved and the date of the latest degree. It should be stressed that this is not a survey, universities and senior secondary schools are obliged to report granted degrees to Statistics Finland. Currently ES cover the years from 1987 to 1999.

We use two alternative measures of the beginning the university studies. The first is the senior secondary school graduation date and the second is the first university enrollment date. Since university admission is competitive, students often apply several times before being admitted to a university. If the policy interest is in the efficiency of the university system, the first enrolment date is the relevant measure. However, if the policy interest is in the eventual graduation age, the time until admission should also count and senior secondary school graduation date appears to be a better measure. This argument is also supported by the fact that students often conduct university level studies in other educational institutions before being admitted to a university. These credits can usually be converted to university credits after admission.[6]

Thus, from the ES sample we select two samples using different definition of the beginning of studies. The first sample includes students who have graduated

[6] Using the senior secondary school graduation date as a beginning of spells also allows us to extend our data beyond the first observation. If the highest degree in 1987 (first year of data) is senior secondary school and graduation date 1985 or 86 we include the observation in data. The year 1985 is chosen so that the students cannot have achieved a post secondary school degree before entering our data.

from senior secondary school between 1985 and 1995 and been enrolled in a university in at least one year betwee 1987 and 1999. The second sample includes students whose first university enrollment year was between 1985 and 1995 and who have been enrolled in a university in at least one year between 1987 and 1999. For both samples we collect information on student's senior secondary school graduation date, first and last university enrolment date, name of the university where enrolled, the field of study, the university graduation date, and degree achieved (if the student has graduated by the end of 1999). In addition, we record the student's sex, age, marital status, presence of children, months in employment and earnings. For the time variant variables these data are collected for each year that a student has been enrolled[7]. In addition, we use the full ES sample (not only the students) to calculate the unemployment rates of recent graduates in the student's field of study and add this information to the student records.

In the ES data the students can be linked to their parents. We were able obtain data on parents' income and education for most students. We converted parents' education into years of schooling using the mean years of schooling required to complete different levels of education. As a measure of parents' income we used mean annual income of mother and father. We measures parents' characteristics at the time when the student enters into university. To reduce measurement error we averaged the parents' information over the first four years that the student was enrolled.

For students who graduated from senior secondary school after 1989, we also obtained their grades in the senior secondary school final exam (matriculation examination). This exam is graduation requirement and it takes place in all schools on the same date in an identical form so that the results are comparable across schools. The results are also standardized to be comparable across years. We used the mean score in four compulsory exams in our analysis.

We added program- and city -specific information to the data from two addiotional sources. We used KOTA-database of the Ministry of Education to calculate student-teacher -ratio in different programs and different universities. We calculated this by dividing the total number of enrolled students in the program by total man-years of the teaching faculty. This was the only resource measure that could be calculated in a consistent way for the whole period. Information on local labor market conditions was available from the regional database of Statistics Finland (ALTIKA). From this database we obtained municipality-level unemployment rates for each year at each University location.

The graduation date is recorded in months. In principle, students may graduate at any time during the academic year. This is reflected in graduation dates that are rather dispersed in data, though there are clear peaks in May

[7] For the two earliest years of data 1985-1986 there is an obvoius problem since our data only extends to 1987. However, for most time-variant variables, such as age and presence of children, we can use the data from later years to construct data for the first years. For variables such as earnings or months in employment we simply copy the 1987 values to the two previous years. Naturally, we also repeat our analysis without these two first years to assess the robustness of the results.

and December (Figure 4). We measure the time until graduation as the number of months from the September of the first enrolment year (or September of the senior secondary school graduation year). For the students who have not achieved a degree by our last observation date (December 1999) we mark the durations as censored at this point. Some of these students are drop-outs, some still continue their studies. To maintain consistency across different student cohorts, we measure all other incomplete spells up to the time of last enrolment date. So if a student started in September 1990, never graduated, and was last enrolled in the fall term of 1994, we record a duration as being censored in September 1994. The duration of this incomplete spell is, therefore, 48 months. Alternatively, one can see our estimates as coming from a competing risks model where the spells are observed until the first failure (graduation or drop-out). If the risks of graduation and drop-out are independent, the drop-outs are not informative for the graduation hazards and can be treated as censored observations. Descriptive statistics of the whole sample, graduates, drop-outs, and undergraduates enrolled at the end of 1999, are in the Appendix.

6 Results

Before turning to the results from estimating parametric models, a look at some nonparametric estimates is instructive. Figures 5 and 6 plot Kaplan-Meyer survival function estimates by the first enrollment year and by the study field. It is difficult to see much differences across the different starting years in figure 6. In particular, the students who started their university studies after the student aid reform in 1992 appear to have almost identical graduation profiles than the students who started earlier. However, censoring due to the last observation date in the end of 1999 makes this comparison rather difficult. Students who started in 1995 are under observation for only four and a half years. It is much easier to notice the differences across the study fields in figure 6 where we plot the survival functions in the five largest fields: Education, Business, Engineering, Natural Sciences and Humanities. Half of the students in teacher education graduate in five years, whereas, in humanities half of the students "survive" for more than 7 years. Survival rates are also high in Natural sciences and in Engineering.

Figure 7 plots the empirical hazard rates for the whole sample. The estimated hazard and survival functions are also shown in Table 1. It is immediately obvious that the graduation hazard is highly non-monotonous. The estimated hazard is almost zero for the first three years and then increases rapidly until the eighth study year. If the student is still enrolled after eight years the conditional probability of graduation starts to decrease. The shape of the estimated hazard function has important implications for the parametric form of the duration model. For example, the constant hazard that is implied by the Exponential distribution and the monotonously increasing or decreasing hazard implied by the Weibull distribution are clearly not consistent with the data. In section 4, we argued the the appropriate duration distribution would be a two-parameter Gamma distribution. However. also the Gamma distribution implies monotonously increasing or decreasing hazard rates. One possible mechanism that could generate non-monotonous hazard is that there is unobserved heterogeneity across individuals. As the students who put more effort in their studies graduate faster, the sample becomes increasingly composed of "lazier" students with lower graduation hazards. Unobserved heterogeneity causes negative duration dependence in the observed hazard rates.

