The Impact of Labor Income Risk on Educational Choices: Estimates and Implied Risk Aversion*

Helena Skyt Nielsen

Annette Vissing-Jorgensen†

January 20, 2006

Abstract

The paper documents that the time-series risk of individual labor incomes has a significant effect on individuals' educational choices. Using a novel Danish data set on labor incomes and educational choices, we classify the full set of post-high school educations into 50 groups based on admissions requirements, length, and topic studied. Income processes are estimated for each education group and the estimates are used in an empirical choice model. The estimated choice model suggests a preference for educations with higher mean incomes and lower risk, with a high variance of permanent income shocks seen as particularly undesirable. Using a structural model of life-time utility maximization we conclude that the parameter estimates of the empirical choice model imply a relative risk aversion coefficient around 5.

*We thank the Danish Social Science Research Council for generous financial support for the project, and Bjarke Christensen and Juanna Joensen for research assistance. Part of the work for the paper was done while Helena Skyt Nielsen was visiting the Finance Department at Kellogg School of Management. We thank the department for its hospitality.

†Department of Economics, University of Aarhus, IZA and CIM, and Kellogg School of Management, Northwestern University, CAM and NBER, respectively.

I. Introduction

In this paper we use data on labor incomes and educational choices to show that the time-series risk of individual labor incomes have a significant effect on individuals' educational choices.

While a very large literature has studied the average return to additional years of schooling, the literature has only recently begun to study what role risk plays in educational choice. Our study is based on a novel panel data set of labor incomes and educational choices for Danish individuals over the period 1984-2000. We sort the total set of all post-high school educations into 50 categories based on admissions requirements, length of study, and topic studied. For each of these 50 education groups we estimate income processes based on data for individuals who started the education and have graduated. The labor income processes estimated are characterized by five parameters: the mean log starting labor incomes, the mean growth rate of labor income up to age 40, the mean growth rate of labor income after age 40, the variance of transitory labor income shocks, and the variance of permanent labor income shocks.

We use these income processes as estimates of the expected income processes a high school graduate (with particular characteristics) would face if choosing a particular education, and estimate a simple model of educational choice. The Danish data, or similar data for another European country, are ideally suited for such a study because admission, with a few exceptions, is based solely on high school GPA and high school math courses taken. This enables us to quite precisely characterize the choice set of educations for each individual. Given the more complex admissions requirements at U.S. schools, such a characterization would be much less precise.

Consistent with individuals having a preference for higher mean incomes and lower risk, the estimates of the empirical choice model document that the predicted log starting labor income and the predicted growth rates have a positive effect on the probability that a particular education is chosen. Conversely, educations for which transitory or permanent labor income shocks have a higher variance are less likely to be chosen. A comparison of the coefficient on mean log starting labor income and the variance of the permanent income shock suggests that individuals require an increase of about 25 percent in their starting labor income in order to be willing to accept an increase in the variance of the permanent shock from its 25th percentile to its 75th percentile. Consistent with theories of life-time smoothing, transitory labor income shocks have a much smaller impact on individuals' choices.

We then take a more structural approach and seek to infer individuals' underlying risk aversion from the coefficients in the empirical choice model. For this purpose we numerically solve a life-time optimization problem and linearize the date zero value function in the five parameters of the labor income process. A given coefficient of relative risk aversion γ will result in a particular set of linearization coefficients. Our estimate of γ is then chosen as that which imply linearization coefficients as close as possible to the coefficients in the empirical choice model. For our baseline specification, this minimum distance approach results in an estimate of γ around 5 with a standard error of about 1.

Our finding of a significant impact of labor income risk on educational choice is of direct relevance for policy making. Within the literature on the average returns to schooling, the effect of educational subsidies on individuals' chosen length of study has been studied. Little work has been done on the role of risk and how policies that alter risk affects the length of study and topic studied. Our work suggests that an increased focus on risk may be valuable, not only for understanding educational choices but also for policy making. Consider, for example, a government (or university) that is deciding how many students to fund or subsidize within each category of educations. Any government objective function that puts weight on the utility of students should reflect both utility differences due to different mean labor income paths, and utility differences due to differences in labor income risk across educations. Ignoring differences in labor income risk could lead to excessive funding of high risk educations, and insufficient funding of low risk educations. Furthermore, consideration of labor income risk leads to potentially important interactions between educational policies and optimal unemployment insurance. Relative to a world with perfect insurance, the presence of idiosyncratic risk that differs across educations will tend to lead to too few individuals choosing high risk educations. Unemployment insurance can help alleviate this. In addition, our results suggest that optimal benefit levels may vary across educations in response to differing labor income risk. The estimation of the effect of income processes on educational choice and of the implied preference parameters serves to provide inputs to such policy calculations.

Our work also contributes to the literature on preference parameter estimation. Enormous attention has been paid to the large value of risk aversion needed to reconcile the equity premium observed in the last century of U.S. stock and T-bill returns. Some have taken this as evidence that individuals in fact are very risk averse. Others have emphasized that the U.S. may be an outlier seen in a world-wide perspective, and that individuals' expected equity premium ex ante (and thus their implied risk aversion) may have been much below the realized average equity premium. In our view, providing evidence on risk aversion from additional sources can help shed new light on these issues. The study of educational choices is particularly useful in this regard because it is one of the most important financial decisions individuals make during their life-time.

The outline of the paper is as follows. Section II describes the relation of our analysis to that of previous papers. Section III outlines our basic theoretical framework of life-time utility maximization. The data are then described in Section IV while Section V describes how we form the educational groups. Section VI contains the estimation of income processes and the empirical choice models is presented in Section VII. Section VIII then turns to the estimation of preferences parameters and Section IX concludes.

II. Relation to prior work

Several authors have shown that individuals are more likely to choose educations with higher predicted present value (PV) of income. Willis and Rosen (1979) use data for WWII veterans for 1968-1971 to show that predicted PV helps explain who goes to college. To account for predicted PV, they include the predicted starting salary gain from going to college and predicted salary growth rates under each choice (college/no college), while accounting for the self-selection into higher education using family background variables. Estimated coefficients on all three earnings variables come out with the expected signs. Boskin (1974) uses the 1967 Survey of Economic Opportunities. He estimates the PV of full-time earnings in 11 broad occupational groups and shows, using a conditional logit model, that the PV of future earnings plays a major role in occupational choice. Studying the choice between five college majors in NLSY data, Berger (1988) finds a positive effect of predicted future earnings on major choice.

As shown by Willis and Rosen (1979) and several authors following them, there is positive self-selection into educational levels. However, placing the focus on choice of major, the evidence for self-selection is weaker. Controlling for IQ and 'Knowledge of the World of

Work' test scores, Berger (1988) finds positive selection into Education and for hourly wages also into Liberal Arts, but not into Business, Engineering or Science. Similarly, in a more recent study, Arcidiacono (2004) shows that the ability distribution across majors is not explained by self-selection. The majority of the ability sorting occurs because individuals with certain ability (e.g. math ability) have a preference to study certain majors (e.g. Natural Sciences). To sum up, the previous literature suggests that there is a high degree of self-selection across levels of education but not across types.

While it is well-established that expected future earnings affect educational choices, there is some disagreement in the literature as to whether risk may affect educational choice. Carneiro, Hansen and Heckman (2003) find that income risk plays little role in the choice between high school and college because it has only a modest effect on life-time utility. Their choice model assumes log utility and does not allow individuals to smooth consumption over the life cycle, since consumption is assumed to equal income at all date. The assumed low risk aversion will tend to bias the effect of risk on utility towards zero. On the other hand, the assumed lack of consumption smoothing will tend to bias the effect of income risk on utility upward, so the overall bias induced by these simplifying assumptions is hard to evaluate. In an analysis of the probability of choosing to attend university in Spain, Diaz-Serrano and Hartog (2002) show that income risk has a negative effect on this probability which disappears for individuals with low risk aversion as measured by a high budget share spent on lottery tickets. Income processes are estimated at the regional level and risk is calculated as the residual variance in a cross-sectional wage equation. Saks and Shore (2004) find that wealthier students tend to choose riskier college majors, especially business.

A number of papers estimate the compensating differential for risk in income. Hartog and Vijverberg (2002) find positive risk compensation using US data. Using Danish data, Diaz-Serrano, Hartog and Nielsen (2003) find positive compensation for both cross sectional and time series risk, the latter of which is largest They test the robustness of the results by application of a number of income mobility measures, and the results tend to be robust. Christiansen, Joensen, and Nielsen (2004) also find evidence of a positive compensation for time-series risk in studies of mean-variance plots for Denmark.

Unfortunately, it is not easy to calculate risk aversion from an observed compensating differential for risk. Admissions requirements mean that individuals are not free to choose among all educations. Some educations could stay more attractive than others in terms of high average pay or low risk. Both the US and Europe feature admissions requirements set by universities or states/the government. Hartog and Vijverberg (2002) make an attempt to back out risk aversion from estimates of risk compensation in earnings based on education-occupation cells. They find that individuals are risk lovers when they use a framework of constant relative risk aversion. When they use trans-log marginal utility, they obtain a relative risk aversion of 0.5. Another attempt to estimate risk aversion has been made by Belzil and Hansen (2004), who estimate risk aversion directly from a structural model of human capital investment (years of schooling), where individuals derive utility directly from income. This assumption could be responsible for their low estimate of risk aversion, 0.93, since a given effect of risk on choice may be consistent with a lower value of risk aversion when individuals are assumed unable to smooth consumption.

Our paper contributes to the literature on risk and education in two ways. First, our analysis focuses on what individuals study, rather than on how long, and the detailed Danish data allows estimation of income processes for fairly detailed education groups. This increases the amount of variation in income processes available to estimate the effect of mean incomes and income risk on choices. Second, unlike previous studies we do not assume that consumption equals income and instead use a life-time optimization simulation when estimating risk aversion. It should also be noted that we explicitly focus on time series income risk throughout the paper, estimated based on income growth rates. This ensures that our risk measures do in fact reflect risk rather than individual effects as could be the case in an estimation of risk based on cross-sectional income differences at a point in time.

III. Theoretical Framework: The Life-time Optimization Problem

Our empirical model (described in more detail below) analyzes the effect of income processes variables on educational choice. In order to estimate risk aversion from the empirical model, additional structure must be imposed. This section lays out our assumed structure.

