Discussion Paper No. 02-06

A Model of Educational Attainment

Application to the German Case

Charlotte Lauer

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 02-06

A Model of Educational Attainment

Application to the German Case

Charlotte Lauer

Non technical summary

This article develops a model of educational attainment which is imbedded in the human capital theory, the reference theory for the economic analysis of issues related to education. The principle of this model is that there is a finite number of possible educational alternatives which may be ordered by level. Each educational alternative yields a certain utility to the individual. The net utility of an educational alternative is expressed in terms of the difference between the "returns" and the "costs" associated with this alternative for any individual with specific characteristics and constraints. The individual chooses, given his specific constraints or characteristics, to attain the education level which maximises his net utility.

In its simplest version, the formalisation of these assumptions leads to an ordered probit model where the threshold values are given by the expected ratio of the marginal costs to the marginal returns. In other words, the individual assesses the marginal costs and the marginal returns associated with the next higher education level, and if the ratio of the former to the latter, given his characteristics, is below a certain threshold, he will opt for the next higher level. In a second step, the educational process is decomposed into two stages: achievement at school and post-school achievement. The same conceptual framework applies at both stages, but a correlation is permitted between them. Empirically, this boils down to estimate a multivariate ordered probit model, where we have one equation for secondary education and as many post-secondary education equations as there are secondary school degrees. The model makes it possible to investigate whether some factors exert a different influence depending on the stage in the educational process, and whether time-variant factors also play a role.

An illustration based on German data shows how this model functions in practice. The explanatory variables used are the traditional ones encountered in the literature on educational attainment: cohort dummies and variables depicting family background. Differences across genders in the parameters have been systematically tested. The example of an application of the basic version of the model shows that it is very easily estimable empirically, while providing an economic interpretation framework which makes sense in economic terms. The application of the decomposition of the educational career points to the usefulness of considering possible correlations between the different steps in the educational career if one intends to analyse specific transitions within the educational process. Moreover, some of the explanatory variables prove to exert a different influence at the school and at the post-school level.

A Model of Educational Attainment

Application to the German case

Charlotte Lauer

Centre for European Economic Research
Zentrum für Europäische Wirtschaftsforschung (ZEW)
Mannheim

Abstract: This paper aims at formulating a theoretical model of educational attainment which is directly implementable empirically. The conceptual framework relies on the human capital theory, where individuals weigh the costs and the returns of education to decide on the optimal amount of education they want to invest in. In an extension of the model, the educational process is decomposed into two stages: achievement at school and post-school achievement. The same conceptual framework applies at both stages, but a correlation is permitted between them. Empirically, this boils down to estimate a multivariate ordered probit model. An illustration based on German data shows how this model functions in practice. Thus, the model confers an economic foundation on the empirical estimation of the impact of certain factors on educational outcomes.

1   Introduction

In most industrialised countries, the level of educational attainment has risen steadily over the past century. This educational expansion has been largely encouraged by public authorities, which devote a non-negligible part of their financial resources to education (5.1 percent of national GDPs on average in OECD countries in 1997, OECD (2000)). Public involvement in the promotion of education is justified by the fact that education is viewed as an essential factor for socio-economic development and welfare. In modern economies, the skills acquired within the education system are an essential production factor and raising them is therefore expected to have a positive impact on economic growth. This is the concept of "human capital" often referred to in economic literature, in analogy to physical capital. On the individual level, it is also of interest to invest resources in education. Beyond the obvious utility of a minimum of instruction for being able to cope with life in general, most people are well aware that investing in further education pays off in the future and improves their prospects in many respects.

If individual and collective interests coincide, why should the State be concerned with educational investments? In fact, intervention from the State is only required to the extent that private educational efforts would lead to sub-optimal outcomes, i.e. to the extent that there are market imperfections. There are some theoretical arguments that this should be the case as far as education is concerned. A first argument relates to the presence of positive externalities, such as the positive effects of a highly educated society in terms of criminality, health etc. Most individuals know that acquiring skills through education will raise their chances in subsequent life, but since they do not reap directly the whole benefit of it, due to the presence of positive externalities, they will tend to under-invest in their education compared to the socially optimal level. Another argument relates to capital market imperfections. As e.g. Kodde (1988) argues, commercial institutions are likely to consider human capital to be a risky, non-liquid and intangible asset. As a consequence, in general, higher credits can only be obtained against higher borrowing rates. Besides, individuals may have different access to financial resources, and for some individuals more than for other, funds may be rationed at specific levels. In the presence of capital market imperfections, individuals of poorer background might not be able to invest in the optimal amount of education. To sum up, left to the mere market forces, individuals would tend to under-invest in their education, especially those individuals facing higher credit constraints.

Beyond the level of compulsory schooling, public authorities can only influence educational participation by setting incentives. This, in turn, can only be efficient if the factors which influence the educational decisions of individuals are known, both in order to identify where the needs are and possibly to gain

evidence on the effectiveness of the means employed. Beside the usefulness of getting knowledge on the structure and determinants of educational attainment with the aim of influencing educational outcomes, it is also essential for policy-makers to gain evidence on what determines educational outcomes if they are to plan educational needs for the future.

This article aims at formulating an economic model of educational attainment which makes it possible to estimate empirically the role of certain factors in the determination of educational outcomes. To this end, a theoretical interpretation framework is developed in such a way that it may be relatively easily implemented empirically with standard statistical packages. The conceptual framework relies heavily on the human capital theory, the basic assumption of which being that individuals weigh up the costs and the returns to education and decide on the optimal amount of education they want to invest in accordingly. In a first step, a basic version of the model is developed, which makes it possible to easily identify empirically the relationship between certain factors and final educational outcome. In a second step, an extension of the model allows to take a closer look at the process of educational attainment and the educational process is decomposed into successive stages. This is of particular interest if the impact of time-variant factors on educational prospects is to be analysed, or if more details are to be gained regarding the stages at which the identified factors play a role within the educational career.

The article is structured as follows. After an overview, in section 2, of the various ways educational outcomes have been modelled in the literature available so far, the article turns to the presentation of the model proposed here. First, section 3 presents the basic version of this model of educational attainment. Then, the extension of the model, which consists in decomposing the educational career into school and post-school attainment, is explained in section 4. For each step, an empirical application aims to show, with the help of German data, how the model can be implemented in practice. Finally, section 5 summarises and concludes.

2   Modelling educational attainment: overview of the literature

The bulk of the work related to the modelling of educational outcomes is more or less explicitly imbedded in the human capital theory, which dates back to Becker (1964) and Mincer (1974). According to the human capital approach, individuals decide on the optimal amount of education they want to invest in by weighing up the costs and the returns of additional schooling. The acquisition of education generates costs, in any case opportunity costs in the form of

foregone earnings during the time education is taken and possibly also direct costs (e.g. books, tuition fees etc.). On the other hand, education is assumed to raise one's productivity and consequently to pay off in terms of future earnings. Therefore, an optimising individual will choose the level of education which maximises his net return and continue to invest in further education until the marginal cost of additional investment equals the marginal return. This theory implies that human capital investment differs among individuals due to differences in either return or cost conditions.

The human capital theory has been criticised for privileging mere monetary considerations, mainly regarding education as a monetary investment and disregarding non-monetary utility aspects or consumption motives. As a matter of facts, education may contribute to non-monetary utility in the short run as well in the long run. As an answer to this objection, some economists extended the model by including the consumption motive in the human capital approach, assuming that education increases the efficiency of allocating leisure (see Heckman (1976) or Kodde (1988), the latter showing that the demand for education is higher if one integrates consumption motives.).

Another extension of the human capital model can be found in the household production model, introduced by Becker (1965), which insists on the necessity of viewing the process of educational attainment as an aspect of family behaviour rather than individual behaviour. According to this approach, children's educational attainment is a commodity which enters the household utility production. The household is viewed as a production unit which generates utility for its members using some market inputs (e.g. school quality) as well household inputs (e.g. time, family characteristics). Adults in the family, typically parents, make decisions regarding the generation and the use of household resources. Parental decisions may affect children's outcomes directly, through for instance the amount of resources allocated to the children, the nature of these resources or their timing, but also indirectly, through for instance decisions related to the family structure or the location the children grow up in. The household production model has been continuously refined and complemented. Among others, Taubman (1976) stresses the importance of genetic inheritance, since scholastic abilities are to some extent passed on from one generation to the next. Also Becker and Tomes (1986), and more recently Ermisch and Francesconi (2001), insist on the inheritability of certain endowments like ability or preferences, which translate into human capital. Hanushek (1992) points to the fact that there is a trade-off for the parents between the number and of the "quality" of the children, here measured by their scholastic performance, because parents maximise the household utility, a function of the quantity and the quality of the children, subject to a budget constraint and a time constraint. More recently, Wilson (2001) tried to integrate in a model of educational attainment aspects of both the human capital

theory and the production function model.