We tried to estimate a Gamma distribution with unobserved heterogeneity terms, but this lead to computational problems. As a premilinary solution to the problem we estimated the models using Weibull baseline hazards with Gamma heterogeneity (frailty). This produces a mixture distribution that is known as the Burr distribution (Lanacaster 1990, 68). The hazard rate in the Burr-distribution is flexible and may generate patterns like the observed hazard rate in figure 7.

We estimate the duration of university studies using two alternative measures of the beginning of the studies. First, we count the elapsed-time-to-degree from the first university enrolment year. Second, we use the senior secondary school graduation date as a starting point. However, in this version only estimations using the first enrolment year as a starting point are reported. We report the results in Table 2. The results are reported as "marginal effects". These marginal effects indicate the expected change in the duration of studies (in years) if the explanatory variable increases by one unit, all other variables at means. For the dummy variables the change from 0 to 1 is reported. The first column presents the basic results. In the second column we include matriculation exam scores and parental characteristics. Unfortunately, we had information on matriculation examination scores only for those who had graduated from senior secondary school between 1990 and 1998. Also parents' characteristics are missing for a number of students. The estimates in column 2 refer to this smaller sample. In column 3 we restrict the sample to those who started after 1987. This serves as a robustness check, since we had to impute some variables for the years 1985 and 1986.

According to the results in Table 2, female students gradutate, on average, 2 - 3.5 months faster than male students. One explanation for this is that some male students are in the military service during their university studies[8]. Presence of children slows graduation by about 3.5 months on average, but the presence of children is likely to affect men and women differently. Therefore, we used an interaction term female*children in addition to gender and children dummies, and found out that for women the presence of children adds 6 months to the graduation times. For men the effect is not statistically significant.

Other results related to the student background indicate that the married or cohabiting students complete their studies considerably faster than single students. Because the most of the entrants are about the same age, age does not have a significant effect on the duration of studies or the effect is small and

[8]Military service is compulsory for men in Finland. The length of the service is 6-12 months, 13 months in the civilian service. However, most men enter university after military service.

the sign of the effect is ambiguous. Swedish speaking student study longer than Finnish speaking students.

We defined 12 study fields: Education; Arts; Humanities and Theology; Physical education; Social sciences; Health care, Veterinary and Dentistry; Medicine; Law; Business; Natural sciences; Agriculture and forestry; and Engineering. Education is used as the reference group. There are huge differences in graduation times between the study fields. The estimated coefficients show that students of education have the shortest elapsed-times-to-degree. Business students also graduate graduate faster than the average. Arts, and humanities and theology take significantly longer than education. In medicine the number of credits required for graduation is higher (250 credits) than in other study fields (160-180 credits), which increases the elapsed-times-to-degree for medical students. A change in the study field increases the time needed for graduation, because the student usually has to complete additional courses and is not likely to get all the previous credits included in the new degree. The students who have changed the study field graduate on average 5 months slower than others.

Prior achievement level affects duration of studies. According to the resluts, the students who have done well in matriculation examination, graduate faster. Student's family background measured as parents' income and/or parent's education did not have any significant effect on graduation times.

Some very interesting resluts were found regarding the effect of unemployment on graduation times. We found that local unemployment rate increases the duration of studies, but unemployment rate of recent graduates decreases study times. A natural explanation would be that local unemployment rate is a good proxy for the student employment opportunities. When there are a number of jobs available, the students are drawn to the labor market and graduation is delayed. In contrast, if the labor market prospects after graduation appear good as the unemployment rate of recent graduates in the field is low, the incentives to graduate are larger, and study times shorter.

The only resource variable that we had access to was the number of students per teaching faculty. The effect is again significant, though rather small in magnitude. Higher teacher student ratio leads to longer graduation times.

Finally, our main coefficient of interest, the effect of student aid reform, appears hard to pin down. There is some evidence that graduation times of the more recent entry cohorts are shorter after controlling for the other time-varying factors such as unemployment rates and university resources. Whether this decline can be attributed to the student aid reform, is a problematic question. Simply adding a variable incicating the level of student aid or a dummy for the new student aid system would lead to biased estimates, beacause this variable would be correlated with study times. For example, the students who begin their studies during the eighties, are influeced by the new student aid system only if they stay at the university until 1992. To get around this endogeneity problem, we experimented by replacing the entry cohort dummies by a fraction of the first five study years that the student would be under the new aid system. With this definition all students in the same entry cohort get the same value irrespective of their eventual study duration. The coefficient is negative implying that the reform shortened study times, but it is not statistically significant at conventional levels.

7 Conclusion

The purpose of this paper was to examine the effects of the changes in the student financial aid and university funding on the duration of university studies. We had a representative sample of Finnish university students from the years 1985-1999 and information on these students' background, university enrolments, field of study, the degree achieved and the time of graduation.

We estimated parametric frailty models with gamma frailty distribution explaining the duration of university studies with students marital status, sex, presence of children, study field, initial achievement level, employment and labor market conditions, and parents' education and income. According to our results female students graduate faster, but the presence of children increase graduation times for female students. We also found that university resources matter. High student teacher ratios lead to longer graduation times. The evidence on the effects of student funding are more uncertain. However, it seems that student aid reform has reduced student employment and shortened graduation times. Labor market conditions have a large impact. Low local unemployment rate at the location of the university increases student employment and delays graduation, labor market prospects after graduation have no significant impact. After controlling for all these factors there are still large differences across different fields. Students in education and business graduate quickly, students in humanities and engineering take considerably longer.