We assume that 19-year olds base their educational choice on a comparison of the value functions as of age 19 for the feasible set of educations. The choice set will differ across individuals depending on their high school GPA and high school math level. An individual

i who chooses education j faces the following life-time optimization problem

$$
\begin{aligned}
V_{i,j,19} &= \max_{\{C_{19},...,C_{80}\}} E\left[\Sigma_{t=19}^{80}\beta^t u\left(C_{i,t}\right)\right] && (1)\\
\text{s.t. } W_{i,t+1} &= R\left(W_{i,t}+Y_{i,j,t}-C_{i,t}\right) && (2)\\
R &= \begin{cases} R_l \text{ if } W_{i,t}+Y_{i,j,t}-C_{i,t}\geq 0 \\ R_b \text{ if } W_{i,t}+Y_{i,,j,t}-C_{i,t}<0 \end{cases}, \qquad R_b > R_l && (3)
\end{aligned}
$$

$$
\begin{aligned}
W_{i,81} &\geq 0 && (4)\\
W_{i,19} &= 0 && (5)\\
Y_{i,j,t} &= a, \text{ for } t=19,...,t_j && (6)\\
\ln Y_{i,j,t} &= p_{i,j,t}+\varepsilon_{i,j,t},\ p_{i,j,t}=g_{i,j,t}+p_{i,j,t-1}+\eta_{i,j,t}, \qquad \text{for } t=t_j+1,...,64 && (7)\\
Y_{i,j,t} &= b_j P_{i,j,64}, \text{ for } t=65,...,80. && (8)
\end{aligned}
$$

where $C_{i,t}$ is consumption at age t, $W_{i,t}$ is financial wealth at age t, $Y_{i,j,t}$ is labor income at age t. R is the after-tax gross real interest rate which is allowed to differ depending on whether the individual is borrowing or lending. a and b are constants, and $P=\exp(p)$. We assume CRRA preferences so $u\left(C_{i,t}\right)=\frac{1}{1-\gamma}C_{i,t}^{1-\gamma}$ where γ is the coefficient of relative risk aversion. β is the utility time discount factor ($\beta=\frac{1}{1+\delta}$ where δ is the annual utility time discount rate). We assume individuals are economically independent from their parents starting at age 19 and have zero financial wealth at that time. [1]

Individuals are assumed to choose an education at age 19, retire at 65, and live until age 80.[2] If education j is chosen, the individual studies until time t_j. During this time an annual real income of a is earned. In the Danish educational system, all admitted students are automatically eligible to receive free tuition and a monthly stipend that does not vary across educations nor with performance.[3] Empirically there is little difference in median annual incomes for students across educational groups.

[1] Hornstrup and Madsen (2000) show that the wealth of Danish individuals below age 24 is close to zero for non-property owners. Even property owners of that age group have wealth as low as DKK 20,000 on average, corresponding to about $3,000.

[2] The overall average retirement age for the male population was 61-62 year during the nineties. For males with a long further education, the average age was 64.5 years, see Danø, Ejrnæs and Husted (2000).

[3] The annual stipend amounted to DKK 3,907 per month in year 2000, or about DKK 41,625 per year after tax.

During working life, the individual's log labor income is assumed to be characterized by a permanent component p and a transitory component ε. The permanent component changes according to a deterministic growth rate function g and in response to permanent income shocks η. The permanent and transitory shocks are assumed independently drawn over time, independent of each other, and normally distributed, $\varepsilon_{i,j,t} \sim N\left(0, \sigma_{\varepsilon,j}^2\right)$, $\eta_{i,j,t} \sim N\left(0, \sigma_{\eta,j}^2\right)$. This income process is identical to that commonly used in the literature on life-cycle consumption and portfolio choice, see e.g. Gourinchas and Parker (2002), Carroll and Samwick (1997), and Cagetti (2003). Our approach differs from that commonly used in that we estimate income processes for different educations, rather than different occupations.

Our life-time simulation makes (at least) three simplifying assumptions. First, for a given education, we assume that individuals expect to graduate with probability one. To partially address differences in graduation rates, we include estimated graduation rates in the empirical model to ensure that coefficients on our labor income variables are not biased due to differences in graduation rates across education groups.[4] Ideally, we would like to include the graduation rate as a parameter in the theoretical simulation and incorporate a separate income process for non-graduates. However, the non-graduates are too few to allow for separate estimation of income processes. Second, our life-time simulation assumes that retirement income equals a known fraction of the last value of the permanent income component at age 64. This fraction, b_j, is calibrated based on predicted income at age 64 in a given education group and the current government retirement benefits.[5] Third, we only allow households to save in the form of a riskless bond.

[4] The minimum time to complete an education is known upon admission (typically 3 or 5 years). We estimate the graduation probability, p_j, as the fraction of admitted students who have graduated within 3 years of the minimum time.

[5] At age 60, members of an unemployment insurance fund become eligible for a public early retirement scheme, which paid DKK 160,000 in year 2000. At age 67 (from 2004: age 65), individuals gain eligibilty for old age pension, which paid DKK 102,000 for singles or 74,400 for individuals living in couples in 2000. For simplicity we assume that all individuals expect to retire at age 65 and calibrate the value of b based on a pension of DKK 74,400 (DKK 58,410 after tax). Details on the Danish retirement schemes and retirement incentives may be found in Bingley, Datta-Gupta and Pedersen (2004).

IV. Data

Our analysis is based on a 10 percent random sample of the Danish population from 1984-2000. It has been made available by Statistics Denmark who gathered the information from different sources, mainly administrative registers. We augmented this data by collecting data on admissions requirements which are publicly available in book form.

Income data:

Income data stem from tax forms, registered pay from employers, and social registers. Several variables represent combinations and cross-checks across these three registers performed by Statistics Denmark who host the data. To arrive at a measure for labor income, we add all wage income, social assistance, unemployment insurance benefits, study grants and other income replacing transfers. For the self-employed, we add net income from their business. If the employee has an employer administered pension plan, the annual contribution to the plan is included in the annual income. To obtain the real values we deflate by the CPI.

We then simulate the tax system over the observation period to calculate estimated after-tax real non-financial income (total after-tax income from tax forms cannot be used because it is affected by financial income, mortgage interest deductions etc.). Taxes are levied at the state, county and municipal level. The state tax consists of three tax brackets that results in progressive taxation. In 2002, the lowest marginal tax rate was 45% and the highest marginal tax rate was 64%. Three major reforms took place in 1987, 1994 and 1998. The first two reforms reduced the highest marginal tax rate. Furthermore, the 1994 reform introduced a labor market contribution that was collected from the gross income (no deductions possible). Furthermore, public income transfers became taxable and were adjusted correspondingly. The 1998 reform mainly concerned taxation of capital income. Our tax calculations account for all changes in taxation of labor income, income transfers and self-employment income over the period. We assume that the individual lives in an average community with respect to county and municipal taxes. We disregard property taxes and tax deductions due to mortgage interest payment deductibility. Hence, we assume that the housing decision is exogenous to the educational choice.

Educational data:

We use detailed information from the Ministry of Education about educations offered and the educational history of individuals.

The Ministry of Education assigns each education a unique identifier. During our sample period 1750 different educations were offered. This universe of educations include educations at all levels (from pre-school to PhD educations). We focus on educations chosen by high school graduates. Educations vary in terms of admissions requirements, length, subject and educational institution. An education in a specific subject may thus be taken at different geographical locations and under different sets of admission requirements, since these are determined at the institutional level.

In addition to detailed education codes, we also have access to the complete educational event histories including entry, drop out and completion dates for both vocational, high school and post-high school educations. This information is systematically reported from educational institutions for all individuals over the period 1974-2001. For educations attended before 1974, information about the educations of individuals stem from surveys, and only information about the highest level of education is available. Educational institutions at the high school level are also required to supply information on GPA and subjects taken. High school GPAs are comparable across schools, since all high school students are faced with identical written exams, and the oral exams are evaluated by both the student's own teacher and a teacher from another school.

It is important to emphasize that tuition data need not be collected. In Denmark tuition is zero, and all individuals who are admitted to an education after the age of 18 are (as discussed above) eligible for a uniform study grant and a favorable loan, independent on performance. Eligibility runs for the recommended study time plus one year.

Admission requirements:

To supplement the individual level information on educations attended, we collect information on admissions requirements published annually by the Ministry of Education who coordinates the admission to most educations. Admission to most Danish post-high school educations is based exclusively on high school GPA and for some educations also on whether the applicant has sufficient mathematics skills. However, most educational institutions do reserve a small number of places for students to enter based on a combined evaluation of their GPA and their curriculum vitae. This proportion is typically 10-25%, and the entry requirement would typically be a couple of years of relevant experience or informal education plus a GPA which is close to the one that is required by individuals admitted under the ordinary quota. Educations within the fields of Education and Health admit a larger proportion of students based on such criteria. For three groups the number exceeded 50% in 2003: nurses (70%), teachers (60%) and pedagogues (70%). Two other educations also have a high fraction of students admitted in this way, that is: Physical therapists (35%) and Midwives (45%). In the empirical choice model we exclude an education from an individual's choice set if the individual would not have been able to enter the chosen education under regular admission criteria.[6]

The availability of both GPA and high school subjects at the individual level combined with GPA and mathematics requirement at the institutional level is one advantage of our analysis over earlier studies because it enables a quite accurate characterization of the educational choice set for each individual. Our current analysis is based on the admissions requirements for year 2001. To the extent that admission requirements vary over time for a given education, this will introduce some errors in characterizing the choice set for individuals who choose an education in earlier years. Admission requirements are generally very stable over time (changing by only a fraction of a grade from one year to the next). However, admission requirements have gone down during the last decade, and therefore our analysis tends to slightly overestimate the choice possibilities for individuals further back in time.[7]

Demographic and geographical data:

The 10 percent sample contains information on gender and age. We focus our analysis on males. It is likely that difficulties in controlling for exit from and reentry into the labor force due to child care will cause income process variables to be less precise measures of the income prospects of women (for example by mistaking large income changes as shocks rather than voluntary decisions). We use information about age in the income process estimation (see below). Finally, data from population registers (contained in the 10 percent sample)

[6] We exclude educations with special entry requirements. For instance educations at the Royal Academies of Musics and at the Royal Academies of Performing Arts require entry tests, and educations within the Police and Defense require physical tests.

[7] We are collecting data for admissions requirements in earlier years and will use this information in the next version of the paper.

provides information about each individual's geographical residence at age 17. We use a distance matrix made available by the Institute of Local Government Studies in Copenhagen to calculate the average distance from an individual's municipality of residence at age 17 to the educational institutions in a particular educational group for use as a control in the empirical choice model.

Our data also contains information about parental income and parents' length and type of education. We use these to construct two control variables used in both income processes and in the choice model. The first is a dummy for whether at least one parent has an education of the same length and topic. The second is the log of father's income.

Sample selection:

We focus on high school graduates only and restrict our estimations to males. Individuals are classified into education groups based on the first education attended after high school. We regard educational switches as a possibility taken into account when the educational choice is made. This means that a person who first enters medical school but graduates in science 6 years later is included in the income process estimation for Medicine since a rational high school student would consider the probability of a switch from Medicine to Science in deciding whether or not to go to medical school. For individuals attending education before 1974, information about the first education attended after high school is not available and we use instead the highest education obtained. We exclude individuals who chose an education which is shared by less than 50 people in the sample.