Another vein of literature has a sociological focus and examines issues like social mobility, social stratification or inequality in educational opportunities, seeking to provide hypotheses to explain why the various social groups do not have the same educational prospects. For instance, the socialisation approach emphasises the transmission by peers or role models, typically the parents, to the child of a pattern of behaviour. Another example is the life course development hypothesis which highlights the role of the age at which educational decisions occur. Indeed, as Müller (1990) states, as students grow older, they will increasingly be able to make their own choices and will be less dependent on their family background. An interesting consequence of this hypothesis is that if this turns out to be true that the timing of educational decisions in the life course matters, educational policy could reduce social inequality in terms of educational prospects by delaying the moment when educational decisions are to be made (see the formulation of alternative hypotheses in Shavit and Blossfeld (1993) to explain differences in educational attainment across countries).

As far as empirical evidence is concerned, the major part of the literature does not aim primarily at establishing causal relationships but rather focus on the nature and the extent of correlations between parental and child's outcomes, without being primarily interested in the process by which these correlations arise. In many studies, empirical analyses are only loosely related to a theoretical model, even though other studies do imbed their empirical analysis in a more or less explicit theoretical framework (see the overview of Haveman and Wolfe (1995)). More concretely, a typical approach has been, following Blau and Duncan (1967) to estimate the impact of selected variables on ultimate educational attainment, typically measured by completed years of schooling. Such an approach assumes that there exists a linear relationship between the explanatory factors, typically family background variables, and years of schooling.

By contrast, Mare (1980), after Boudon (1974), points to the usefulness, as a complement to the traditional approach, of viewing final educational attainment as the result of a process which can be decomposed into a finite number of stages at which the student makes decisions. In other words, educational attainment can be viewed as a sequence of transitions within the education system, and the product of the transition probabilities generates the probability of final educational attainment. Beyond the interest of analysing specific transitions per se, Mare (1980) argues that restricting the analysis to the determinants of only final educational attainment is not informative with respect to the stage at which the explanatory factors exert an influence, whether they affect the probability of transiting to the highest level of education attained or

the probability of successfully completing some earlier transition. The approach he proposes therefore consists in analysing school continuation decisions. Given information on the highest grade finally completed, the school continuation decisions at each grade can be deduced, i.e. the probability that a person attended a particular grade of schooling, given that the person completed the previous grade. Beside empirical studies which focus on one particular transition in isolation from other transitions (e.g. Willis and Rosen (1979) and Manski and Wise (1983) examine the transition from high school to college attendance), many studies (among other Shavit and Blossfeld (1993)) are directly based on the schooling-transition model of Mare (1980).

In a recent work, however, Cameron and Heckman (1998) have criticised the schooling-transition model as applied in the empirical literature and questioned the empirical conclusions usually drawn from the application of this model. Their main objection relates to educational selectivity. Indeed, schooling transition models generally ignore the consequence of factors who are not observable to the scientist but might influence transitions, such as ability or motivation. These unobserved factors may be correlated across transitions. As a consequence, a dynamic selection bias arises, since from one schooling stage to the next, only a certain type of students remains and the sample composition becomes progressively selective with respect to unobserved heterogeneity.

In section 3, a model of educational attainment is formulated which draws on the model of Cameron and Heckman (1998). The basic model presented differs less in terms of real economic content than in terms of formulation from that of Cameron and Heckman (1998). Nevertheless, it is reformulated in such a way that it can be directly implemented and interpreted empirically. Section 4 proposes an extension of this model which consists in decomposing educational attainment into school attainment and post-school attainment, taking problems of dynamic selectivity into account. In each section, an example of empirical application of the model based on German data is presented with a view to illustrating the empirical opportunities offered by both versions of the model.

3 Basic model

3.1 Formulation of the model

Let us suppose that each individual chooses, given some constraints, how much he wants to invest in education, i.e. chooses the education E he wants to acquire among J educational alternatives E_j of increasing levels, with $j \in \{1...J\}$[1].

---

1 Cameron and Heckman (1998) express the educational decision in terms of years of schooling.

Let E^* be the desired educational level. The desired level of education is not observable, but only the actual decision E_j of the individual[2] i.e. the education level j actually chosen.

The decision is assumed to be rational in the sense that it maximises the perceived utility for the individual, subject to some constraints. The perceived utility is defined as the expected net returns, i.e. the difference between expected returns and expected costs of each educational alternative E_j, given some characteristics. The concept of utility may cover monetary as well as non monetary aspects. Consequently, the terms "returns" and "costs" are used for ease of language, but may also be non pecuniary. Thus, the optimal educational decision for an individual with a given vector of characteristics x is given by:

$$Max_{j \in \{1...J\}} \; r(E_j \,|\, x) - c(E_j \,|\, x) \qquad (1)$$

where r denotes the expected return and c the expected cost associated with educational level E_j. The return function is assumed to be strictly positive, weakly convex and increasing with the education level. The cost function is assumed to be positive, concave and increasing with the education level. The costs function starts with the value zero, while the return function is strictly positive. These assumptions ensure that the net expected return will be concave and positive for at least the lowest possible education level. The costs and returns of education are allowed to differ across individuals and it is assumed that they are affected by characteristics observable to the scientist and by some other unobserved factors. Thus, the cost and return functions can be decomposed in the following way:

$$\begin{aligned} r(E_j \,|\, x) &= r(E_j)\,\varphi_r(x)\,\varepsilon_r \qquad (2) \\ c(E_j \,|\, x) &= c(E_j)\,\varphi_c(x)\,\varepsilon_c \end{aligned}$$

where $\varphi_r(x)$ is a positive function defining the effects of the observed characteristics on the expected returns to education and ε_r is a random variable accounting for the effect of unobserved individual heterogeneity on the expected returns[3]. Similarly, $\varphi_c(x)$ is a positive function which defines the effects of the observed characteristics on the expected costs of education and ε_c is a positive random variable representing the impact of unobserved individual

---

2 It does not matter who in fact makes the decision, whether it is the individual himself or somebody else (the parents, for instance). What counts is the outcome of the decision among the possible alternatives. For ease of presentation, we will speak of the individual rather than of the decision-maker, since the outcome concerns the individual.

3 Due to the multiplicative structure of the model, $\varphi_r(x) > 1$ (resp. < 1) implies that the observed characteristics of an individual, taken together, increase (resp. decrease) the return expectations. Similarly, $\varepsilon_r > 1$ (resp. < 1) means that the unobserved factors increase (resp. decrease) the return expectations.

heterogeneity. Thus, the observed characteristics as well as unobserved individual heterogeneity may affect the expected returns and the expected costs in different ways[4]. However, the personal shifters φ_r, φ_c, ε_r, and ε_c are assumed not to depend on the specific education level. Without loss of generality, it is assumed that $E(\varepsilon_r) = E(\varepsilon_c) = 1$, meaning that unobserved heterogeneity has on average a neutral effect on the return as well as on the cost expectations.