References

[1] Ben-Porath, Y (1967). The production of human capital and the life cycle of earnings, Journal of Political Economy 75.

[2] Blomster, Peter (2000). Yliopisto-opiskelijoiden toimeentulo ja opintotuki 1900-luvun Suomessa. KELA sosiaali- ja terveysturvan tutkimuksia 56. Helsinki.

[3] Booth, Alison L. and Satchell, Stephen E. (1995). The Hazards of Doing a PhD: An Analysis of Completion and Withdrawal Rates of British PhD Students in the 1980s, Journal of the Royal Statistical Society. Series A (Statistics in Society), Volume 158, Issue 2, 297-318.

[4] Cox, R. and Oakes, D. (1984). Analysis of Survival Data. London: Chapman and Hall.

[5] Ehrenberg, Ronald G. and Mavros, Panagiotis G. (1995). Do Doctoral Students' Financial Support Patterns Affect Their Times-to-Degree and Completion Probabilities, The Journal of Human Resources, 30(3) 581-609.

[6] Havén, Heikki (1999). Education in Finland. Statistics and Indicators. Statistics Finland. Abridged version also available through the Internet http://www.stat.fi/tk/he/edufinland/

[7] Härkönen, Hanna (2001). Opiskelijan rahapussilla. Selvitys Helsingin yliopiston opiskelijoiden toimeentulosta ja asumisesta keväällä 2001. Helsingin yliopiston ylioppilaskunta.

[8] Kalbfleisch, J. D. and Prentice, R. L. (1980). The Statistical Analysis of Failure Time Data. John Wiley and Sons.

[9] Katz, L. and Mayer, B. (1990). The Impact of Potential Duration of Unemployment Benefits on the Duration of Unemployment, Quarterly Journal of Economics, November, 973-1002

[10] KELA Student financial aid 2000-2001. http://www.kela.fi/opintotuki/esite.htm

[11] Kennan (1985). The Duration of Contract Strikes in U.S. Manufacturing, Journal of Econometrics 28(1), 5-28.

[12] Kiefer, Nicholas M. (1998). Economic Duration Data and Hazard Functions, Journal of Economic Literature 26(2) 646-679.

[13] Lancaster, Tony (1990). The Econometric Analysis of Transition Data Cambridge University Press

[14] Löfgren, Curt and Ohlsson, Henry (1999). What determines when undergraduates complete their thesis? Evidence from two economics departments, Economics of Education Review 18, 79-88.

[15] Meyer, B. (1990). Unemployment Insurance and Unemployment Spells, Econometrica 58 (July), 757-782.

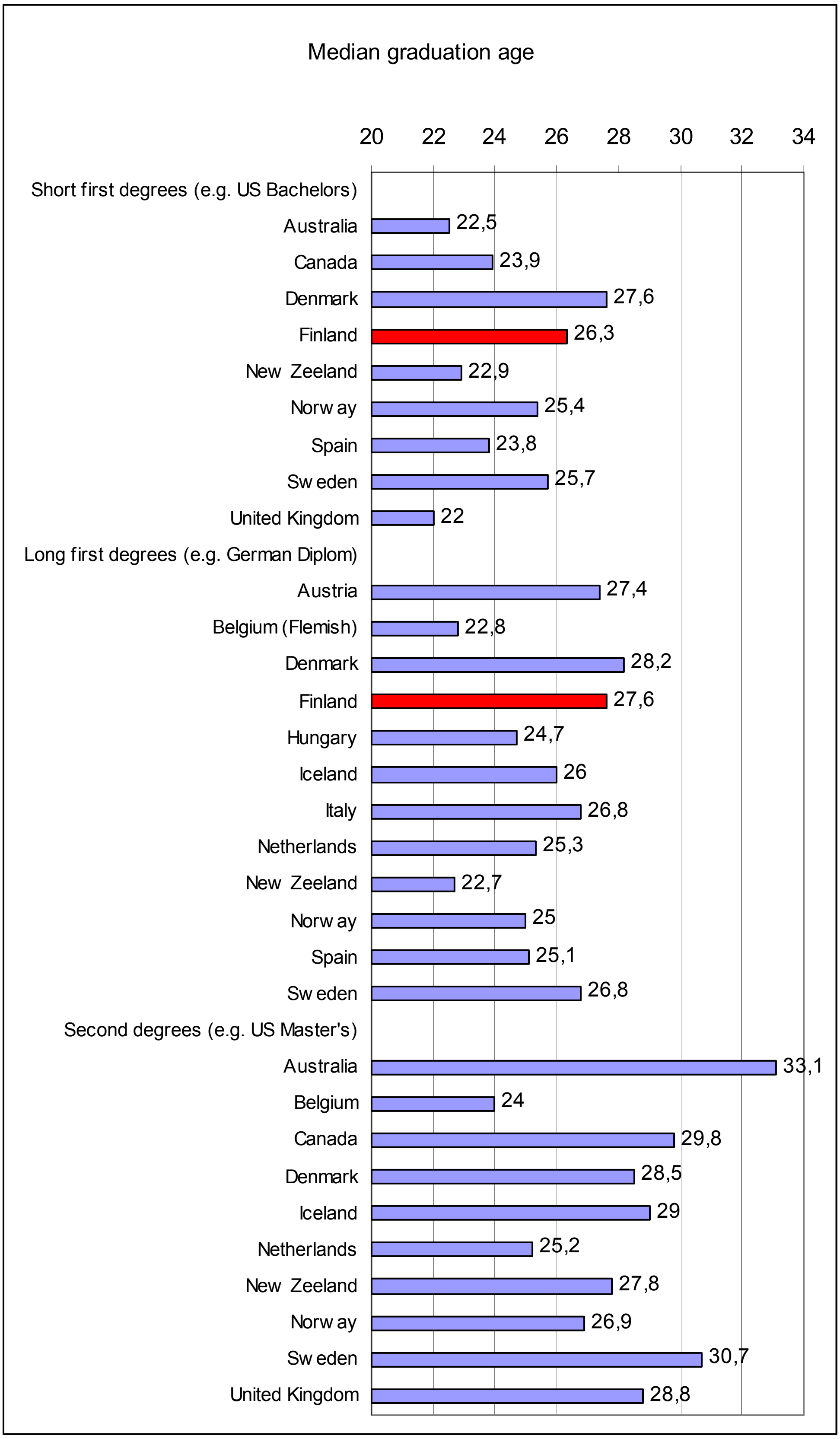
[16] Mincer, Jacob (1974). Schooling, experience and earnings, New York Columbia University Press.