We form two separate samples, one for estimation of income processes and the other for estimation of an empirical choice model. The income process estimation sample consists of individual-year observations for 1984-2000 where the individual has graduated. The educational choice sample consists of individuals who start an education no later than age 25 in the period 1984-2000. The fact that we consider the same time period means that educational choices are made during the same period for which the parameters of the income process are estimated.

V. Educational Group Classification

We classify the set of post-high school educations according to admissions requirements, length, and subject. A numerical grading system is used at the high school (and post-high school) level. The scale goes from 0 to 13 and possible grades are 0, 3, 5, 6, 7, 8, 9, 10, 11, and 13. We define five categories of educations based on the high school GPA required for admission: (1) High GPA required (GPA$>$9), (2) Medium GPA required (8.2$<$GPA$\leq$ 9.0), (3) Low GPA required (6$<$GPA$\leq$8.2), (4) Completed high school education required (graduating from high school requires a GPA of at least 6), (5) No high school education required.[8]

Many educations have mathematics requirements for admission. These requirements state that the applicant must have passed either high or medium level high school mathematics courses. We classify educations into two groups according to whether any mathematics requirements are stated in the admissions rules or not.

As for length of study, we classify educations into two groups. We define short educations as those with a minimum time to complete the education of four years or less, and long educations as those requiring five years of study. We classify educations that lead to Bachelor degrees as long educations because students are admitted with the option to do a Master's degree upon completion of a Bachelor degree. With the exception of certain business degrees, the majority of students admitted to such programs do continue to the Master's level. The minimum time to complete a Bachelor degree is three years, with an extra two years required for a Master's degree. The option to leave after three years with a Bachelor's degree was introduced only in the early 1990s.

Finally, we categorize educations by nine subject categories: (1) Education, (2) Humanities, (3) Agriculture, including veterinarians, gardening etc., (4) Business, (5) Social sciences, (6) Health care, (7) Biological sciences, (8) Physical sciences, and (9) Engineering and technical fields.

We end up with $5 \times 2 \times 2 \times 9 = 180$ possible groups, 124 of which are empty and 6 of which contain less than 50 individuals and are therefore dropped. This leaves us with a total

[8] For the oldest part of the sample that is needed for income process estimation, we do not have this information. Therefore, we randomly distribute the oldest part of the sample within a given length-subject cell across these five groups assuming that the distribution is proportional to that of the younger part of the sample.

of 50 education groups. The groups are listed in Panel A of Table 1 which also shows the distribution across groups of individuals in the income process estimation sample.

The sub-division by GPA and math admission requirements allows us to rule out certain educations for each individual in the choice model. It furthermore is likely to ensure substantial variation in pecuniary payoffs across groups. That will be the case if attractive pecuniary benefits forces up admission requirements for the more attractive educations. We document the relation between income process variables and admission requirements, length, and subject below.

Note that we do not group educations by geographical location. However, in the choice model, we control for average migration or transportation cost to the educations in a given cell by including the average distance from location of residence at age 17 to the educations in a particular cell.

VI. Income Process Estimation

A. Methodology

The income process used in the empirical work is the one described for the life-time optimization problem

$$\ln Y_{i,j,t} = p_{i,j,t} + \varepsilon_{i,j,t},\ p_{i,j,t} = g_{i,j,t} + p_{i,j,t-1} + \eta_{i,j,t} \tag{9}$$

$$\varepsilon_{i,j,t} \sim N\left(0, \sigma^2_{\varepsilon,j}\right),\ \eta_{i,j,t} \sim N\left(0, \sigma^2_{\eta,j}\right) \tag{10}$$

$$cov\left(\varepsilon_{i,j,t}, \varepsilon_{i,j,s}\right) = 0, \forall t \neq s,\ cov\left(\eta_{i,j,t}, \eta_{i,j,s}\right) = 0, \forall t \neq s,\ cov\left(\varepsilon_{i,j,t}, \eta_{i,j,s}\right) = 0, \forall t, s \tag{11}$$

Growth rate estimation:
In first differences, (9) implies

$$\ln Y_{i,j,t} - \ln Y_{i,j,t-1} = g_{i,j,t} + \eta_{i,j,t} + \varepsilon_{i,j,t} - \varepsilon_{i,j,t-1}. \tag{12}$$

We allow g to differ across educations and individuals. Furthermore, for a given education j we estimate a separate growth rate for young individuals (age$\leq$ 40) and for old individuals (age$>$ 40) to allow growth to differ over the life-cycle. Thus the equation estimated is

$$\ln Y_{i,j,t} - \ln Y_{i,j,t-1} = \Sigma_{j=1}^{50} g_j D_{ij} + x'_{ij}\beta + \underbrace{\eta_{i,j,t} + \varepsilon_{i,j,t} - \varepsilon_{i,j,t-1}}_{u_{i.j,t}} \tag{13}$$

where D_{ij} is a dummy equal to one if individual i's first post-high school education was education j. We estimate (13) separately for young individuals and for old individuals.

x_{ij} includes five variables: (1) A dummy for having taken additional courses in high school in the subject of a given education j, D(HS Relevant)$_{ij}$, (2) a dummy for having at least one parent with an education of the same subject and length as education j, D(Parent With This Education)$_{ij}$, (3) high school GPA, HS GPA$_i$, (4) a dummy for medium or high level high school math courses, D(HS Math)$_i$, (5) log of father's income, ln(Father's Income)$_{it}$. We include these five variables as controls in (13), but do not allow their coefficients to vary across educations (since a lot of such estimated variation in a finite sample may be a result of realized income shocks, rather than ex ante predictable variation in e.g. the returns to GPA across educations). Above age 40 few individuals have observations for these variables, so we drop them.

The inclusion of the x-variables in the growth rate regression serves two purposes. Similar comments apply to the regression for starting income discussed below. Firstly, different individual characteristics may be associated with different income growth rates regardless of which education is chosen. Including controls for such characteristics ensures that the g_j estimates pick up differences in growth rates due to the nature of the education rather than (potentially) due to differences in the distribution of individual characteristics across educations. For example, if Physical Sciences tends to be chosen by individuals with high GPA (in high school) and high GPA is associated with high income growth rates, then omitting GPA in the above regression would lead to values of g for Physical Sciences that incorrectly predict that all individuals could obtain high g if they chose Physical Sciences. Secondly, two of our x-variables (D(HS Relevant)$_{ij}$ and D(Parent With Same Education)$_{ij}$) vary not only across individuals, i, but also across educations, j. Including these two variables in the income process leads to increased variation in predicted income growth rates across educations for a particular individual. This improves our ability to estimate the effect of income processes on educational choices. For example, an individual whose father has chosen a particular subject may have an inherited/learned ability for that subject and may thus have a higher predicted income growth rate (or level, in the regression for starting income below) in that subject than in other subjects.

The predicted mean labor income growth rate for age$\leq$ 40 for an individual i in education

j is thus

$$\widehat{g^y_{ij}} = \widehat{g^y_j} + x'_{ij}\widehat{\beta^y} \tag{14}$$

where g^y_j is the coefficient on D_{ij} in the estimation of (13) on the sub-sample of individual-year observations within the income process estimation sample where age≤ 40 and β^y is the estimate of β from that estimation.

The predicted mean labor income growth rate for age>40 for an individual in education j is

$$\widehat{g^o_j} \tag{15}$$

i.e. the coefficient on D_{ij} in the the estimation of (13) on the sub-sample of individual-year observations within the income process estimation sample where age> 40.

Estimation of variances:

The estimation of the variance of permanent and transitory income shocks is based on the residuals from (13) and follows Carroll and Samwick (1997). We use both the residuals from the young and the old sub-sample. We start by calculating 1-period growth rate residuals, $\widehat{u_{i,j,t}}$, 2-period growth rate residuals $\widehat{u_{i,j,t}} + \widehat{u_{i,j,t-1}}$, 3-period growth rate residuals $\widehat{u_{i,j,t}} + \widehat{u_{i,j,t-1}} + \widehat{u_{i,j,t-2}}$ and so on. Since

$$\ln Y_{i,j,t} - \ln Y_{i,j,t-d} = \text{constant} + \left(\eta_{i,j,t} + ... + \eta_{i,j,t-d}\right) + \left(\varepsilon_{i,j,t} - \varepsilon_{i,j,t-d}\right) \tag{16}$$

we have

$$\sigma^2\left(\widehat{u_{i,j,t}} + ... + \widehat{u_{i,j,t-d}}\right) = d\sigma^2_{\eta,j} + 2\sigma^2_{\varepsilon,j}. \tag{17}$$

Thus, to estimate $\sigma^2_{\eta,j}$, $\sigma^2_{\varepsilon,j}$ we run an OLS regression of $\left(\widehat{u_{i,j,t}} + ... + \widehat{u_{i,j,t-d}}\right)^2$ on d and 2, for each j. It is important to emphasize that the estimated variances are based on income growth rates, and thus are unaffected by possible unobserved individual effects in income levels.

To make the estimated variances robust to serial correlation in the transitory shocks up to MA(2), we impose the restriction $d \geq 3$ meaning that we only use 3-period squared growth rate residuals, 4-period squared growth rate residuals etc. Measurement error (of the commonly specified type) is indistinguishable from the transitory shock and will add to the transitory variance. However, when imposing $d \geq 3$, the estimated variance on the permanent shock will be consistently estimated unless the measurement error term is more persistent that MA(2).[9]

Estimation of log starting labor incomes:
The log starting labor income under the assumed income process is given by $\ln Y_{i,j,s} = p_{i,j,s} + \varepsilon_{i,j,s}$, where s is the year of graduation. We assume that $p_{i,j,s}$ is a linear function of the set of education dummies, observables x_{ij} from above, year dummies ($D_{i\tau}$ equal to one if individual i graduated in year τ) and an unobservable component $v_{i,j,s}$, so $p_{i,j,s} = \Sigma_{j=1}^{50}\beta_j^1 D_{ij} + x_{ij}'\delta + \Sigma_{\tau=1943}^{2000}\beta_\tau D_{i\tau} + v_{i,j,s}$. Thus, the log starting labor income is given by the following model

$$\begin{aligned} \ln Y_{i,j,s} &= p_{i,j,s} + \varepsilon_{i,j,s} \\ &= \Sigma_{j=1}^{50}\beta_j D_{ij} + x_{ij}'\delta + \Sigma_{\tau=1943}^{2000}\beta_\tau D_{i\tau} + \underbrace{v_{i,j,s} + \varepsilon_{i,j,s}}_{\text{Error term in regression}} \end{aligned} \tag{18}$$

A lot of individuals are not in our data set in the year they graduate. We can, however, use the level of their log labor income in years where data are available and use our estimates $\widehat{g_{ij}^y}$ and $\widehat{g_j^o}$ to construct estimates of what the starting salary was for such individuals. This results in several starting salary estimates for a given individual and we exploit all of them. For individuals for which we have the starting salary one can similarly construct additional estimates of the starting salary based on income data for later years. We exploit such observations too since salaries in the year after graduation may be noisy due to individuals working only part of the year, waiting for a spouse to graduate etc.