The optimal educational decision E_{j*} is such that the expected net return is maximised, i.e. the net return associated with E_{j*} must be positive, larger than the net returns expected from the next lower education level E_{j*-1} and at least as large as those expected from the next higher education level E_{j*+1}[5]:

$$
\begin{aligned}
r(E_{j*})\,\varphi_r(x)\,\varepsilon_r - c(E_{j*})\,\varphi_c(x)\,\varepsilon_c &> 0 \\
r(E_{j*})\,\varphi_r(x)\,\varepsilon_r - c(E_{j*})\,\varphi_c(x)\,\varepsilon_c &> r(E_{j*-1})\,\varphi_r(x)\,\varepsilon_r - c(E_{j*-1})\,\varphi_c(x)\,\varepsilon_c \\
r(E_{j*})\,\varphi_r(x)\,\varepsilon_r - c(E_{j*})\,\varphi_c(x)\,\varepsilon_c &\geq r(E_{j*+1})\,\varphi_r(x)\,\varepsilon_r - c(E_{j*+1})\,\varphi_c(x)\,\varepsilon_c
\end{aligned}
$$

This is equivalent to:

$$
\begin{aligned}
\varphi_c(x)\varepsilon_c \left[r(E_{j*}) \cdot \frac{\varphi_r(x)\varepsilon_r}{\varphi_c(x)\varepsilon_c} - c(E_{j*}) \right] &> 0 \\
\varphi_c(x)\varepsilon_c \left[r(E_{j*}) \cdot \frac{\varphi_r(x)\varepsilon_r}{\varphi_c(x)\varepsilon_c} - c(E_{j*}) \right] &> \varphi_c(x)\varepsilon_c \left[r(E_{j*-1}) \cdot \frac{\varphi_r(x)\varepsilon_r}{\varphi_c(x)\varepsilon_c} - c(E_{j*-1}) \right] \\
\varphi_c(x)\varepsilon_c \left[r(E_{j*}) \cdot \frac{\varphi_r(x)\varepsilon_r}{\varphi_c(x)\varepsilon_c} - c(E_{j*}) \right] &\geq \varphi_c(x)\varepsilon_c \left[r(E_{j*+1}) \cdot \frac{\varphi_r(x)\varepsilon_r}{\varphi_c(x)\varepsilon_c} - c(E_{j*+1}) \right]
\end{aligned}
$$

Let us define: $\varphi(x) = \dfrac{\varphi_r(x)}{\varphi_c(x)}$ and $\varepsilon = \dfrac{\varepsilon_r}{\varepsilon_c}$.

$\varphi(x)$ measures the net impact of observed characteristics x and ε the net effect of unobserved individual heterogeneity on the expected relation of returns to costs. Since $\varepsilon_r > 0$, $\varepsilon_c > 0$, $\varepsilon > 0$, $\varphi_r(x) > 0$ and $\varphi_c(x) > 0$, one obtains after simplification:

$$
\begin{aligned}
r(E_{j*})\,\varphi(x)\,\varepsilon - c(E_{j*}) &> 0 \\
r(E_{j*})\,\varphi(x)\,\varepsilon - c(E_{j*}) &> r(E_{j*-1})\,\varphi(x)\,\varepsilon - c(E_{j*-1}) \\
r(E_{j*})\,\varphi(x)\,\varepsilon - c(E_{j*}) &\geq r(E_{j*+1})\,\varphi(x)\,\varepsilon - c(E_{j*+1})
\end{aligned}
$$

---

4 A high scholastic ability, for instance, a typically unobserved factor, reduces the cost of educational investment, but might not increase to the same extent the return to educational investments, since the way the acquired knowledge is "transformed" into wages depends on another kind of ability.

5 If the net return is equal for E_j and E_{j+1}, the individual is assumed to choose alternative E_j.

or

$$0 < \frac{c(E_{j^*})}{r(E_{j^*})} \cdot \frac{1}{\varphi(x)} < \varepsilon$$

$$\frac{c(E_{j^*}) - c(E_{j^*-1})}{r(E_{j^*}) - r(E_{j^*-1})} \cdot \frac{1}{\varphi(x)} < \varepsilon \leq \frac{c(E_{j^*+1}) - c(E_{j^*})}{r(E_{j^*+1}) - r(E_{j^*})} \cdot \frac{1}{\varphi(x)}$$

Thus, for any individual with observed characteristics x, the expected net return is positive at the optimum and the unobserved individual component is bounded by the expected ratios, given x, of marginal costs to marginal returns of moving from the next lowest education level for the lower bound, and of moving to the next highest level for the upper bound. Consequently, the probability that an individual chooses E_j is given by:

$$Pr(E_j \,|\, x) = Pr \left[\frac{c(E_j) - c(E_{j-1})}{r(E_j) - r(E_{j-1})} \cdot \frac{1}{\varphi(x)} < \varepsilon \leq \frac{c(E_{j+1}) - c(E_j)}{r(E_{j+1}) - r(E_j)} \cdot \frac{1}{\varphi(x)} \right] \quad (3)$$

To simplify the notation, let us call:

$$\mu_j = \ln \left(\frac{c(E_{j+1}) - c(E_j)}{r(E_{j+1}) - r(E_j)} \right) \text{ and } \gamma_j = \mu_j \cdot \frac{1}{\varphi(x)} \quad (4)$$

If a change, or a difference across individuals, in x results, for example, in higher expected costs of E_{j+1}, this raises the threshold γ_{j+1} and results in a lower probability that the individual chooses E_{j+1} or higher and a higher probability that he chooses E_j at most[6]. If conversely a change in x brings about higher returns for E_{j-1}, this raises the lower bound γ_j and therefore increases the probability of choosing at most E_{j-1} rather than E_j or higher. Thus, any change in the observed characteristics x may change educational decisions to the extent that it affects the expected ratios of marginal costs to marginal returns from the next lower level and to the next higher level. Note that in this model, it is not necessary to assess the *actual* costs and returns of each educational alternative, but it is enough to determine how the observed characteristics influence the *perceived* ratio of costs to returns. Taking the logarithm of the expression in brackets in equation (3) and assuming that $\varphi(x) = \exp[\beta x]$, we obtain:

$$Pr(E_j \,|\, x) = Pr \left[\mu_{j-1} - \beta x < \ln \varepsilon \leq \mu_j - \beta x \right] \quad (5)$$

---

6 Note that we can only determine the exact impact of a change in x on the probability of completing exactly E_j - and not at most or at least - if we consider the impact of x on both bounds and if we know the distribution function of ε.

If we assume that $\ln \varepsilon$ is normally distributed with mean 0 and variance σ^2, equation (5) may be rewritten as

$$Pr(E_j \mid x) = \Phi\left(\frac{\mu_j - \beta x}{\sigma}\right) - \Phi\left(\frac{\mu_{j-1} - \beta x}{\sigma}\right) \tag{6}$$

where and Φ is the cumulative standard normal distribution function.

This expression takes the familiar form of an ordered probit model, where the μ_j's are the cut values[7]. The ordered probit model written above can only be identified up to some factor of proportionality. Since this is the ratio of the parameters to σ which matters, it is common practive to normalise σ to 1 (Maddala, 1983, p.23). The parameters β and the cut values μ_j can be estimated by maximising the likelihood function for this model:

$$\mathcal{L} = \prod_{i=1}^{n} \prod_{j=1}^{J} \left[\Phi(\mu_j - \beta x_i) - \Phi(\mu_{j-1} - \beta x_i)\right]^{I_{ij}} \tag{7}$$

where I_{ij} is an indicator variable equal to 1 if the individual i, $i \in \{1...n\}$ opts for educational level E_j and 0 otherwise. Maximising L boils down to maximising $\ln L$, since L is a positive function and $\ln L$ a monotone increasing transformation of L. Therefore, the model can be easily estimated by maximising the following log-likelihood function[8]:

$$\ln \mathcal{L} = \sum_{i=1}^{n} \sum_{j=1}^{J} I_{ij} \ln \left[\Phi(\mu_j - \beta x_i) - \Phi(\mu_{j-1} - \beta x_i)\right] \tag{8}$$

3.2 An example

This section presents a simple example of an application of this model, based on a German data set, the GSOEP, a widely used data set in German socio-economic research. The GSOEP is a longitudinal household survey conducted on an annual basis since 1984 (for further information on the GSOEP, see Haisken-DeNew and Frick (2000)). For the empirical application, the sample of West German residents has been used and the cohorts born between 1929 and 1968 have retained[9]. This leaves us with a sample of about 6,000 individuals for the estimation.

---

7 Note that given the assumptions made for the cost and return functions, we have at the extremes $\mu_0 = -\infty$ and $\mu_J = +\infty$.

8 It can be estimated directly with standard statistical package like Stata 7.0, which include a command for the estimation of ordered probit models.

9 In order to maximise the number of observations, cohorts born before 1948 are taken from the 1985 wave of the panel, while cohorts born after 1948 are drawn from the 1999 wave.