[17] van Ophem, Hans and Jonker, Nicole (1999). The Duration of Higher Education: Semiparametric Estimation of a Dependent Risks Model, University of Amsterdam.

[18] Siegfried, John J. and Stock, Wendy A. (2001). So You Want to Earn a Ph.D in Economics? How Long Do You Think it Will Take? Journal of Human Resources 36(2) 364-378.

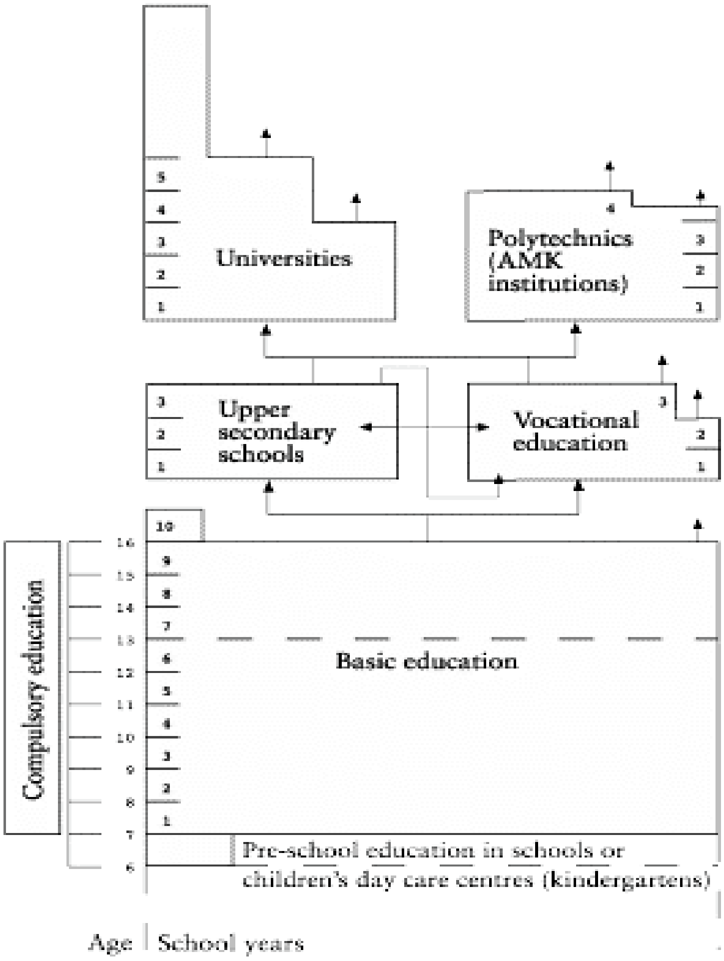
[19] Raivola, Reijo, Zechner, Minna and Vehviläinen, Jukka (2000). Opintotuki - opiskelijapalkka vai koulutusinvestointi. Opetusministeriö 14/2000.