Using the estimates of the coefficients in (18), we can then predict the mean log starting labor income for each individual i, were he or she to choose education j, as

$$\widehat{\ln Y_{i,j,s}} = \Sigma_{j=1}^{50}\widehat{\beta}_j + x_{ij}'\widehat{\delta} + \Sigma_{\tau=1984}^{2000}\widehat{\beta_\tau} D_{i\tau} \tag{19}$$

where $D_{i\tau}$ is the year the education is chosen.

When estimating (18) and when estimating (13) we drop the top and bottom two percent of income observations, calculated by education group and age, in order to reduce the effect of possible measurement error.

[9] This approach is also used by Carroll and Samwick (1997).

B. Estimates

Panel A of Table 1 shows the resulting estimates of the five income process variables for each of the education groups.[10] $\widehat{\ln Y_{i,j,s}}$ and $\widehat{g_{ij}^{y}}$ has variation at the individual level within education groups, and the values shown in the panel is the average within the group across individuals in the educational choice sample. Panel B of Table 1 summarizes the distribution of income process variables across the 50 educational groups.

The first column of numbers show that inequality in predicted mean log starting salaries is substantial. The difference between the 10th and the 90th percentile corresponds to a percentage difference in the predicted level of the starting labor income of 68 percent.

Compared to the estimates for the U.S. in Carroll and Samwick (1997), the estimated time series income risk facing Danish individuals is much smaller. The median across the 50 Danish educational groups is σ_{η}^{2} =0.0046 and σ_{ε}^{2} =0.0152 (which implies standard deviations of 6.8 percent and 12.3 percent, respectively). These variances are less than half of the variances presented in Carroll and Samwick's Table 1 for their full sample. Part of the difference is likely due to higher transfers to low-income individuals in Denmark. Part of the difference is also likely to be due to lower measurement error in the Danish data which are obtained directly from tax forms, employers, and the government, rather than being based on household interviews.

Five F-tests strongly reject that either of the five income process variables are identical across the 50 educations (each with p-values less than 0.1 percent). This is encouraging for our ability to use the five income process variables to analyze the effect of income processes on educational choice.

Below we will omit four educations in the educational choice model because the variance estimates are based on less than 100 squared income growth residuals. We also show results when dropping six additional education groups that have between 100 and 500 squared income growth residuals.

It is interesting to characterize which of the four variables we used to group educations generate variation in the resulting income processes. Table 2 Panel A shows regressions run at the education group level (with 46 education groups). Not surprisingly, educations with

[10] For five of the 50 education groups we have no individuals above age 40, resulting in missing values for g^{o}. In the empirical choice model we set g^{o} for these educations equal to the sample average value of g^{o}.

stricter admission requirements have higher average predicted log starting incomes. In terms of the subjects studied, Business (Subject 4) and Social Sciences (Subject 5) stand out with high income growth rates but also high permanent income shock variances. Table 2 Panel B shows results when dropping the six additional small education groups. F-tests for the joint significance of the regressors now have substantially lower p-values in the regressions for the two income variance measures. This suggest that measurement error in the variance measures (and likely also the other income process variables) may be important and motivates showing results on educational choice for this more restricted sample of educations in addition to those for our baseline set of 46 education groups.

Finally, Panel C of Table 2 shows the potential value (for both reasons discussed above) of including individual characteristics as control variables in the starting income and growth rate regressions. This is in particular the case in the starting salary regression where all five controls have economically and statistically strong effects, all with the expected signs.

VII. Empirical Model of Educational Choice

We start by estimating a reduced form conditional logit model motivated by a first order approximation of the age 19 value functions in the five parameters of the income processes. The linearization parameters will depend on the preference parameters and we subsequently turn to a minimum distance estimation of relative risk aversion.

A. Methodology

From the life-time optimization problem it follows that the value function $V_{i,j,19}$ for individual i as of age 19 if education j is chosen will be a function of (a) the preference parameters β, γ, and (b) the five parameters characterizing the income process, i.e. $\widehat{\ln Y_{i,j,s}}$, $\widehat{g^y_{ij}}$, $\widehat{g^o_j}$, $\widehat{\sigma^2_{\eta,j}}$, and $\widehat{\sigma^2_{\varepsilon,j}}$. Consider a first order approximation of $V_{i,j,19}$ in the income process parameters $Z'_{i,j} = \left(\widehat{\ln Y_{i,,j,s}},\ \widehat{g^y_{ij}},\ \widehat{g^o_j},\ \widehat{\sigma^2_{\eta,j}},\ \widehat{\sigma^2_{\varepsilon,j}}\right)$

$$V_{i,j,19} \simeq Z'_{i,j}\alpha \tag{20}$$

where α is a 5×1 vector. α will depend on the preference parameters β and γ.

It is likely that individuals differ in their non-pecuniary utility from a given education. Let $NP_{i,j} = W'_{i,j}\theta + e_{i,j}$, where $W_{i,j}$ is a vector of observable individual characteristics and

$e_{i,j}$ is unobservable (to the econometrician, but known by the individual). Then

$$V_{i,j,19} \simeq Z'_{i,j}\alpha + W'_{i,j}\theta + e_{i,j} \tag{21}$$

Suppose $e_{i,j}$ are i.i.d. across individuals and are extreme value distributed with cumulative distribution function

$$F(e) = \Pr(e_{i,j} < e) = \exp(-\exp(-e)). \tag{22}$$

Let the dummy variable $D_{i,j}$ be one if individual i chooses education j. Then

$$D_{i,j} = \begin{cases} 1 & \text{if } V_{i,j,19} = \max(V_{i,1,19}, ..., V_{i,J,19}) \\ 0 & \text{otherwise} \end{cases}. \tag{23}$$

Under the extreme value distribution it follows that

$$\Pr(D_{i,j} = 1) = \frac{\exp\left(Z'_{i,j}\alpha + W'_{i,j}\theta\right)}{\sum_{j=1}^{J} \exp\left(Z'_{i,j}\alpha + W'_{i,j}\theta\right)} \tag{24}$$

which in turn determines the likelihood function.

The assumption that $e_{i,j}$ are i.i.d. may be too restrictive if some educations are closer substitutes than other educations due to high correlation between the unobservable non-pecuniary benefits. We try different approaches and investigate whether the i.i.d. assumption affects the main parameters of interest. The most straightforward approach is to include subject dummies to pick up subject specific non-pecuniary benefits. In addition to that, we estimate a nested logit model.

With a nested logit model, the choice set is divided into K so-called nests, within which $e_{i,j}$ are allowed to be correlated. In our case we define the nests as subject-length combinations and have $K = 15$ such that each nests (we have 9*2 such cells, 3 cells are empty). The probability that individual i chooses education j in nest k is:

$$\Pr(D_{i,j,k} = 1) = \frac{\exp(V_{i,j,19}/\lambda_k)\,(\sum_{j\in B_k} \exp(V_{i,j,19}/\lambda_k))^{\lambda_k - 1}}{\sum_{l=1}^{K} (\sum_{j\in B_l} \exp(V_{i,j,19}/\lambda_l))^{\lambda_l}} \tag{25}$$

where $(1-\lambda_k)$ is a measure of correlation (or substitutability) between educations in nest k. If $\lambda_k = 1$, there is no correlation among the unobserved components of utility for alternatives within a nest, and the choice probabilities collapse into conditional logit probabilities. If $\lambda_k < 0$, the nested logit model is not consistent with utility maximization, if $\lambda_k > 1$, it is consistent with utility maximization, whereas if $0 < \lambda_k < 1$,it may or may not be consistent with utility maximization. In addition to α and θ, we also estimate the K values of λ_k.

B. Results

Table 3 presents the estimation results for the empirical choice model based on the sample of males. Our baseline estimates are shown in regression (1). All five income process variables enter with the expected signs. An increase in the predicted mean log starting labor income increases the probability that an education is chosen, as does an increase in the expected growth rate of labor income up to age 40 or after age 40. With 46 educations in the sample, the average probability that an education is chosen is about 2 percent. The coefficient of 1.08 on mean log starting labor income implies that the probability that an education which initially has a 2 percent probability of being chosen increases to about 2.4 percent if the mean log starting labor income is increased from the 25th percentile of 11.51 to the 75th percentile of 11.69.[11] This is an economically important effect. As for the income shock variances, the variance of the permanent income shock has a strong negative effect on the probability that an education is chosen. An increase in this variance from its 25th percentile of 0.0027 to its 75th percentile of 0.0071 is estimated to decrease the probability that an education which initially has a 2 percent probability of being chosen to about 1.5 percent, suggesting that individuals strongly dislike risk from permanent income shocks. The effect of transitory income shocks is comparatively small. A comparison of the coefficients on the mean log starting labor income and the variance of the permanent income shock suggests that individuals require an increase of about 25 percent in their starting labor income in order to be willing to accept an increase in the variance of the permanent shock from its 25th percentile to its 75th percentile.[12]

In addition to the five income process variables, regression (1) includes two control variables. The fraction of enrolled students who graduate within 3 years of the minimum required completion time is included to control for differences across educations in the chance of completing the education. The average distance, measured in hundreds of kilometers, from the individuals' municipality of residence at age 17 to the educational institutions included in a particular education group, is included to control for pecuniary and/or non-pecuniary benefits of living close to one's parents. Excluding these controls has only a small effect on results. Regressions (2) to (7) provide evidence on the robustness of the baseline results. Regression (2) allows for an effect of the two dummy variables D(HS Relevant)$_{ij}$ and D(Parent With

[11] This is calculated based on the expression for the choice probability in equation (24).

[12] The required increase in the starting labor income is calculated as $\exp((0.0071\text{-}0.0027)\times 55.27/1.08) - 1$.

This Education)$_{ij}$ directly on utility (as opposed to only via income processes). This is a simple approach to account for possible non-pecuniary benefits of studying something one finds interesting. Also included is a dummy variable for whether the education is a short or long education in order to control for differences in foregone income during years spent in education and for any possible direct (positive or negative) utility effects of being a student for a longer period. While both relevant high school classes and having a parent with a given education strongly increases the probability that an education is chosen, results for the five income process parameters are similar to regression (1).

Regressions (3) and (4) attempt a more ambitious approach to controlling for possible non-pecuniary benefits and includes dummies for eight of the nine possible topics. This substantially reduces the amount of variation left in the income process variables since these vary mainly at the education level. Nonetheless, the five income process variables retain their significance and economic importance, though the coefficients do change somewhat.