We cannot model the decision as such, since we have no information on what motivates individuals in their educational choices, but we can postulate that the educational outcome we observe is the result of the decision made by the individual. Therefore, the dependent variable is defined as the highest degree obtained, in five ordered categories. The lowest education level E_1 (no vocational degree) is assigned to individuals who do not hold any vocational qualification and at most a school degree which do not entitle them to pursue tertiary level studies. Education level E_2 (apprenticehip) corresponds to an apprenticeship or a vocational degree of comparable level and a school degree below the maturity level. Education level E_3 is assigned to individuals who hold either an advanced vocational qualification (like a degree from a technical college or a master craftsman degree) and/or are in possession of the maturity certificate (the so-called *Abitur* or *Fachhochschulreife*) but do not hold a tertiary level degree. E_4 is the higher technical college (in German *Fachhochschule*) level, and the highest level E_5 is the university level.

To keep in line with usual pratice in empirical literature and keep the illustration of the model as simple as possible, the independent variables consist of birth cohort dummies (born between 1929 and 1938 as the reference category, born between 1939 and 1948, born between 1949 and 1958 and born between 1959 and 1968) in order to control for cohort effects and of a series of variables depicting family background. As we saw in section 2, the social background individuals grew up in is likely to influence the educational outcomes. Parental background might affect offspring's educational outcomes through the availability of financial resources within the family. In the presence of imperfect capital markets, and in case the parents do not dispose of enough money, then, investment in education might be limited by credit constraints (Rosenzweig and Wolpin (1993)). The GSOEP contains no direct information on parental income. However, there is information on the occupational situation of the father (not on that of the mother, though) during the childhood of the individual. This can serve as an indicator of the probable financial situation of the household the individual grew up in, which affects cost to return expectations, especially through the cost side. The occupational position of the father is depicted through a set of dummy variables: worker (reference category), farmer, self-employed, senior manager, middle manager or employee. A dummy for missing information on the father's occupation is also included in order to avoid problems of selective sample composition due to non-randomly missing information. Furthermore, another set of dummies describe the education level of the parents, defined as no vocational degree (reference category), apprenticeship, technical college/maturity or higher education. Here again, a dummy for missing information has been added.

Table 1 presents the results of specification tests. The first series of χ^2-tests aims at examining whether the parameters differ significantly between men

Table 1: Test results

Null hypothesis		χ^2	$\mathbf{p>}\chi^2$
Female * Birth cohort = 0		61.59	0.00
Female * Mother's education = 0		2.52	0.64
Female * Father's education = 0		7.37	0.12
Female * Father's occupation = 0		1.82	0.94
μ_1: Male = Female		227.86	0.00
μ_2: Male = Female		182.69	0.00
μ_3: Male = Female		154.67	0.00
μ_4: Male = Female		114.82	0.00
Males:	$\mu_1 = \mu_2$	2,042.40	0.00
	$\mu_2 = \mu_3$	814.33	0.00
	$\mu_3 = \mu_4$	211.85	0.00
Females:	$\mu_1 = \mu_2$	2,145.92	0.00
	$\mu_2 = \mu_3$	532.03	0.00
	$\mu_3 = \mu_4$	87.08	0.00

and women, in which case the specification should account for this. The tests on gender differences are based on an estimation of the model developed in section 3, with the five-level dependent variable described above, which includes in addition to the explanatory variables further dummies constructed by interacting all the explanatory variables with a dummy variable for gender (1 if the individual is a female, 0 if he is a male). The null hypothesis that all the interaction dummies for one set of variables are not jointly significantly different from zero is only rejected for the birth cohort variables (at a 1% significance level). For the other variables, the null hypothesis cannot be rejected at the 10% level, which means that gender differences in the impact of the family background variables are not significant in a statistical sense. A significance level of at least 10% has been retained as a criterion for the interaction dummies to be included in the final specification. Thus, only the interaction terms between the birth cohort and the sex of the individual have been retained in the final specification.

Furthermore, possible gender differences in the threshold values μ_j have also been tested, since there is no *a priori* reason to assume that the thresholds should be the same for men and women, even though this is a common assumption in empirical literature as far as ordered probit estimations are concerned. The test results in table 1 are based on a specification which includes only the significant interaction terms for birth cohort and allow the threshold

11

values to differ between men and women[10]. As can be seen, all the threshold values differ significantly between men and women at the 1% level. Therefore, it proves useful to allow the threshold values to differ across genders in the finally retained specification. In a next step, χ^2-tests on the threshold values were run in order to ensure that the categories of the dependent variable are really distinct. The hypothesis that the thresholds are not distinct has been tested, for both males and females. If two consecutive thresholds proved not to be statistically different, then, the educational categories should be aggregated. The test results show that the educational categories have been defined in an appropriate way since all threshold values differ significantly from each other, both for males and females.

Table 2 reports the estimation results. The estimated coefficients should be interpreted in a qualitative way: a positive and significant coefficient means that a value of one for the dummy variable is associated with a higher probability of reaching a higher level of educational attainment compared to the reference category[11]. A Wald test performed on the hypothesis that all coefficients except the threshold values are zero is rejected at the 1% level.

To obtain the overall effect of belonging to one cohort instead of belonging to the reference cohort 1929-38, one adds the coefficient of the simple cohort dummy to that of the same cohort dummy interacted with the gender dummy[12]. As can be seen, the cohort coefficients are significant, positive and increasing, which points to the presence of educational expansion and means that, all else equal, the cost to return ratio associated with education has decreased over the generations. However, the rise in the net utility of education from one generation to the next seems to follow a decreasing path[13]. There are significant differences between men and women in the cohort effects. The interactions with the female dummy are positive and increasing, which implies that the gap in the cost to return ratio in favour of men has decreased over time and that educational expansion has been stronger among women than

---

10 This cannot be estimated directly with the ordered probit command of Stata 7.0 but is programmable within the framework of maximum likelihood estimation (see Gould and Sribney (1999)).

11 It also means that the variable increases the probability of achieving the highest education category E_5 and decreases that of achieving only the lowest category E_1. For the intermediate categories, one can only say that a positive coefficient increases the probability of holding at least E_2 or of holding at least E_3. However, in order to obtain a precise measure of the impact of a variable on the probability of achieving exactly one specific education level, one needs to compute equation (6).

12 0.27+0.14=+0.41 for the 1939-48 cohort, 0.34+0.45=+0.79 for the 1949-58 generation and 0.22+0.60=+0.82 for the 1959-68 generation.

13 +0.41 for the 1939-48 cohort compared to the previous cohort 1929-38, 0.79-0.41=+0.38 for the 1949-58 cohort compared to the previous cohort 1939-48 and 0.82-0.79=0.03 for the 1959-68 cohort compared to the previous cohort 1949-58.

Table 2: Determinants of final educational attainment

Variable	coef.	(s.e)
Birth cohort (ref.: 1929-38)		
1939-48	0.27**	(0.06)
1949-58	0.34**	(0.06)
1959-68	0.22**	(0.06)
Female * 1939-48	0.14†	(0.08)
Female * 1949-58	0.45**	(0.09)
Female * 1959-68	0.60**	(0.08)
Mother's education (ref.: No voc. degree)		
Apprenticeship	0.22**	(0.03)
Tech. college/maturity	0.36**	(0.09)
Higher education	0.96**	(0.13)
Missing	0.05	(0.06)
Father's education (ref.: No voc. degree)		
Apprenticeship	0.27**	(0.05)
Tech. college/maturity	0.40**	(0.06)
Higher education	0.87**	(0.08)
Missing	0.13†	(0.07)
Father's occupation (ref.: Worker)		
Farmer	0.18**	(0.07)
Self-employed	0.48**	(0.06)
Senior manager	0.89**	(0.06)
Middle manager	0.56**	(0.05)
Employee	0.34**	(0.07)
Missing	0.14**	(0.04)
Thresholds		
μ_1 Male	-0.55**	(0.06)
Female	0.44**	(0.06)
μ_2 Male	1.07**	(0.06)
Female	1.98**	(0.07)
μ_3 Male	1.82**	(0.06)
Female	2.74**	(0.07)
μ_4 Male	2.13**	(0.07)
Female	2.97**	(0.07)
Observations	6,005	
Log-likelihood	-7,122.45	
	χ^2	p$>\chi^2$
Wald test	1,436.25	0.00

Significance level : $\dagger$: 10% $*$: 5% $**$: 1%

Dependent variable: E_1 = no voc. degree, E_2 = apprenticeship,

E_3 = tech. college or maturity, E_4 = higher tech. college, E_5 = university.

among men.