Figure 1. Median graduation age in tertiary education



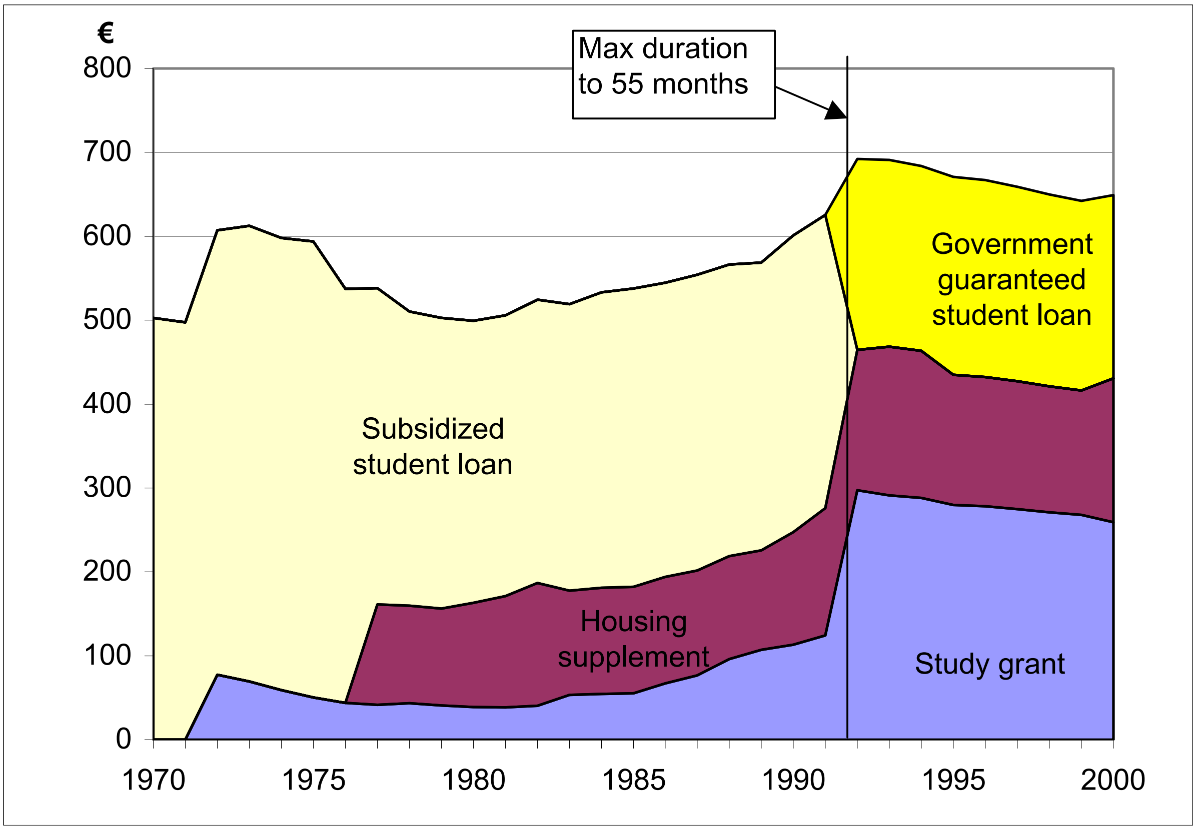
Source: OECD Education at a Glance 1998

Figure 2. Finnish education system



Source: National Board of Education 2001.

Figure 3. Maximum amount of Student financial aid in Finland 1970-2000 (€).



In 2000 prices (euro) deflated by cost of living index 1951=100

Figure 4. Graduation Months for Master's Degree Graduates in 1990-1998.

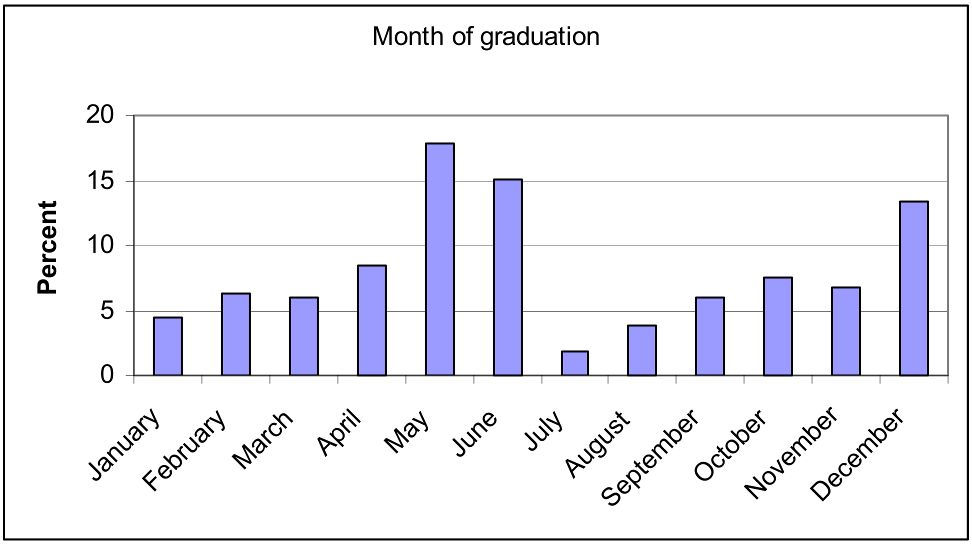


Figure 5. Kaplan-Meier survival estimates by first enrolment year.

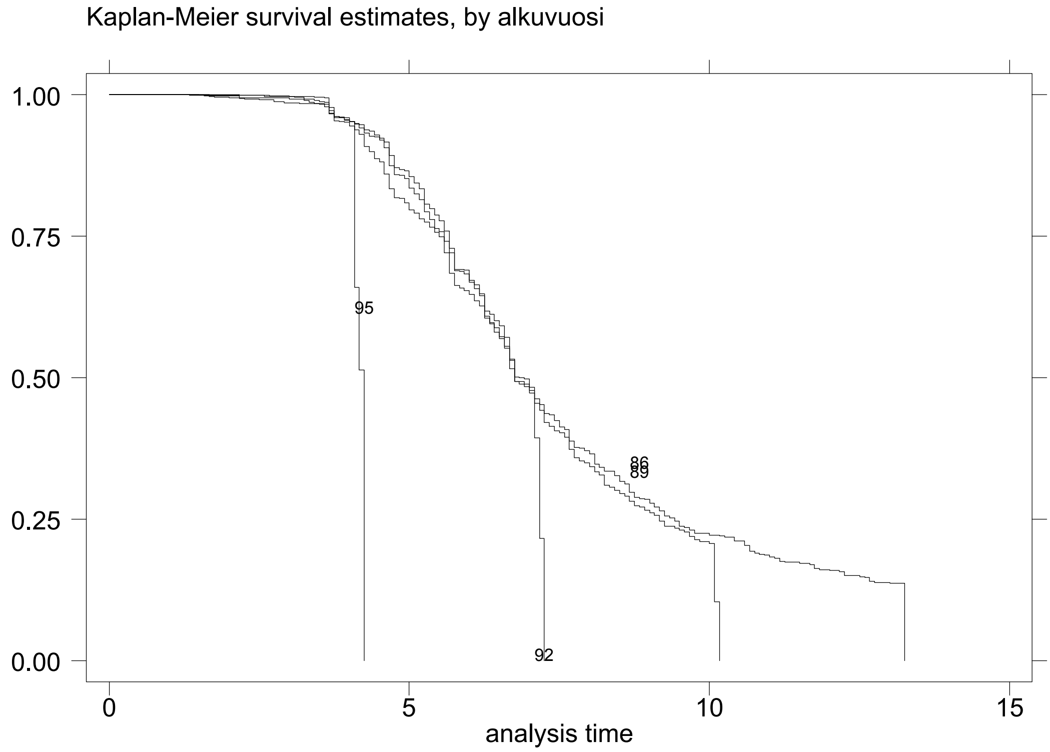


Figure 6. Kaplan-Meier survival estimates by study field.