Regression (5) addresses the issue of measurement error (estimation error) in the income variables used as regressors. It increases the required number of squared income growth residuals, $\left(\widehat{u_{i,j,t}} + ... + \widehat{u_{i,j,t-d}}\right)^2$, used for estimating the income variances from at least 100 per education in the baseline estimation to 500. This increases the effects of both the predicted mean log starting labor income and the income variances.

Regression (6) and (7) uses an alternative measure of the predicted mean log starting labor income that allows the effects of control variables in the x-vector in equation (18) to vary across educations. This strengthens the effect of starting salary but this result is somewhat sensitive to the exact choice of the variables included in the x-vector.

Overall the results in Table 3 strongly confirms the theoretical predictions that both mean incomes and income risk affect individuals' choice of post-high school education. Finally, to allow for correlation of the non-pecuniary benefits among groups of educations, we also estimate a nested logit model with 15 nests (subject*length). The results are presented in Table 4. The coefficients of main interest changes slightly, but the main conclusions are not affected. Four of the 15 $\lambda's$ are above unity and none are below zero, which is reassuring since we would like to interpret the results as consistent with utility maximization.

VIII. Risk Aversion Estimation

We estimate the coefficient of relative risk aversion γ using a simulated minimum distance (SMD) estimator. The basic idea is related to that of Gourinchas and Parker (2002) and Cagetti (2003). Gourinchas and Parker (2002) estimate preference parameters by comparison of simulated and actual consumption-age profiles, and estimate the coefficient of relative risk aversion to be between 0.5 or 1.4 depending on the weighting scheme. The discount factor is estimated to be 0.96. Cagetti (2003) compares simulated and actual wealth paths over the life cycle, and arrive at risk aversion estimates around 3 to 4 combined with discount factors around 0.8-0.9 depending on which education group is considered.

We are also concerned with preference parameters but use data on educational choices rather than consumption or wealth data. We estimate risk aversion as the value that makes the linearized value function as of age 19 best fit the coefficients of the empirical choice model of educational choice presented above. We do not attempt to estimate individuals' discount factor β because the linearization coefficients are relatively insensitive to this parameter. We confirmed that our estimates of γ are not very sensitive to the assumed value of β (within a reasonable range for β).

We apply a solution method similar to that of Gourinchas and Parker (2002), where we exploit the homogeneity of utility in the permanent income component. We maximize the expected discounted life-time utility by grid search over cash on hand and consumption and linearly interpolate the value function between the grid points.

A. Linearizing the age 19 value function

For simplicity we drop the i, j subscript in the derivations below. Let $z = \left(g^y,\ g^o,\ \sigma_\eta^2,\ \sigma_\varepsilon^2\right)'$. Furthermore, let $X_t = W_t + Y_t$ denote total cash on hand at time t and $P_t = \exp(p_t)$ denote the permanent component of Y_t. Under the assumed income process, the value function is homogeneous of degree $(1-\gamma)$, see e.g. Haliassos and Michaelides (2002). The value function at time t can therefore be rewritten as follows

$$\begin{aligned} V_t\left(X_t, P_t; z, \beta, \gamma\right) &= P_t^{1-\gamma} V_t\left(X_t/P_t, 1; z, \beta, \gamma\right) \\ &= P_t^{1-\gamma} \nu_t\left(x_t; z, \beta, \gamma\right) \\ &= e^{(1-\gamma)\ln P_t} \nu_t\left(x_t; z, \beta, \gamma\right) \qquad (26) \end{aligned}$$

where $x_t = X_t/P_t$ and $\nu_t(x_t; z, \beta, \gamma) = V_t(X_t/P_t, 1; z, \beta, \gamma)$. Linearizing $V_t(X_t, P_t; z, \beta, \gamma)$ in $(\ln P_t, z)$ around the median values $(\ln P_t^*, z^*)$ gives

$$\begin{aligned} V_t(X_t, P_t; \theta, \beta, \gamma) &= e^{(1-\gamma)\ln P_t}\nu_t(x_t; z, \beta, \gamma) \\ &\simeq e^{(1-\gamma)\ln P_t^*}\nu_t(x_t; z^*, \beta, \gamma) \\ &\quad + (1-\gamma)\, e^{(1-\gamma)\ln P_t^*}\nu_t(x_t; z^*, \beta, \gamma)(\ln P_t - \ln P_t^*) \\ &\quad + e^{(1-\gamma)\ln P_t^*}\frac{d\nu_t(x_t; z^*, \beta, \gamma)'}{dz}(z - z^*) \\ &= (P_t^*)^{1-\gamma}\nu_t(x_t; z^*, \beta, \gamma) \\ &\quad + (1-\gamma)(P_t^*)^{1-\gamma}\nu_t(x_t; z^*, \beta, \gamma)(\ln P_t - \ln P_t^*) \\ &\quad + (P_t^*)^{1-\gamma}\frac{d\nu_t(x_t; z^*, \beta, \gamma)'}{dz}(z - z^*) \end{aligned} \tag{27}$$

$$= \text{constant} + \left[\underbrace{(1-\gamma)(P_t^*)^{1-\gamma}\nu_t(x_t; z^*, \beta, \gamma)}_{\alpha_1(\beta,\gamma)_{1x1}} \quad \underbrace{(P_t^*)^{1-\gamma}\frac{d\nu_t(x_t; z^*, \beta, \gamma)'}{dz}}_{\alpha_2(\beta,\gamma)_{1x4}}\right]\left[\begin{array}{c}\ln P_t - \ln P_t^* \\ z - z^*\end{array}\right]. \tag{28}$$

The empirical choice model is only identified up to scale. Thus we focus on the ratio

$$\frac{\alpha_2(\beta, \gamma)}{\alpha_1(\beta, \gamma)} = \frac{\frac{d\nu_t(x_t; z^*, \beta, \gamma)'}{dz}}{(1-\gamma)\nu_t(x_t; z^*, \beta, \gamma)} \tag{29}$$

with $t = 19$. $\nu_t(x_t; z^*, \beta, \gamma)$ is calculated as part of the simulation. $\frac{d\nu_t(x_t; z^*, \beta, \gamma)}{dz}$ is calculated (also based on the simulation) in step 2 of the minimum distance approach outlined below. Denote $\frac{\alpha_2(\beta,\gamma)}{\alpha_1(\beta,\gamma)}$ by $\alpha_2^{**}(\gamma)$, where we for simplicity suppress the dependence of α_2^{**} on β (since we do not estimate β).

Prior to choosing an education at $t = 19$ the individual does not know what the realized value of the first $\ln P_t$ will be. Thus, the $\ln P_t$ in V_{19} is the predicted value from equation (19), i.e. $\widehat{\ln Y_{i,j,s}}$ in our earlier notation.

B. Minimum distance estimation of relative risk aversion

Our minimum distance approach to estimate the coefficient of relative risk aversion γ can be summarized in the following steps:

1. Pick a value of γ.

2. For this γ:

   (a) Perform the life-time simulation nine times to enable a simple approximation of $\frac{d\nu_t(x_t;z^*,\beta,\gamma)}{dz}$. First perform the life-time simulation at the median values of the parameters $z = \left(g^y,\ g^o,\ \sigma_\eta^2,\ \sigma_\varepsilon^2\right)'$ of the income process. Then perform it eight more times, each time with one of the four parameters increased to its 75$^{\text{th}}$ percentile or lowered to its 25$^{\text{th}}$ percentile. For each combination, save the scaled value function $\nu_{19}(x_t, z, \beta, \gamma)$. Panel B of Table 1 shows the percentiles used, calculated in the educational choice sample.

   (b) Approximate $\frac{d\nu_t(x_t;z^*,\beta,\gamma)}{dz}$ as

$$\begin{bmatrix} \frac{\nu_{19}(x_t,z,\beta,\gamma) \text{ at } g^y_{75th} - \nu_{19}(x_t,z,\beta,\gamma) \text{ at } g^y_{25th}}{g^y_{75th} - g^y_{25th}} \\ \frac{\nu_{19}(x_t,z,\beta,\gamma) \text{ at } g^o_{75th} - \nu_{19}(x_t,z,\beta,\gamma) \text{ at } g^o_{25th}}{g^o_{75th} - g^o_{25th}} \\ \frac{\nu_{19}(x_t,z,\beta,\gamma) \text{ at } \sigma^2_{\eta,75th} - \nu_{19}(x_t,z,\beta,\gamma) \text{ at } \sigma^2_{\eta,25th}}{\sigma^2_{\eta,75th} - \sigma^2_{\eta,25th}} \\ \frac{\nu_{19}(x_t,z,\beta,\gamma) \text{ at } \sigma^2_{\varepsilon,75th} - \nu_{19}(x_t,z,\beta,\gamma) \text{ at } \sigma^2_{\varepsilon,25th}}{\sigma^2_{\varepsilon,75th} - \sigma^2_{\varepsilon,25th}} \end{bmatrix} \tag{30}$$

   and calculate $\alpha_2^{**}(\gamma) = \frac{\alpha_2(\beta,\gamma)}{\alpha_1(\beta,\gamma)} = \frac{\frac{d\nu_t(x_t;z^*,\beta,\gamma)}{dz}'}{(1-\gamma)\nu_t(x_t;z^*,\beta,\gamma)}$.

3. Repeat steps 1-2 for L different values of γ.

4. The empirical choice model provides an estimate of $\alpha_2^{**}(\gamma_0)$ where γ_0 is the true value of γ. Denote this estimate by $\widehat{\alpha_2^{**}}$. Let $g\left(\widehat{\alpha_2^{**}}, \gamma\right) = \left[\widehat{\alpha_2^{**}} - \alpha_2^{**}(\gamma)\right]$. Then our minimum distance estimator of γ is defined as

$$\widehat{\gamma} = \underset{\gamma \in \{\gamma_1,\dots,\gamma_L\}}{\arg\min}\ g\left(\widehat{\alpha_2^{**}}, \gamma\right)' W g\left(\widehat{\alpha_2^{**}}, \gamma\right) \tag{31}$$

   for a positive definite matrix weighting matrix W.