Parental education has a significant impact on children's educational outcomes. This is true for both mother's and father's education. In particular, having parents who hold a higher education degree seems to particularly favour educational prospects. This may be because children growing up in families where the education level of the parents is high inherit to some extent the learning ability of their parents, which diminishes the costs of acquiring education and therefore - to stick to the model notation - lowers the thresholds. A high ability might also help to better convert education into utility (e.g. wages, free time use) and thus increase the return, which further lowers the thresholds. Moreover, highly educated parents generally place greater value on education and are therefore more likely to encourage their offsprings to pursue further studies. Thus, the perception of the return to education is expected to be higher among sons and daughters of highly educated persons. The impact of the father's education has the same order of magnitude as that of the mother's education, and is even slightly stronger - except for the higher education dummy - even though the occupational position of the father is controlled for, while we have no information on the mother's occupational position[14].

As far as the occupational position of the father is concerned, all coefficients are positive and significant, which means that the offsprings of blue collar workers (the reference category) have the worst educational prospects. The best educational opportunities, all other things equal, have children of senior managers, followed by those of intermediate level managers and self-employed. Children of farmers have only a slight advantage in terms of educational attainment compared to worker's children, but are significantly disadvantaged compared to children of employees. These results might be the consequence of financial constraints which raise the costs of investing in education and thus enhance the cost to return ratio. This could also reflect different systems of preferences, e.g. worker families value less education than other and children growing up in those families would expect a lower utility from education or higher costs than other families.

As seen before, there are significant differences in the threshold values between men and women. To be more specific, women have systematically higher thresholds than men. This means that for a given family background and a given cohort, women expect higher costs and/or lower returns to education and invest less in education than their male counterparts.

---

14 Therefore, the influence of mother's education may be of indirect nature and partly stem from the correlated, but unobserved effect of the mother's occupational position.

4 Extension: decomposition of the educational career

The model described and applied above is very simple to implement empirically and provides a convenient interpretation framework for the analysis of the determinants of educational attainment. In particular, it allows an interpretation of the cut values obtained through the ordered probit estimation which makes sense in economic terms. This is a very convenient model if one intends to explain final educational attainment of individuals depending on certain background characteristics. In a further step, an extension of this model is proposed in order to allow the analysis to go into the detail of the educational process. Indeed, after having determined the overall impact of certain factors on final educational achievement, it may be of interest to examine at which stage of the educational career which factors exert an influence. As a matter of fact, some factors might influence educational decision more decisively (or exclusively) at an earlier stage of the educational career, other at a later stage. For instance, has the observed educational expansion taken place at all levels? Does parental background play a more important role for earlier or for later stages of the educational process? Are gender differences similar throughout the educational career? Moreover, changes in the environment at a given stage, such as a reform in the educational system or changing economic conditions, changing situation of the parents, new information, or simply personality development over time, might modify the appreciation of returns and costs and therefore induce a reorientation of decisions for subsequent stages. Finally, at certain stages of the educational career, explicit decision-taking is required, which implies that a reorientation is possible, under the constraints of a new context and of previous choices which might have limited the possible subsequent alternatives.

In order to examine these issues, the educational career is now decomposed into successive stages: primary education, general secondary education and post-secondary education. This decomposition is general enough to be applicable to most educational systems. Given that all individuals are assumed to receive primary education, the final educational outcome is the result of two decisions: the decision on general secondary education and that on post-secondary education, given the level attained in general secondary education.

In practical terms, an appealing approach owing to its simplicity would be to consider that for each transition, the model defined in section 3 applies, i.e. after having completed secondary education, individuals decide on their post-secondary education, given the new set of variables and the choices available, independently of their previous choices. Practically, this would result in a sequential-response model, which only differs from the textbook models as

presented in Amemiya (1986) pp.310, Maddala (1983) pp.49 or Gouriéroux (1989) pp.249 through the fact that we have an ordered choice instead of a binary choice. However, this approach is only valid under the assumption that the probability of the choice at the second decision point is independent of the choice at the first point, i.e. that the random factors influencing responses at various stages are independent. Otherwise, the estimated coefficients will be biased. Since a selection with respect to unobservable factors may take place, we do not want to impose a priori such a restriction and estimate the correlation between the first and the second stage along with the other parameters instead of *a priori* assuming it equal to zero.

4.1   Formulation of the model

Let us call E^S the level of general secondary education attained by an individual among K alternatives E_k^S of increasing levels, with $k \in \{1...K\}$, and E^{S*} the unobservable desired secondary school attainment. Furthermore, let us define K variables E^{Fk}, with $k \in \{1...K\}$, which represent the final educational level an individual attains given that he holds a secondary school degree of level k, and E^{Fk*} the corresponding unobservable desired educational attainment. Given his secondary school degree of level k, the individual may choose one of L_k educational alternatives E_ℓ^{Fk} of increasing levels, with $\ell \in \{1...L_k\}$. To sum up, we have the following variables:

$$
\begin{array}{llll}
\text{Secondary schooling:} & E^S & = E_k^S & , \ k \in \{1...K\} \\
\text{Final education:} & E^{F1} & = E_\ell^{F1} & , \ \ell \in \{1...L_1\} \\
& E^{F2} & = E_\ell^{F2} & , \ \ell \in \{1...L_2\} \\
& \ldots & & \\
& E^{FK} & = E_\ell^{FK} & , \ \ell \in \{1...L_K\}
\end{array}
$$

Like in section 3, we can define K+1 functions $\varphi^S(x^S)$, $\varphi^{F1}(x^{F1})$, ..., $\varphi^{FK}(x^{FK})$, and K+1 error terms ε^S, ε^{F1}, ..., ε^{FK}. Furthermore, we assume that $\varphi^t(x^t) = \exp[\beta_t x_t]$, with $t \in \{S, F1...FK\}$ and, analogue to equation (5), obtain for all $k \in \{1...K\}$ and $\ell \in \{1...L_k\}$:

$$Pr(E_k^S \,|\, x^S, E_\ell^{Fk} \,|\, x^{Fk}) = \tag{9}$$
$$Pr\left[\mu_{k-1}^S - \beta^S x^S < ln\varepsilon^S \leq \mu_k^S - \beta^S x^S \,,\ \mu_{\ell-1}^{Fk} - \beta^{Fk} x^{Fk} < ln\varepsilon^{Fk} \leq \mu_\ell^{Fk} - \beta^{Fk} x^{Fk}\right]$$

where

$$\mu_j^S = \ln\left(\frac{c(E_{j+1}^S) - c(E_j^S)}{r(E_{j+1}^S) - r(E_j^S)}\right) \ \text{and} \ \mu_j^{Fk} = \ln\left(\frac{c(E_{j+1}^{Fk}) - c(E_j^{Fk})}{r(E_{j+1}^{Fk}) - r(E_j^{Fk})}\right) \tag{10}$$

If we knew for sure that $E_k^S \mid x^S$ and $E_\ell^{Fk} \mid x^{Fk}$ are independent from each other, we would have $Pr(E_k^S \mid x^S, E_\ell^{Fk} \mid x^{Fk}) = Pr(E_k^S \mid x^S) \cdot Pr(E_\ell^{Fk} \mid x^{Fk})$ and we could estimate the equation for secondary schooling and those for final education given secondary schooling separately and in the same way as in section 3. However, some unobserved factors might affect educational attainment both at the secondary level and at the final educational level and these equations might be correlated. If this is the case, estimating the equations separately would lead to biased estimates. Therefore, we allow the error terms to be correlated and assume that they follow a multivariate normal distribution of the following form[15]:

$$
\begin{pmatrix} ln\varepsilon^S \\ ln\varepsilon^{F1} \\ ln\varepsilon^{F2} \\ \vdots \\ ln\varepsilon^{FK} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \varrho_1 & \varrho_2 & \cdots & \varrho_K \\ \varrho_1 & 1 & 0 & \cdots & 0 \\ \varrho_2 & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ \varrho_K & 0 & \cdots & 0 & 1 \end{pmatrix} \right)
$$

Like in section 3, only the ratio of the β's to the σ's can be identified, therefore the σ's are normalised to 1. We have K correlation parameters $\varrho_k, k \in \{1...K\}$, corresponding to the correlation between secondary school choice and final educational choice given the secondary school degree obtained, for each secondary school degree.