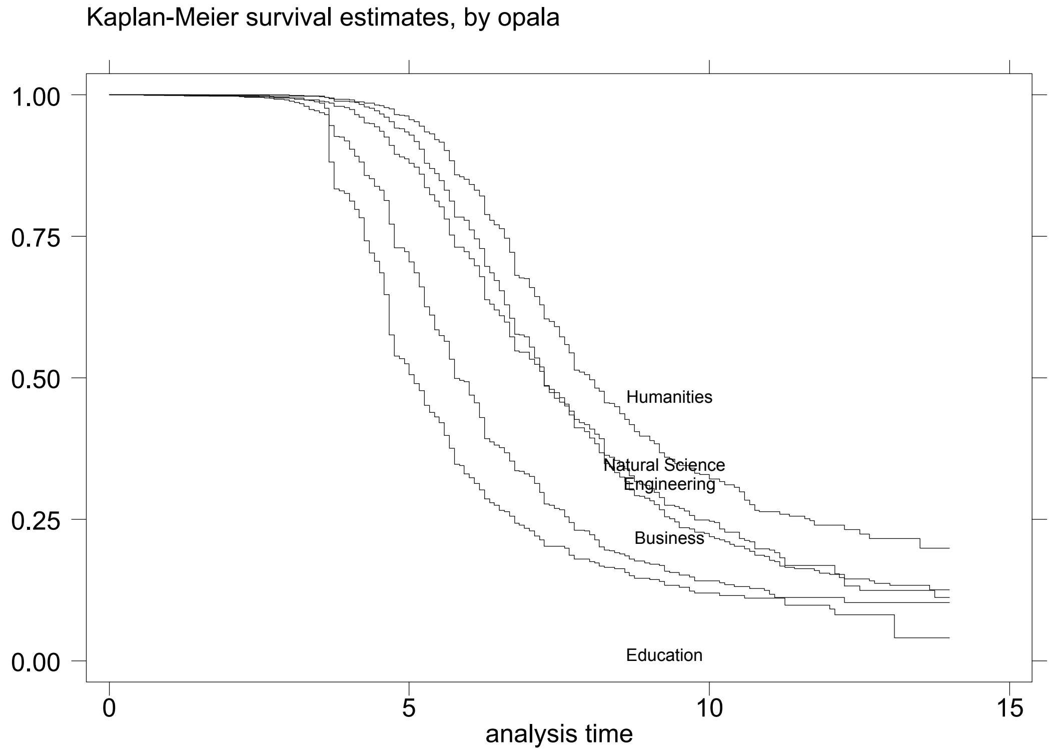
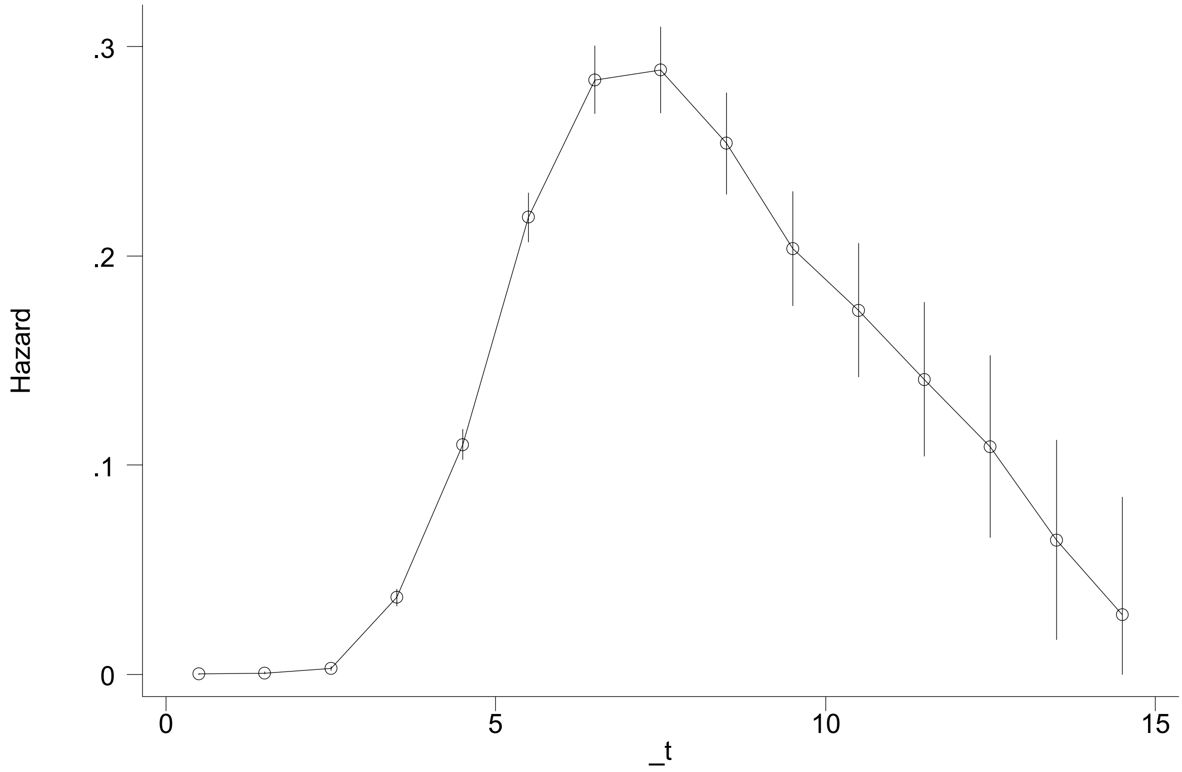


Figure 7. Empirical Hazard Rate.



Note: To estimate the empirical hazard the data have been split into 12 month intervals. The estimated hazard rate is number of graduates during the year divided by the population at risk at the beginning of the year. The vertical lines indicate the 95% confidence intervals.

Table 1. Life – table.