5. Assuming that the set $\gamma_1, ..., \gamma_L$ includes a value sufficiently close to the true γ, and that the regularity conditions set forward in Newey and McFadden (1994) are satisfied, then $\widehat{\gamma}$ is consistent and asymptotically normal. The asymptotic distribution can be derived by expanding the first-order condition around the true value for γ. The first order condition for γ is

$$\left[\nabla_\gamma g\left(\widehat{\alpha_2^{**}}, \widehat{\gamma}\right)\right]' W g\left(\widehat{\alpha_2^{**}}, \widehat{\gamma}\right) = 0 \tag{32}$$

Expanding $g\left(\widehat{\alpha_2^{**}}, \widehat{\gamma}\right)$ around γ_0 implies

$$\left[\nabla_\gamma g\left(\widehat{\alpha_2^{**}}, \widehat{\gamma}\right)\right]' W \left[g\left(\widehat{\alpha_2^{**}}, \gamma_0\right) + \nabla_\gamma g\left(\widehat{\alpha_2^{**}}, \overline{\gamma}\right)(\widehat{\gamma} - \gamma_0)\right] = 0 \tag{33}$$

and thus

$$\sqrt{I}\left(\widehat{\gamma} - \gamma_0\right) = -\left(\left[\nabla_\gamma g\left(\widehat{\alpha_2^{**}}, \widehat{\gamma}\right)\right]' W \nabla_\gamma g\left(\widehat{\alpha_2^{**}}, \overline{\gamma}\right)\right)^{-1}\left[\nabla_\gamma g\left(\widehat{\alpha_2^{**}}, \widehat{\gamma}\right)\right]' W \sqrt{I} g\left(\widehat{\alpha_2^{**}}, \gamma_0\right) \tag{34}$$

where $\overline{\gamma}$ lies between $\widehat{\gamma}$ and γ_0. Furthermore, a first-order expansion of $g\left(\widehat{\alpha_2^{**}}, \gamma_0\right)$ in α_2^{**} around the true value α_{20}^{**} implies

$$g\left(\widehat{\alpha_2^{**}}, \gamma_0\right) = \underbrace{g\left(\alpha_{20}^{**}, \gamma_0\right)}_{=0} + \nabla_{\alpha_{20}^{**}} g\left(\overline{\alpha_{20}^{**}}, \gamma_0\right)\left(\widehat{\alpha_2^{**}} - \alpha_{20}^{**}\right) \tag{35}$$

and thus

$$\sqrt{I}\left(g\left(\widehat{\alpha_2^{**}}, \gamma_0\right) - g\left(\alpha_{20}^{**}, \gamma_0\right)\right) = \nabla_{\alpha_{20}^{**}} g\left(\overline{\alpha_{20}^{**}}, \gamma_0\right)\sqrt{I}\left(\widehat{\alpha_2^{**}} - \alpha_{20}^{**}\right). \tag{36}$$

Assume $\sqrt{I}\left(\widehat{\alpha_2^{**}} - \alpha_{20}^{**}\right) \xrightarrow{d} N\left(0, V_\alpha\right)$ and let g_γ denote $\nabla_\gamma g\left(\alpha_{20}^{**}, \gamma_0\right)$, and g_α denote $\nabla_{\alpha_{20}^{**}} g\left(\alpha_{20}^{**}, \gamma_0\right)$. Then

$$\sqrt{I}\left(\widehat{\gamma} - \gamma_0\right) \xrightarrow{d} N\left(0, V_\gamma\right) \tag{37}$$

with

$$V_\gamma = \left(g_\gamma' W g_\gamma\right)^{-1} g_\gamma' W g_\alpha' V_\alpha g_\alpha W' g_\gamma \left(g_\gamma' W g_\gamma\right)^{-1}. \tag{38}$$

For simplicity, we use $W = I$, so

$$V_\gamma = \left(g_\gamma' g_\gamma\right)^{-1} g_\gamma' g_\alpha' V_\alpha g_\alpha g_\gamma \left(g_\gamma' g_\gamma\right)^{-1}. \tag{39}$$

6. $g\left(\alpha_{20}^{**}, \gamma_0\right) = \left[\alpha_{20}^{**} - \alpha_2^{**}\left(\gamma_0\right)\right]$ implies that

$$g_\alpha = I \tag{40}$$

$$g_\gamma = -\nabla_\gamma \alpha_2^{**}\left(\gamma_0\right). \tag{41}$$

Let l^{sol} denote the value of l for which γ_l (approximately) solves the first order condition. We then estimate g_γ as $-\left[\alpha_2^{**}\left(\gamma_{l^{sol}}\right) - \alpha_2^{**}\left(\gamma_{l^{sol}-1}\right)\right] / \left(\gamma_{l^{sol}} - \gamma_{l^{sol}-1}\right)$.

V_α is estimated from the estimated covariance matrix in the conditional logit estimation, using the delta method. The derivations can be augmented to account for estimation uncertainty in the income variables in the conditional logit estimation. Since this had little effect on the resulting standard error of the γ estimate we omit this complication.

C. Results

Table 5 shows the results of the minimum distance estimation. Panel A assumes a real interest on savings of 1 percent and a real interest rate on borrowing of 7.5 percent. The spread of 6.5 percent is consistent with data from the Central Bank of Denmark and Finanstilsynet (a government agency) on average interest rates on all types of bank loans to individuals in the 1990s. This spread is however not representative of the spread on total household borrowing since a lot of mortgages are issued by mortgage companies. Mortgage rates on new mortgages are currently about 2.5 percent above the interest rate on short government debt. As an alternative we also show results for an overall spread of 3 percent between borrowing and lending rates. The top part of each column shows the scaled linearization coefficients from the life-time simulation, $\alpha_2^{**}(\gamma)$, while the scaled coefficients from the empirical choice model $\widehat{\alpha_2^{**}}$ is given at the bottom of each column for reference.

The empirical choice model suggested that individuals' educational choices are affected much more strongly by the variance of the permanent labor income shock σ_η^2 than the variance of the transitory labor income shock σ_ε^2. The life-time simulation shows that this is consistent with rational behavior since the scaled linearization coefficients on σ_η^2 in the simulation is much larger than the scaled linearization coefficient on σ_ε^2. The scaled coefficients on the two expected income growth rates are generally lower in the simulation (for the various values of γ considered) than in the empirical choice model.

Of the four scaled linearization coefficients, the coefficient on σ_η^2 not surprisingly depends most strongly on the value of γ. This implies than the minimum distance estimation will identify γ mainly based on that coefficient. Consistent with this, the minimum distance estimate of γ based on $\beta = 0.98$ and an interest spread of 3 percent equals 4, the value that makes the scaled linearization coefficient on σ_η^2 closest to the value from the empirical choice model. The standard error on the risk aversion estimate is 0.5. This standard error accounts for the estimation uncertainty in the coefficients from the empirical choice model. The results are fairly insensitive to the assumed value of β (results not shown). Estimation of β could be done based on the life-cycle of consumption and wealth accumulation, but that is outside the scope of this paper (and the availability of data). Worsening households' borrowing terms in the high hand side of the table makes the value function a bit less sensitive to the two income shock variances, and therefore increases the risk aversion estimate to 5. This finding is due to the fact that worsened borrowing conditions increase precautionary savings, and consequently decreases the sensitivity of the value function to the income shock variances.

The reported risk aversion estimates are based on the baseline specification for males (model (1) in Table 3). If instead we used model (2) or (3), which controls for preference variables, the estimate would be higher but still less than 10. Model (5) which drops the six additional small education groups also leads to risk aversion estimates around 5.

IX. Conclusion

Labor income risk, particularly the variance of permanent income shocks, has a statistically significant and economically important effect on educational choices.

Our analysis documents this using a large Danish panel data set on labor incomes and educational choices. The Danish data (or similar data for another European country) are ideally suited for the study because admissions to post-high school educations is based largely on high school GPAs and high school math courses passed. This enables a fairly accurate characterization of the educational choice set for a particular individual. Furthermore, the Danish data contains a large number of individuals, allowing for a fairly detailed grouping of educations.

Income processes estimated for each of 50 education groups suggest quite large and significant differences in income risk across educations. An empirical choice model documents the effect of labor income risk, and of predicted mean log starting labor incomes and predicted mean labor income growth rates, on high school graduates' choice of post-high school education. Our baseline results suggest that an increase in the variance of permanent labor income shocks from its 25th percentile of 0.0027 to its 75th percentile of 0.0071 decreases the probability that an education which initially has a 2 percent probability of being chosen to about 1.5 percent. Furthermore, a comparison of the coefficients on mean log starting labor income and the variance of the permanent income shock suggests that individuals require an increase of about 25 percent in their starting labor income in order to be willing to accept an increase in the variance of the permanent shock from its 25th percentile to its 75th percentile, suggesting that individuals strongly dislike risk from permanent income shocks. The effect of transitory income shocks is comparatively small, as theory suggests should be the case.

A life-time utility maximization problem is then solved using numerical simulation methods in order to determine which value of relative risk aversion is implied by the empirical choice model estimates. The maximization problem allows consumption to be chosen endogenously and the robustness of results to different assumptions about borrowing and lending rates is analyzed. The resulting estimate of relative risk aversion is around 5 with a standard error of about 1.

References

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

[2] Belzil, C., and J. Hansen (2004), Earnings Dispersion, Risk Aversion and Education, Forthcoming in Research in Labor Economics.

[3] Berger, M. (1988),Predicted Future Earnings and Choice of College Major, Industrial and Labor Relations Review, 41: 418-429.

[4] Bingley, P., N. Datta Gupta and P. J. Pedersen (2004), The Impact of Incentives on Retirement Age in Denmark, Ch. 3. in J. Gruber and D. A. Wise, (eds.)Social Retirement Programs and Retirement around the World, NBER: U Chicago Press.

[5] Boskin, M. J. (1974), A Conditional Logit Model of Occupational Choice, Journal of Political Economy, 82: 389-398.

[6] Cagetti, M. (2003), Wealth Accumulation over the Life Cycle and Precautionary Savings, Journal of Business and Economics Statistics 21(3): 339-353.

[7] Carneiro, P., K. Hansen, J. J. Heckman (2003), Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice, International Economic Review 44(2) 361-422.

[8] Carroll, C. and Samwick (1997), The Nature of Precautionary Wealth, Journal of Monetary Economics 40: 41-71.

[9] Christiansen C., J. S. Joensen and H. S. Nielsen (2004), The Risk-Return Trade-off in Human Capital Investment, manuscript, University of Aarhus, Denmark.

[10] Danø, A. M., M. Ejrnæs and L. Husted (2000), How does the Reform of Early Retirement Affect Retirement Age? (In Danish: Hvordan påvirker efterlønsreformen de ældres tilbagetrækningsalder?), Nationaløkonomisk Tidsskrift 138(2):205-221.

[11] Diaz-Serrano, L., J. Hartog and H. S. Nielsen (2003), Compensating Wage Differentials for Schooling Risk in Denmark, IZA DP No. 963, IZA, Bonn.

[12] Diaz-Serrano, L and J. Hartog (2002), Earnings Risk and Demand for Higher Education: A Cross-Section Test for Spain, IZA DP No. 641, Bonn.

[13] Fuchs-Schündeln, N. and M. Schündeln (2003), Precautionary Savings and Self-Selection: Evidence From the German Reunification 'Experiment', manuscript, Yale University.