To simplify the presentation, let us define, similarly to equation (4), $\gamma_k^S = \mu_k^S - \beta^S x^S$ and $\gamma_\ell^{Fk} = \mu_\ell^{Fk} - \beta^{Fk} x^{Fk}$. Thus, equation (9) becomes:

$$ Pr(E_k^S \mid x^S, E_\ell^{Fk} \mid x^{Fk}) = Pr \left[\gamma_{k-1}^S < ln\varepsilon^S \leq \gamma_k^S, \gamma_{\ell-1}^{Fk} < ln\varepsilon^{Fk} \leq \gamma_\ell^{Fk} \right] $$

Given the assumptions on $ln\varepsilon^S$ and $ln\varepsilon^{Fk}$, we obtain:

$$ Pr(E_k^S \mid x^S, E_\ell^{Fk} \mid x^{Fk}) = \tag{11} $$
$$ \Phi_2(\gamma_k^S, \gamma_\ell^{Fk}, \varrho_k) - \Phi_2(\gamma_k^S, \gamma_{\ell-1}^{Fk}, \varrho_k) - \Phi_2(\gamma_{k-1}^S, \gamma_\ell^{Fk}, \varrho_k) + \Phi_2(\gamma_{k-1}^S, \gamma_{\ell-1}^{Fk}, \varrho_k) $$
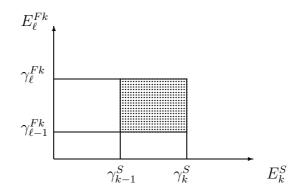
where Φ_2 is the bivariate standard normal distribution. The probability corresponds to the grey area in figure 1:

Thus, the log-likelihood for this model may be written as:

$$ \ln \mathcal{L} = \sum_{i=1}^{n} \sum_{k=1}^{K} \sum_{\ell=1}^{L_k} I_{ik\ell} \ln \left[Pr(E_k^S \mid x_i^S, E_\ell^{Fk} \mid x_i^{Fk}) \right] \tag{12} $$

---

15 See the general overview of bivariate and multivariate probit models in Greene (2000) pp.849 and p.927, Ronning (1991) pp.95 or Gouriéroux (1989) pp.86.

Figure 1: Joint probability of E_k^S and E_ℓ^{Fk}

where $Pr(E_k^S \mid x_i^S, E_\ell^{Fk} \mid x_i^{Fk})$ is defined such as in equation (11) for all individuals i, $i \in \{1...n\}$ and $I_{ik\ell}$ is an indicator variable equal to 1 if individual i obtained school degree k and post-school degree ℓ.

The estimation of this sequential model is not as straight-forward as that of the basic version in section 3, but remains reasonably feasible. As a matter of fact, since we have determined the expression of the log-likelihood function to be maximised (in equations (12) and (11)), we just need to write this function down and start the maximisation procedure, and we will be able to recover the parameters (the β's μ's and ϱ's). Using Stata 7.0, the maximisation is done numerically with the help of the Newton-Raphson iterative procedure[16]. This may be quite time-consuming.

4.2 An example

In this section, the decomposition of the educational career such as explained previously is applied to the GSOEP data, the purpose being to provide a simple illustration of how the model functions in practice on the basis of a widely used data set.

The ultimate educational outcome is, following the model, the result of two decisions: the first decision concerns attainment in general secondary education, and the second one is about post-school educational achievement. Here again, it is assumed that the educational outcome observed is the result of the decision of the individual given some constraints and the dependent variables are defined by the highest degree obtained in general secondary education and in post-school education respectively, as shown in table 3.

---

16 The principle of the Newton-Raphson procedure is the following: starting with a vector of initial values b_0, the algorithm calculates a new vector b_1 such that $b_1 = b_0 + g(-H)^{-1}$, where g is the gradient vector and H the matrix of second derivatives, then calculates $b_2 = b_1 + g(-H)^{-1}$ and repeats the procedure until convergence is achieved (see details in Gould and Sribney (1999)).

Table 3: Definition of dependent variables

General secondary education:		
E_1^S Hauptschule	E_2^S Realschule	E_3^S Gymnasium

Post-(general) secondary education of:		
Hauptschule-leavers	**Realschule-leavers**	**Gymnasium-leavers**
E_1^{F1} No voc. degree	E_1^{F2} No voc. degree	E_1^{F3} No higher education
E_2^{F1} Apprenticeship	E_2^{F2} Apprenticeship	E_2^{F3} Higher tech. college
E_3^{F1} Tech. college	E_3^{F2} Tech. college	E_3^{F3} University
	E_4^{F2} Higher tech. college/ university	

Attainment in general secondary education in Germany can be defined, using the GSOEP, by three levels E_k^S, with $k \in \{1...3\}$: the lowest category E_1^S is attributed to those individuals who obtained at most a degree from the Hauptschule, which provides pupils with basic general education. The intermediate category E_2^S consists of those people holding a degree from the Realschule, which provides pupils with a general education of intermediate level, and E_3^S is given to individuals who obtained a degree from the highest type of general secondary education, the Gymnasium, which qualifies for university studies (see a detailed description of the German education system in Lauer (2001)).

As far as post-school attainment is concerned, the definition of the dependent variable depends on the level attained in general secondary education. Since we have three possible attainments in secondary education, we have three different variables E_ℓ^{Fk} for post-school attainment, with $k \in \{1...3\}$, the number and the type of categories $\ell \in \{1...L_k\}$ of each depending on the value of k, i.e. on the type of secondary school degree possessed. In the end, we therefore estimate, through the maximisation of the log-likelihood function defined in equation 12, a model of four simultaneous equations (the secondary school equation on the one hand, and the post-Hauptschule, the post-Realschule and the post-Gymnasium equations on the other hand), with as many correlations ϱ_k, $k \in \{1...3\}$ between attainment in secondary education and subsequent achievement as there are secondary school degrees.

For those people having at most the Hauptschule degree, the dependent variable is E_1^{F1} if the person has no vocational degree, E_2^{F1} if the individual completed a basic vocational training (typically an apprenticeship but also in full-time vocational schools) and E_3^{F1} if the person has completed at least[17]

---

17 Only a few individuals with Hauptschule degree managed to get a higher education degree in the end, through the obtention of the vocational maturity certificate which entitles to lower tertiary studies.

an advanced vocational qualification (like the technical college or the master craftsman degree, but also the vocational maturity certificate). For the group of Realschule graduates, the dependent variable is in four ordered categories: no vocational degree (E_1^{F2}), a basic vocational degree (E_2^{F2}), an advanced vocational degree (E_3^{F2}) and a higher education degree[18]. Finally, the Gymnasium leavers may either have no higher education degree (E_1^{F3}), in which case they stopped their studies immediately after the obtention of the maturity certificate or only completed a vocational training, or hold a lower tertiary level degree (higher technical college - Fachhochschule - or equivalent, E_2^{F3}), or a university degree (E_3^{F3}).

One issue of interest is to investigate whether some factors play a more important role in the first stage, other factors in the second stage of the educational career. Therefore, the same variables on family background as in section 3 (parental education and father's occupation) have been included in all four equations. Moreover, we want to allow time-variant factors to influence educational choices. Ideally, we would like to have information on e.g. the financial situation of the household, on changes in the household structure etc. at the time when the individual decided upon his secondary education and when he decided upon his post-secondary education. Since this kind of information is not available, the approach adopted here was the following. It is assumed that the decision concerning secondary school choice is made at the end of primary schooling, at the age of 10, since this is the stage at which pupils are streamed into different types of secondary schools in Germany. Therefore, we can examine the effect of the context prevailing at that time by looking at the effect of finishing primary education in that year rather than in this other year. Concretely, a set of dummy variables indicating in which year the individual ended primary education, grouped in five-year primary school cohort dummies has been included in the secondary school equation. Similarly, it is assumed that the decision concerning post-secondary education has been made at the end of secondary education. Thus, we examine the effect of finishing secondary education at a specific point in time rather than at another point in time and include in the post-school equations a set of dummy variables indicating the year in which the person finished general secondary education, grouped in five-year secondary school cohorts. The computation of the year in which the individual ended secondary education is based on information on the type of secondary education completed and its duration.