Study Year	Begin Total	Censored	Graduates	Hazard	Cum. Graduation
0-1	9961	91	1	0.0001	0.0001
1-2	9869	118	6	0.0006	0.0007
2-3	9745	255	26	0.0027	0.0034
3-4	9464	147	338	0.0367	0.0393
4-5	8979	1008	882	0.1098	0.1393
5-6	7089	833	1314	0.2184	0.3088
6-7	4942	608	1154	0.2842	0.4808
7-8	3180	433	748	0.2889	0.6118
8-9	1999	330	413	0.2538	0.6992
9-10	1256	238	210	0.2035	0.7548
10-11	808	192	114	0.1740	0.7940
11-12	502	153	56	0.1409	0.8212
12-13	293	121	24	0.1088	0.8396
13-14	148	71	7	0.0642	0.8496
14-15	70	69	1	0.0286	0.8538

Table 2. Determinants of the duration of university studies.

	(1) _t	(2) _t	(3) _t
Female	-0.04 (0.01)**	-0.02 (0.01)	-0.04 (0.01)**
Female*Children	0.24 (0.05)**	0.16 (0.12)	0.28 (0.06)**
Children	0.00 (0.05)	-0.06 (0.10)	-0.03 (0.05)
Married	-0.20 (0.02)**	-0.14 (0.04)**	-0.20 (0.02)**
Unemployment rate of the study field	0.39 (0.28)	-2.60 (1.19)*	0.31 (0.34)
Local unemployment rate	-0.59 (0.14)**	-0.52 (0.23)*	-0.69 (0.14)**
Swedish speaking	0.04 (0.01)**	0.02 (0.02)	0.03 (0.01)*
Age at the beginning of studies	0.00 (0.00)	-0.02 (0.01)*	0.00 (0.00)
Students/teachers	0.01 (0.00)**	0.01 (0.00)**	0.01 (0.00)**
Starting year 85 (Starting year 95 as reference group)	0.08 (0.03)*		
Starting year 86	0.04 (0.03)		
Starting year 87	0.06 (0.03)		0.04 (0.03)
Starting year 88	0.07 (0.03)**		0.06 (0.03)*
Starting year 89	0.06 (0.02)*		0.05 (0.02)*
Starting year 90	0.07 (0.02)**	0.34 (0.11)**	0.07 (0.02)**
Starting year 91	0.07 (0.02)**	0.35 (0.11)**	0.07 (0.02)**
Starting year 92	0.05 (0.02)*	0.32 (0.11)**	0.06 (0.02)**
Starting year 93	0.04 (0.02)	0.30 (0.11)**	0.04 (0.02)*
Starting year 94	0.03 (0.02)	0.15 (0.11)	0.03 (0.02)
Changed study field	0.35 (0.05)**	0.18 (0.07)*	0.35 (0.05)**
Study field: (Education as reference group)			
Humanities	0.34 (0.03)**	0.51 (0.13)**	0.34 (0.04)**
Theology	0.22 (0.05)**	0.36 (0.15)*	0.21 (0.05)**
Industrial Arts	0.37 (0.05)**	9.48 (4,708.91)	0.36 (0.06)**
Music	0.52 (0.05)**	0.79 (0.18)**	0.53 (0.06)**
Theatre	-0.04 (0.06)	0.39 (0.20)*	-0.06 (0.07)
Visual Arts	0.39 (0.12)**	0.56 (0.21)**	0.39 (0.12)**
Physical Education	0.19 (0.04)**	0.02 (0.08)	0.18 (0.05)**
Social Sciences	0.15 (0.02)**	0.26 (0.08)**	0.17 (0.03)**
Psychology	0.21 (0.04)**	0.25 (0.09)**	0.19 (0.04)**
Healthcare	0.10 (0.08)	-0.06 (0.17)	0.10 (0.08)
Law	0.03 (0.03)	0.17 (0.08)*	0.05 (0.03)
Business	-0.02 (0.02)	0.16 (0.08)*	0.02 (0.03)
Natural Sciences	0.28 (0.02)**	0.36 (0.08)**	0.28 (0.02)**
Agriculture and Forestry	0.26	0.37	0.28

	(0.03)**	(0.07)**	(0.03)**
Engineering	0.26	0.29	0.26
	(0.02)**	(0.06)**	(0.02)**
Medicine	0.32	0.33	0.32
	(0.02)**	(0.05)**	(0.02)**
Dentistry	0.17	0.18	0.16
	(0.04)**	(0.06)**	(0.04)**
Veterinary	0.32	0.47	0.34
	(0.05)**	(0.11)**	(0.05)**
Pharmacy	0.32	0.29	0.31
	(0.04)**	(0.09)**	(0.04)**
Parents' income 2 (Lowest income group as reference group)	-	-0.02	-
		(0.02)	
Parents' income quartile 3	-	-0.01	-
		(0.02)	
Parents' income quartile 4	-	-0.00	-
		(0.02)	
Parents' years of education	-	0.00	-
		(0.00)	
Mean score in Matriculation Examination	-	-0.04	-
		(0.01)**	
Constant	1.44	1.93	1.52
	(0.06)**	(0.23)**	(0.07)**
Ln p: (Weibull exponent)	2.21	2.26	2.22
	(0.02)**	(0.06)**	(0.03)**
Ln theta: (heterogeneity)	1.22	0.01	1.18
	(0.04)**	(0.32)	(0.05)**
Number of Students	9,942	4,584	8,483
Number of Graduates	5,281	605	4,164
Number of Observations	72,021	22,524	60,298

Standard errors in parentheses, * significant at 5%; ** significant at 1%. All models are based on Weibull distribution and estimated with gamma frailty.

Appendix. Descriptive statistics at the beginning of studies (first enrolment year).

All students.