[14] Gourinchas and Parker (2002), Consumption over the Life Cycle, Econometrica 70: 47-89.

[15] Gourieroux; C. and A. Montfort (1996), Simulation-Based Econometric Methods, CORE lectures: Oxford University Press.

[16] Guiso, L., M. Haliassos and T. Jappelli (2002), Household Portfolios, MIT Press.

[17] Hartog, J. and W. Vijverberg (2002), Do Wages Really Compensate for Risk Aversion and Skewness Affection? IZA DP No. 426, IZA, Bonn.

[18] Hornstrup, M. and B. Madsen (2000), Wealth in Denmark (In Danish: Formuerne i Danmark), Report, Arbejderbevægelsens Erhvervsråd, Denmark.

[19] Montgomery, M. (2002), A Nested Logit Model of the Choice of a Graduate Business Schools, Economics of Education Review 21: 471-480.

[20] Rosen, S. and R. Willis (1979), Education and Self-Selection, Journal of Political Economy 87(5), S7-S36.

[21] Saks, R. E. and S. H. Shore (2004), Risk and Career Choice, working paper, Wharton.

Table 1. Summary Statistics for Predicted Income Process Parameters
Panel A. Statistics By Education in Income Process Estimation Sample

Educational Group	Mean Log Starting Income	Mean Income Growth Rate, Age≤ 40	Mean Income Growth Rate, Age> 40	Variance of Permanent Income Shock, σ_η^2	Variance of Transitory Income Shock, σ_ϵ^2	Number of Individuals in Income Process Estimation
Required GPA > 9						
Math Requirements						
Long Educations						
Subject 3	11.92	0.0630	0.0265	0.0028	0.0236	98
Subject 6	11.87	0.0752	0.0373	0.0071	0.0126	1,144
No Math Requirements						
Long Educations						
Subject 2	11.64	0.0587	0.0241	0.0034	0.0225	201
Subject 5	11.62	0.0701	0.0418		0.0107	757
Subject 9	11.70	0.0590	0.0352	0.0066	0.0132	237
Required GPA $> 8.2, \leq 9$						
Math Requirements						
Long Educations						
Subject 6	11.77	0.0800	0.0449	0.0075	0.0172	379
Subject 7	11.61	0.0693	0.0356	0.0012	0.0255	257
No Math Requirements						
Long Educations						
Subject 2	11.56	0.0696	0.0343	0.0069	0.0140	816
Subject 4	11.30	0.0918	0.0363	-0.0175	0.1013	80
Subject 5	11.61	0.0883	0.0368	0.0059	0.0225	1,190
Subject 7	11.64	0.0683	0.0444	0.0015	0.0094	84
Subject 9	11.69	0.0636	0.0375	0.0015	0.0223	153
Short Educations						
Subject 1	11.66	0.0469	0.0327	0.0037	0.0067	404
Subject 5	11.62	0.0573	0.0302	-0.0013	0.0064	17
Subject 6	11.86	0.0269	0.0481	0.0248	-0.0263	51

Educational Group	Mean Log Starting Income	Mean Income Growth Rate, Age$\leq$ 40	Mean Income Growth Rate, Age$>$ 40	Variance of Permanent Income Shock, σ_η^2	Variance of Transitory Income Shock, σ_ϵ^2	Number of Individuals in Income Process Estimation
Required GPA $> 6, \leq 8.2$						
Math Requirements						
Long Educations						
Subject 4	11.23	0.1617	0.0363	0.0016	0.0448	47
Subject 5	11.72	0.0871	0.0448	0.0062	0.0091	378
Subject 6	11.77	0.0731	0.0590	0.0057	0.0126	142
Subject 7	11.61	0.0734	0.0411	-0.0004	0.0224	152
Subject 9	11.74	0.0728	0.0489	0.0056	0.0097	2,257
No Math Requirements						
Long Educations						
Subject 2	11.63	0.0613	0.0297	0.0032	0.0165	482
Subject 4	11.37	0.1028	0.0556	0.0113	0.0205	1,396
Subject 8	11.17	0.0786	0.0363	0.0934	-0.0218	55
Short Educations						
Subject 1	11.66	0.0485	0.0234	0.0022	0.0059	488
Subject 2	11.59	0.0552	0.0336	0.0042	0.0085	98
Subject 5	11.79	0.0311	0.0255	0.0134	-0.0138	47
Subject 6	11.61	0.0535	0.0267	-0.0005	0.0106	32
Required GPA $= 6$ (HS Diploma)						
Math Requirements						
Long Educations						
Subject 3	11.67	0.0740	0.0452	0.0045	0.0213	346
Subject 5	11.59	0.0989	0.0427	0.0137	0.0049	555
Subject 7	11.55	0.0831	0.0353	0.0013	0.0230	355
Subject 8	11.61	0.0811	0.0365	0.0055	0.0158	1,647
Subject 9	11.77	0.0638	0.0328	0.004	0.0136	111
Short Educations						
Subject 5	11.72	0.1090	0.0363	0.0123	0.0018	274
Subject 9	11.69	0.0694	0.0392	0.0053	0.0101	2,186

Educational Group	Mean Log Starting Income	Mean Income Growth Rate, Age$\leq$ 40	Mean Income Growth Rate, Age$>$ 40	Variance of Permanent Income Shock, σ_η^2	Variance of Transitory Income Shock, σ_ϵ^2	Number of Individuals in Income Process Estimation
No Math Requirements						
Long Educations						
Subject 2	11.49	0.0731	0.0332	0.0027	0.0240	1,093
Subject 4	11.27	0.1169	0.0564	0.0123	0.0335	835
Subject 5	11.45	0.0895	0.0263	0.0028	0.0324	250
Subject 9	11.79	0.1145	0.0363	0.0244	0.0060	159
Short Educations						
Subject 1	11.64	0.0504	0.0249	0.003	0.0056	1,809
Subject 2	11.62	0.0505	0.0357	0.0047	0.0212	276
Subject 3	11.63	0.0631	0.0204	0.0045	0.0156	89
Subject 4	11.34	0.0854	0.0438	0.0195	0.0049	247
Subject 5	11.51	0.0791	0.0471	0.0014	0.0238	335
Subject 6	11.61	0.0536	0.0431	0.0038	0.0192	180
Subject 9	11.43	0.0677	0.0206	0.0032	0.0356	442
No GPA Requirement						
No Math Requirements						
Short Educations						
Subject 3	11.10	0.0855	0.0492	0.0056	0.0569	913
Subject 4	11.53	0.0671	0.0409	0.0047	0.0147	1,547
Subject 5	10.92	0.0864	0.0258	0.0084	0.0451	854
Subject 6	11.60	0.0450	-0.0012	0.0015	0.0132	46
Subject 9	11.25	0.0751	0.0386	0.0031	0.0431	1,437
Total						27425

Note: The Subjects studied are denoted with numbers: (1) Education, (2) Humanities, (3) Agriculture, (4) Business, (5) Social Sciences, (6) Health Care, (7) Biological Sciences, (8) Physical Sciences, (9) Engineering and Technical Fields. The table is based on predicted values of income process variables for all 11170 individuals in the educational choice sample. The last column indicates the number of individuals in the income process estimation sample used to estimate the income process variables.

Panel B. Overall Percentiles in Educational Choice Sample

Income Variable	Percentile 10th	 25th	 Median	 75th	 90th
Mean Log Starting Income	11.26	11.51	11.61	11.69	11.78
Mean Income Growth Rate, Age$\leq$ 40	0.039	0.053	0.071	0.089	0.104
Mean Income Growth Rate, Age$>$ 40	0.025	0.030	0.036	0.043	0.048
Variance of Transitory Income Shocks, σ_ϵ^2	0.0049	0.0091	0.0152	0.0230	0.0394
Variance of Permanent Income Shocks, σ_η^2	0.0013	0.0027	0.0046	0.0071	0.0136

Table 2. Determinants of Variation in Income Processes
Panel A. Effect of Four Variables Used to Categorize Educations on Income Processes

	(1) Log Starting Income		(2) Income Growth Rate, Age$\leq$ 40		(3) Income Growth Rate, Age$>$ 40		(4) Variance of Trans. Shock		(5) Variance of Perm. Shock	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
Omitted: D(No GPA Requirement)										
D(Required GPA = 6 (HS Diploma))	0.277	3.6	0.005	0.4	0.005	0.8	-0.021	-3.2	0.004	1.4
D(Required GPA $> 6, \leq 8.2$)	0.344	3.9	-0.009	-0.7	0.008	1.2	-0.028	-3.7	0.003	0.9
D(Required GPA $> 8.2, \leq 9$)	0.384	4.0	-0.005	-0.4	0.108	1.5	-0.031	-3.7	0.006	1.7
D(Required GPA > 9)	0.441	4.2	-0.018	-1.2	0.001	0.1	-0.028	-3.1	0.004	0.9
Omitted: D(No Math Requirements)										
D(Math Requirements)	0.129	2.2	-0.002	-0.3	0.005	1.2	-0.005	-1.0	0.000	0.0
Omitted: D(Short Education)										
D(Long Education)	-0.040	-0.7	0.014	1.6	0.003	0.8	0.011	2.2	-0.001	-0.5
Omitted: D(Subject 1)										
D(Subject 2)	-0.052	-0.5	0.005	0.3	0.004	0.5	0.004	0.4	0.002	0.6
D(Subject 3)	-0.034	-0.3	0.015	0.9	0.009	1.0	0.011	1.1	0.003	0.7
D(Subject 4)	-0.148	-1.3	0.028	1.7	0.024	2.8	-0.002	-0.2	0.011	2.4
D(Subject 5)	-0.085	-0.9	0.029	2.0	0.008	1.1	0.000	0.0	0.006	1.6
D(Subject 6)	0.057	0.6	-0.006	-0.4	0.008	1.0	-0.003	-0.3	0.005	1.3
D(Subject 7)	-0.122	-1.1	0.011	0.6	0.004	0.5	0.008	0.7	-0.001	-0.3
D(Subject 8)	-0.077	-0.4	0.005	0.2	0.004	0.3	-0.002	-0.1	0.040	0.5
D(Subject 9)	0.004	0.0	0.010	0.7	0.008	1.0	0.003	0.3	0.005	1.2
Constant	11.232	100.6	0.042	2.6	0.101	1.3	0.027	2.7	-0.001	-0.1
P-value for F-test for all Regressors	0.0007		0.0409		0.2114		0.0489		0.4271	
N	46		46		46		46		46	

Note: The Subjects studied are denoted with numbers: (1) Education, (2) Humanities, (3) Agriculture, (4) Business, (5) Social Sciences, (6) Health Care, (7) Biological Sciences, (8) Physical Sciences, (9) Engineering and Technical Fields. The regressions are run at the education group level.