Table 4 presents overall statistics and tests for the estimation. The overall Wald test at the bottom of the table tests the null hypothesis that all slope coefficients of all equations are jointly insignificant. The null hypothesis

---

18 Generally a lower tertiary degree obtained via the obtention of the vocational maturity certificate degree.

Table 4: Overall results and tests

			Equation					
	Secondary		Post-Hauptschule		Post-Realschule		Post-Gymnasium	
	χ^2	$p > \chi^2$	χ^2	$p > \chi^2$	χ^2	$p > \chi^2$	χ^2	$p > \chi^2$
Coefficients								
Wald test	1,481.89	0.00	369.15	0.00	126.77	0.00	159.71	0.00
Female * School cohort	45.32	0.00	35.03	0.00	12.04	0.10	9.61	0.21
Female * Mother's educ.	2.44	0.68	2.63	0.45	4.62	0.20	2.04	0.73
Female * Father's educ.	1.03	0.90	9.44	0.01	4.86	0.30	3.32	0.51
Female * Father's occ.	5.66	0.46	10.31	0.19	1.85	0.93	2.81	0.83
Thresholds	$t = S$		$t = F1$		$t = F2$		$t = F3$	
μ_1^t: Male = Female	4.11	0.04	77.91	0.00	24.97	0.00	4.76	0.11
μ_2^t: Male = Female	37.19	0.00	57.00	0.00	36.39	0.00	2.24	0.03
μ_3^t: Male = Female					38.61	0.00		
Males: $\mu_1^t = \mu_2^t$	1,001.96	0.00	215.99	0.00	22.72	0.00	31.04	0.00
$\mu_2^t = \mu_3^t$					23.39	0.00		
Females: $\mu_1^t = \mu_2^t$	1,206.35	0.00	144.64	0.00	28.83	0.00	18.81	0.00
$\mu_2^t = \mu_3^t$					27.69	0.00		
Correlation secondary/post-secondary			$k = 1$		$k = 2$		$k = 3$	
ϱ_k: Male = Female			0.30	0.58	0.45	0.50	0.05	0.82
Observations			5,857					
Log-likelihood			-10,021.67					
			χ^2	$p > \chi^2$				
Overall Wald test			1,842.89	0.00				

is rejected at a 1% significance level. Partial Wald tests on the coefficients of each equation separately are also all rejected at a 1% significance level and at-test that the slope parameters are jointly significant in all equations[19]. Gender differences in the coefficients as well as in the threshold values and in the correlations were also tested in the same way as in section 3. There are no significant (at a 10% significance level) gender differences in the slope coefficients with the exception of the school cohort variables in all equations (except for the post-Gymnasium equation), and of the coefficients of the father's education dummies in the post-Hauptschule equation. Similarly, there are no differences between males and females in the correlation between secondary school attainment and post-secondary education. However, all the threshold values differ significantly between men and women (though weaker for the second threshold to university education). In the final specification, only the significant interac-

---

19 The value of the χ^2-statistics, however, is all the lower since the level of education gets higher, which suggests that the variables used have more explanatory power for secondary school decision than for post-school decisions.

tions between the female dummy and the set of coefficients have been retained, as well as different thresholds for men and women, but one single correlation coefficient is estimated for males and females. The results of further tests on the statistical distinctness of successive threshold values show that the chosen educational categories do not need to be aggregated since all the resulting thresholds are statistically different, for men like for women.

Table 5: Determinants of secondary school attainment

Variable	coef.	(s.e)
Primary school cohort (ref.: 1939-1950)		
1951-55	0.28**	(0.08)
1956-60	0.39**	(0.08)
1961-65	0.42**	(0.08)
1966-70	0.57**	(0.07)
1971-78	0.49**	(0.07)
Female * 1951-55	-0.21†	(0.11)
Female * 1956-60	-0.28*	(0.11)
Female * 1961-65	0.24*	(0.11)
Female * 1966-70	0.20*	(0.10)
Female * 1971-78	0.34**	(0.10)
Mother's education (ref.: No voc. degree)		
Apprenticeship	0.27**	(0.04)
Tech. college/maturity	0.42**	(0.10)
Higher education	0.85**	(0.14)
Missing	0.05	(0.07)
Father's education (ref.: No voc. degree)		
Apprenticeship	0.21**	(0.06)
Tech. college/maturity	0.33**	(0.07)
Higher education	0.80**	(0.09)
Missing	0.01	(0.09)
Father's occupation (ref.: Worker)		
Farmer	0.10	(0.08)
Self-employed	0.65**	(0.06)
Senior manager	1.12**	(0.07)
Middle manager	0.68**	(0.05)
Employee	0.38**	(0.08)
Missing	0.28**	(0.05)
Thresholds		
μ_1^S Male	1.15**	(0.07)
Female	1.28**	(0.07)
μ_2^S Male	2.10**	(0.07)
Female	2.56**	(0.08)

Significance level : $\dagger$: 10% $*$: 5% $**$: 1%

The estimation results of the secondary school equation are reported in table 5 and the results of the three post-school equations in table 6. Observing table 5, it appears that there has been a strong educational expansion at the secondary school level. Indeed, the primary school cohort dummies all

Table 6: Determinants of post-secondary school attainment

Variable	Level of secondary school attained					
	Hauptschule		Realschule		Gymnasium	
	coef.	(s.e)	coef.	(s.e)	coef.	(s.e)
Secondary school cohort (ref.: 1945-1955)						
1956-60	0.34**	(0.10)	0.32*	(0.14)	0.34	(0.21)
1961-65	0.30**	(0.10)	0.38**	(0.15)	0.08	(0.20)
1966-70	0.24*	(0.11)	0.27†	(0.16)	-0.09	(0.19)
1971-75	0.23†	(0.12)	0.53**	(0.14)	-0.28	(0.19)
1976-80	0.36**	(0.11)	0.35*	(0.16)	-0.50**	(0.18)
1981-86	0.19	(0.13)	0.20	(0.15)	-0.68**	(0.17)
Female * 1956-60	0.28†	(0.15)	0.09	(0.20)		
Female * 1961-65	0.29†	(0.15)	-0.02	(0.22)		
Female * 1966-70	0.79**	(0.15)	0.56**	(0.22)		
Female * 1971-75	0.57**	(0.15)	0.44*	(0.19)		
Female * 1976-80	0.69**	(0.14)	0.62**	(0.19)		
Female * 1981-86	0.79**	(0.18)	0.62**	(0.18)		
Mother's education (ref.: No voc. degree)						
Apprenticeship	0.26**	(0.05)	0.24**	(0.06)	-0.08	(0.08)
Tech. college/maturity	0.47**	(0.15)	0.52**	(0.12)	-0.17	(0.15)
Higher education					0.09*	(0.04)
Missing	0.00	(0.07)	0.20†	(0.10)	0.00	(0.18)
Father's education (ref.: No voc. degree)						
Apprenticeship	0.20**	(0.08)	0.17†	(0.09)	0.06	(0.14)
Tech. college/maturity	0.54**	(0.12)	0.28*	(0.11)	0.09*	(0.03)
Higher education	1.20**	(0.25)	0.70**	(0.16)	0.41*	(0.19)
Missing	0.13	(0.11)	-0.02	(0.13)	-0.46*	(0.22)
Female * Apprenticeship	0.29**	(0.10)				
Female * Tech. coll./mat.	0.10	(0.15)				
Female * Higher educ.	-0.24	(0.30)				
Female * Missing	0.19	(0.13)				
Father's occupation (ref.: Worker)						
Farmer	0.12	(0.08)	0.38**	(0.14)	0.03	(0.19)
Self-employed	0.50**	(0.09)	0.46**	(0.12)	0.32*	(0.15)
Senior manager	0.68**	(0.16)	0.87**	(0.17)	0.39*	(0.16)
Middle manager	0.51**	(0.08)	0.53**	(0.11)	0.21	(0.14)
Employee	0.32**	(0.10)	0.39**	(0.11)	0.16	(0.19)
Missing	0.10†	(0.06)	0.19*	(0.08)	0.07	(0.12)

to be continued...