Variable	Obs	Mean	Std. dev.	Min	Max
Female	9,276	0.53	0.50	0	1
Children	9,276	0.01	0.11	0	1
Female*Children	9,276	0.01	0.10	0	1
Married	9,271	0.02	0.15	0	1
Months in employment	9,276	3.14	3.68	0	12
Student aid reform	9,276	0.43	0.49	0	1
Unemployment rate	9,276	0.06	0.05	0	0.18
Yearly earnings (€)	9,276	3,205	3,581	0	39,527
Age at the beginning of studies	9,276	20.40	2.29	17	54
Swedish speaking	9,276	0.08	0.27	0	1
Parents’ yearly income (€)	9,276	51,539	34,883	0	964,669
Parents' education	9,177	12.28	2.83	9	22
Study field changes	9,276	0.14	0.35	0	1
Mean score in the matriculation exam	4,566	5.08	0.80	0	6
Humanistics and theology	9,276	0.17	0.38	0	1
Arts	9,276	0.02	0.15	0	1
Physical education	9,276	0.005	0.07	0	1
Social sciences	9,276	0.10	0.30	0	1
Healthcare, veterinary, dentistry	9,276	0.02	0.13	0	1
Law	9,276	0.03	0.18	0	1
Business	9,276	0.10	0.30	0	1
Natural sciences	9,276	0.17	0.37	0	1
Agriculture and forestry	9,276	0.02	0.15	0	1
Engineering	9,276	0.20	0.40	0	1
Medicine	9,276	0.04	0.20	0	1

Graduates.

Variable	Obs	Mean	Std. dev.	Min	Max
Female	4,822	0.55	0.50	0	1
Children	4,822	0.01	0.10	0	1
Female*Children	4,822	0.01	0.09	0	1
Married	4,819	0.02	0.15	0	1
Months in employment	4,822	3.34	3.68	0	12
Student aid reform	4,822	0.23	0.42	0	1
Unemployment rate	4,822	0.04	0.04	0	0.18
Yearly earnings (€)	4,822	3,463	3,337	0	30,808
Age at the beginning of studies	4,822	20.17	1.90	17	54
Swedish speaking	4,822	0.07	0.26	0	1
Parents’ yearly income (€)	4,822	52,102	33,375	0	845,441
Parents’ education	4,781	12.24	2.83	9	22
Study field changes	4,822	0.13	0.34	0	1
Mean score in the matriculation exam	1,671	5.16	0.77	2	6
Humanistics and theology	4,822	0.13	0.34	0	1
Arts	4,822	0.02	0.13	0	1
Physical education	4,822	0.008	0.09	0	1
Social sciences	4,822	0.09	0.29	0	1
Healthcare, veterinary, dentistry	4,816	0.02	0.15	0	1
Law	4,816	0.04	0.19	0	1
Business	4,816	0.13	0.33	0	1
Natural sciences	4,816	0.14	0.35	0	1
Agriculture and forestry	4,816	0.02	0.16	0	1
Engineering	4,816	0.21	0.40	0	1
Medicine	4,816	0.06	0.23	0	1

Drop-outs.

Variable	Obs	Mean	Std. Dev.	Min	Max
Female	1,057	0.57	0.49	0	1
Children	1,057	0.03	0.16	0	1
Female*Children	1,057	0.02	0.14	0	1
Married	1,057	0.04	0.19	0	1
Months in employment	1,057	4.14	4.07	0	12
Student aid reform	1,057	0.34	0.47	0	1
Unemployment rate	1,057	0.05	0.05	0	0.18
Yearly earnings (€)	1,057	4,165	4,371	0	36,910
Age at the beginning of studies	1,057	21.14	2.99	18	54
Swedish speaking	1,057	0.10	0.30	0	1
Parents' yearly income (€)	1,057	46,054	40,494	0	964,669
Parents' education	1,031	11.62	2.65	9	22
Study field changes	1,057	0.05	0.21	0	1
Mean score in the matriculation exam	374	4.72	0.84	2.75	6
Humanistics and theology	1,057	0.21	0.41	0	1
Arts	1,057	0.03	0.17	0	1
Physical education	1,057	0.002	0.04	0	1
Social sciences	1,057	0.10	0.30	0	1
Healthcare, veterinary, dentistry	1,057	0.005	0.07	0	1
Law	1,057	0.01	0.12	0	1
Business	1,057	0.04	0.20	0	1
Natural sciences	1,057	0.31	0.46	0	1
Agriculture and forestry	1,057	0.02	0.13	0	1
Engineering	1,057	0.11	0.32	0	1
Medicine	1,057	0.006	0.08	0	1

Students still enrolled in 1999.

Variable	Obs	Mean	Std. dev.	Min	Max
Female	3,397	0.48	0.50	0	1
Children	3,397	0.01	0.12	0	1
Female*Children	3,397	0.02	0.12	0	1
Married	3,397	0.02	0.15	0	1
Months in employment	3,397	2.54	3.46	0	12
Student aid reform	3,397	0.73	0.44	0	1
Unemployment rate	3,397	0.09	0.05	0	0.18
Yearly earnings (€)	3,397	2,541	3,528	0	39,527
Age at the beginning of studies	3,397	20.48	2.48	18	53
Swedish speaking	3,397	0.08	0.27	0	1
Parents' yearly income (€)	3,397	52,446	34,950	0	832,769
Parents' education	3,365	12.54	2.85	9	22
Study field changes	3,397	0.18	0.38	0	1
Mean score in the matriculation exam	2,521	5.08	0.80	0	6
Humanistics and theology	3,397	0.21	0.41	0	1
Arts	3,397	0.03	0.17	0	1
Physical education	3,397	0.003	0.06	0	1
Social sciences	3,397	0.11	0.31	0	1
Healthcare, veterinary, dentistry	3,397	0.01	0.11	0	1
Law	3,397	0.03	0.18	0	1
Business	3,397	0.08	0.28	0	1
Natural sciences	3,397	0.15	0.36	0	1
Agriculture and forestry	3,397	0.03	0.16	0	1
Engineering	3,397	0.22	0.42	0	1
Medicine	3,397	0.03	0.18	0	1