Panel B. Effect of Four Variables Used to Categorize Educations on Income Processes, Small Education Groups Dropped

	(1) Log Starting Income		(2) Income Growth Rate, Age $\leq$ 40		(3) Income Growth Rate, Age$>$ 40		(4) Variance of Trans. Shock		(5) Variance of Perm. Shock	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
Omitted: D(No GPA Requirement)										
D(Required GPA = 6 (HS Diploma))	0.320	4.1	0.005	0.4	-0.004	-0.7	-0.023	-3.9	0.002	1.0
D(Required GPA $> 6, \leq 8.2$)	0.424	4.6	-0.005	-0.3	-0.001	-0.1	-0.032	-4.7	0.002	1.0
D(Required GPA $> 8.2, \leq 9$)	0.448	4.3	-0.001	-0.1	-0.003	-0.4	-0.029	-3.9	0.003	1.3
D(Required GPA > 9)	0.532	5.0	-0.015	-0.9	-0.010	-1.4	-0.033	-4.3	0.003	1.3
Omitted: D(No Math Requirements)										
D(Math Requirements)	0.192	2.9	-0.004	-0.4	0.004	0.8	-0.013	-2.7	0.003	1.3
Omitted: D(Short Education)										
D(Long Education)	-0.081	-1.3	0.013	1.2	0.005	1.2	0.013	2.7	-0.002	-1.2
Omitted: D(Subject 1)										
D(Subject 2)	-0.029	-0.3	0.005	0.3	0.003	0.4	0.004	0.6	0.002	1.0
D(Subject 3)	-0.027	-0.2	0.018	1.0	0.006	0.8	0.015	1.9	0.002	0.7
D(Subject 4)	-0.107	-1.0	0.030	1.6	0.019	2.6	-0.003	-0.3	0.011	3.9
D(Subject 5)	-0.078	-0.8	0.028	1.7	0.006	0.9	0.003	0.4	0.006	2.2
D(Subject 6)	-0.017	-0.2	-0.002	-0.1	0.014	2.0	0.011	1.4	0.003	0.9
D(Subject 7)	-0.179	-1.5	0.011	0.5	0.001	0.2	0.028	2.0	-0.003	-0.8
D(Subject 8)	-0.080	-0.5	0.010	0.4	0.002	0.2	0.005	0.4	0.003	0.6
D(Subject 9)	-0.022	-0.2	0.010	0.6	0.006	0.9	0.009	1.2	0.002	0.7
Constant	11.147	102.4	0.042	2.3	0.021	2.9	0.035	4.4	0.000	0.0
P-value for F-test for all Regressors	0.0002		0.2964		0.0680		0.0031		0.0255	
N	40		40		40		40		40	

Note: The Subjects studied are denoted with numbers: (1) Education, (2) Humanities, (3) Agriculture, (4) Business, (5) Social Sciences, (6) Health Care, (7) Biological Sciences, (8) Physical Sciences, (9) Engineering and Technical Fields. The regressions are run at the education group level.

Panel C. Effect of Individual Characteristics on Log Starting Incomes and Mean Income Growth Rate, Age$\leq$ 40

	Log Starting Income		Income Growth Rate, Age$\leq$ 40	
	Coef.	t	Coef.	t
D(HS Relevant)$_{i,j}$	0.059	20.5	0.0033	2.1
D(Parent With This Education)$_{i,j}$	0.086	13.9	-0.014	-4.4
HS GPA$_i$	0.021	13.5	0.0071	9.3
D(HS Math)$_i$	0.205	55.8	0.0031	1.7
ln(Father's Income)$_{it}$	0.020	7.6	0.0048	3.7
N	174,991		110,797	

Note: The regressions for this panel do not include the 50 education group dummies. Dummies for missing parental information and missing high school GPA are included, and dummies for cohort (year of graduation) are included in the log starting income regression.

Table 3. Conditional Logit Model of Educational Choice

	(1) Baseline		(2) Additional regressors		(3) Additional regressors		(4) Additional regressors		(5) Drop small groups		(6) Alternative $\ln \widehat{Y_{i,j,s}}$		(7) Alt. $\ln \widehat{Y_{i,j,s}}$ and drop small groups	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
$\ln \widehat{Y_{i,j,s}}$	1.08	8.9	0.72	5.1	1.13	6.0	0.68	2.7	1.42	10.1	2.31	24.9	2.69	24.9
$\widehat{g_{ij}^y}$	18.45	18.6	18.23	16.5	24.63	13.5	40.47	20.2	27.82	21.7	15.06	14.6	27.99	21.8
$\widehat{g_j^o}$	37.41	23.4	32.55	19.9	42.31	20.5	39.45	18.8	26.62	15.0	37.36	23.2	19.95	11.3
$\widehat{\sigma_{\varepsilon,j}^2}$	-4.49	-2.3	-7.99	-3.7	-2.99	-0.9	-14.81	-3.9	-17.96	-7.9	9.03	5.3	-5.07	-2.6
$\widehat{\sigma_{\eta,j}^2}$	-55.27	-14.3	-59.04	-15.0	-97.25	-15.8	-130.39	-20.0	-86.36	-11.8	-30.06	-7.4	-29.94	-5.6
Graduation Rate$_j$	0.25	3.1	0.11	1.1	-0.85	-7.0	-1.19	-8.5	0.94	10.8	0.26	3.1	1.1	12.4
Distance$_{ij}$	-0.55	-21.8	-0.55	-21.5	-0.59	-22.6	-0.64	-23.2	-0.50	-19.1	-0.52	-21.2	-0.49	-18.7
D(HS relevant)$_{ij}$			0.59	16.9			1.13	27.3						
D(Parent w/this)$_{ij}$			0.88	18.0			0.80	15.7						
D(long)$_j$			-0.0075	-0.2			-0.34	-8.3						
D(humanities)$_j$					-0.56	-7.6	-0.55	-7.4						
D(agriculture)$_j$					-0.9	-9.6	-1.32	-13.6						
D(business)$_j$					-0.05	-0.6	-0.77	-8.4						
D(social science)$_j$					-0.79	-10.6	-1.59	-19.0						
D(health care)$_j$					-1.3	-14.0	-1.79	-18.8						
D(biological sci.)$_j$					-2.13	-20.8	-2.76	-25.2						
D(physical sci.)$_j$					0.51	6.1	0.07	0.8						
D(engineering)$_j$					0.26	3.4	-0.04	-0.5						
N	230544		230544		230544		230544		201962		230544		201962	

Note: Standard errors are adjusted for the use of generated regressors.

Table 4. Nested Logit Model of Educational Choice

	Coef.	t
$\widehat{\ln Y_{i,j,s}}$	1.00	6.1
$\widehat{g_{ij}^{y}}$	15.11	17.3
$\widehat{g_{j}^{o}}$	50.74	24.0
$\widehat{\sigma_{\varepsilon,j}^{2}}$	-12.95	-5.6
$\widehat{\sigma_{\eta,j}^{2}}$	-37.3	-11.6
Graduation Rate$_j$	0.40	3.5
Distance$_{ij}$	-3.9	-15.3
$\widehat{\lambda_1}$	1.09	14.5
$\widehat{\lambda_2}$	0.31	6.7
$\widehat{\lambda_3}$	0.42	6.8
$\widehat{\lambda_4}$	0.07	4.3
$\widehat{\lambda_5}$	0.68	19.5
$\widehat{\lambda_6}$	0.21	8.1
$\widehat{\lambda_7}$	2.14	47.3
$\widehat{\lambda_8}$	1.52	27.3
$\widehat{\lambda_9}$	0.71	5.9
$\widehat{\lambda_{10}}$	0.50	5.9
$\widehat{\lambda_{11}}$	1.01	29.1
$\widehat{\lambda_{12}}$	1.50	10.8
$\widehat{\lambda_{13}}$	0.53	13.6
$\widehat{\lambda_{14}}$	1.00	-
$\widehat{\lambda_{15}}$	0.36	11.4
N	230,544	

Table 5. Minimum Distance Estimation of Relative Risk Aversion

Left: Borrowing Rate=4%, Lending Rate=1%, Right: Borrowing Rate=7.5%, Lending Rate=1%

		g^{y}	g^{o}	σ_{η}^2	σ_{ϵ}^2			g^{y}	g^{o}	σ_{η}^2	σ_{ϵ}^2
		Scaled Linearization Coefficient in Life-time Simulation Value Function, $\alpha_2^{**}(\gamma)$						Scaled Linearization Coefficient in Life-time Simulation Value Function, $\alpha_2^{**}(\gamma)$			
β	γ					β	γ				
0.98	2	11.25	6.73	-24.25	-0.11	0.98	2	10.18	5.07	-18.92	-0.10
0.98	3	10.95	5.97	-36.18	-0.14	0.98	3	9.57	4.09	-30.17	-0.24
0.98	4	10.88	5.26	-52.60	-0.29	0.98	4	9.10	3.18	-44.68	-0.30
0.98	5	10.78	4.75	-65.10	-0.26	0.98	5	8.70	3.05	-52.51	-0.49
0.98	6	10.87	4.59	-79.03	-0.36	0.98	6	8.79	2.64	-61.63	-0.41
0.98	7	10.79	4.23	-92.96	-0.32	0.98	7	8.60	2.74	-67.44	-0.47
0.98	8	11.18	4.21	-172.2	-0.45	0.98	8	8.65	2.77	-73.87	-0.35
0.98	9	10.93	4.50	-226.0	-0.53	0.98	9	8.21	2.88	-78.70	-0.40
0.98	10	11.84	4.26	-1422	-0.72	0.98	10	7.77	2.85	-90.55	-0.37
0.98	11	10.34	4.54	-1793	-0.78	0.98	11	7.33	2.86	-107.7	-0.50
0.98	12	9.72	3.60	-15696	-1.21	0.98	12	6.98	2.69	-145.1	-0.64
0.98	14	6.27	5.19	-88677	-1.47	0.98	14	6.78	2.52	-360.9	-1.12
0.98	16	6.81	9.00	-376684	-1.29	0.98	16	6.75	3.23	-2170	-1.65
0.98	18	11.30	14.37	-1853619	-1.05	0.98	18	6.53	4.95	-35444	-1.71
0.98	20	20.79	22.37	-11544777	-0.82	0.98	20	5.66	7.11	-766311	-1.32
		Scaled Coefficient In Empirical Model, $\widehat{\alpha_2^{**}}$						Scaled Coefficient In Empirical Model, $\widehat{\alpha_2^{**}}$			
		17.07	34.62	-51.14	-4.16			17.07	34.62	-51.14	-4.16

For Borrowing Rate=4%: Minimum distance estimate of γ: 4. Standard error: 0.5.

For Borrowing Rate=7.5%: Minimum distance estimate of γ: 5. Standard error: 1.0.