| | | Level of secondary school attained | | | | | |
| | | Hauptschule | | Realschule | | Gymnasium | |
		coef.	(s.e)	coef.	(s.e)	coef.	(s.e)
Thresholds		$k = 1$		$k = 2$		$k = 3$	
μ_1^{Fk}	Male	-0.69**	(0.08)	-0.28	(0.57)	2.31**	(0.33)
	Female	0.41**	(0.09)	0.56*	(0.32)	2.47**	(0.28)
μ_2^{Fk}	Male	1.15**	(0.14)	1.00**	(0.32)	2.49**	(0.35)
	Female	2.27**	(0.21)	1.89**	(0.24)	2.74**	(0.30)
μ_3^{Fk}	Male			1.82**	(0.18)		
	Female			2.74**	(0.14)		
Correlations		$k = 1$		$k = 2$		$k = 3$	
ϱ_k		0.64**	(0.15)	0.70**	(0.18)	0.86**	(0.10)

Significance level :   † : 10%   ∗ : 5%   ∗∗ : 1%

prove extremely significant for the determination of attainment in secondary education and exhibit an increasing pattern[20]. In other words, the net utility of a high level secondary school degree has increased over time. Judging from the coefficients of the interaction terms, educational expansion started later on for women. Until the cohorts ending primary education the 1950s, the increase in the utility of a higher general education used to be stronger for men. Then, however, educational expansion has been much stronger for women than for men, and the gap in the rate of expansion has increased over time. In the post-school equations (table 6), the pattern is more differentiated. Among those people holding as the highest general secondary school degree the Hauptschule degree, those who obtained their degree later have attained a higher level of subsequent education. Female Hauptschule-graduates have experienced a particularly strong educational expansion since the Hauptschule cohorts of the end of the late 1950s, much stronger than men, especially since the late 1960s Hauptschule cohorts. A similar pattern is observable for Realschule-degree holders, even though the stronger educational improvement for women started for the late 1960s Realschule cohorts only. For the Gymnasium leavers, however, their has been a decline in subsequent educational attainment. This means that the propensity to study among those entitled to do so has declined over time. This phenomenon starts becoming significant for those cohorts who obtained their maturity certificate from the middle of the 1970s onwards.

Mother's education is strongly significant for achievement at school, but also

---

20 The overall impact of belonging to a specific cohort on the educational outcome is obtained by adding the coefficient of the cohort dummy to that of the interaction term between that same cohort and the gender dummy.

for post-Hauptschule and post-Realschule achievement[21]. The coefficients have the same order of magnitude in the post-Hauptschule and in the post-Realschule equation. Thus, the higher the education of the mother is, the lower the cost to return ratio for the children and the higher the educational prospects of the children in general and vocational education are. At a higher level of secondary education, however, the influence of mother's education attenuates strongly. Mother's education has virtually no influence for Gymnasium leavers, except when the mother is a university graduate, but even there, the effect is small.

Father's education is also a strong determinant of secondary school achievement. As far as post-school education is concerned, it depends on the level of general education attained. The influence of father's educational attainment is all the weaker since the level attained in general education is high. Thus, the impact is highest for Hauptschule degree holders, it is less strong for Realschule leavers and very weak for maturity certificate holders. For the latter category, father's education has more effects than mother's education, especially in case the father has a higher education degree. The interaction terms in the post-Hauptschule equation show that the educational attainment of the father seems to have slightly more importance for women: the fact that the father has completed an apprenticeship improves much more the prospects in vocational education for females than for males, but the other interaction terms are not significant.

Like for secondary school achievement, sons and daughters of senior managers have by far the best prospects also with respect to post-school education. The effect is strongest for secondary education, but it also particularly strong for post-Realschule achievement. Overall, the impact of father's occupation is weakest for holders of the maturity certificate. In this group, only individuals with a senior manager or a self-employed as a father achieve significantly better than the reference category of workers' offsprings. For all other groups, also sons and daughters of middle staff managers and to a lesser extent of employees achieve better than workers' offsprings. The group of the Realschule-graduate is the only one for which having a farmer as a father is recorded to have a positive effect on subsequent education compared to the reference category.

The threshold values are significantly higher for females than for males at all educational levels. This implies that for the same characteristics and coefficients, the cost to return ratio will be higher for women than for men and the educational prospects poorer. Moreover, even with this very simple

---

21 As far as the education level of the mother is concerned, the too small proportion a mothers with a higher education degree made it necessary to aggregate the categories "technical college/maturity" and "higher education" into one single category for the post-school attainment of Hauptschule and Realschule degree holders.

specification in terms of variables included, the correlations between the secondary school equation and the post-secondary education equations all prove extremely significant[22]. This implies that the decisions at the different stages are not independent and that a purely sequential model - or a model focussing on one specific transition within the educational career in isolation from previous transitions - would be inappropriate. Therefore, the additional computational burden resulting from a time-intensive maximisation procedure seems to be worth it[23].

5 Conclusion

This article develops a model of educational attainment which is imbedded in the human capital theory, the reference theory for the economic analysis of issues related to education. The principle of this model is that there is a finite number of possible educational alternatives which may be ordered by level. Each educational alternative yields a certain utility to the individual. The net utility of an educational alternative is expressed in terms of the difference between the "returns" and the "costs" associated with this alternative for any individual with specific characteristics and constraints. The costs and the returns may be non-pecuniary and can be decomposed into one education-specific component, one individual-specific component and one random component. The individual chooses, given his specific constraints or characteristics, to attain the education level which maximises his net utility.

In a first step, the formalisation of these assumptions leads to an ordered probit model where the threshold values are given by the expected ratio of the marginal costs to the marginal returns. In other words, the individual assesses the marginal costs and the marginal returns associated with the next higher education level, and if the ratio of the former to the latter, given his characteristics, is below a certain threshold, he will opt for the next higher level. In a second step, the educational process is decomposed into two stages: achievement at school and post-school achievement. The same conceptual framework applies at both stages, but a correlation is permitted between them. Empirically, this boils down to estimate a multivariate ordered probit model, where we have one equation for secondary education and as many post-secondary education equations as there are secondary school degrees. The model makes it possible to investigate whether some factors exert a different influence depending on the stage in the educational process, and whether time-variant factors

---

22 A speculation of the sign of the correlation between the error terms of the different would be too hazardous since we do not know exactly what these unobserved factors are.

23 If the correlations had proved to be insignificant, the estimations should better be run separately, because it would be more efficient in the sense that fewer parameters are estimated.

also play a role.

A simple example shows how this model functions in practical terms, using data from the German Socio-Economic Panel, a widely used data set in German socio-economic research. The theoretical model can be estimated directly by maximising the log-likelihood function derived theoretically. Final educational attainment has been defined by five ordered categories. The explanatory variables used are the traditional ones encountered in the literature on educational attainment: cohort dummies and variables depicting family background. Differences across genders in the parameters have been systematically tested. Little variation has been found in the coefficients for males and females, except for cohort dummies. The thresholds have also been allowed to differ between males and females, which is not usual in empirical studies but proved extremely useful. As a matter of fact, the threshold values proved systematically higher for women than for men, which means that for the same characteristics and coefficients, the (perceived) cost to return ratios are higher for women and their final educational attainment poorer.

For the application of the decomposition of the educational process into two steps, attainment in general secondary education has been defined by three levels, and we have therefore three further equations for post-school educational attainment, one for each secondary school degree. Overall, the family background variables play a more important role for achievement at school than for post-school education. Educational expansion proved particularly strong at the level of secondary education, especially for women, but not at the post-maturity level, where the propensity to undertake tertiary level studies has decreased over generations, all other things equal. Thus, there might be different effects at work depending on the stage observed within the educational career. The three estimated correlations between school achievement and post-school education all proved extremely significant. This points to the usefulness, even with this very simple specification in terms of variables included, of allowing non-zero correlations, which would have been the implicit assumption if the two stages had been estimated separately. The model promises to be even more useful with a more sophisticated specification, in particular if more time-variant variables depicting the costs or the returns of education are included, provided these are available in the data set used. Overall, the model formulated here has the advantage of being directly estimable empirically, while providing an economic interpretation framework consistent with the human capital theory.

References

Amemiya, T. (1986). *Advanced Econometrics* (1985 ed.). Oxford: Basil Black-well Ltd.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education.* New York: Columbia University Press.

Becker, G. S. (1965). A theory of the allocation of time. *Economic Journal, 75*, 493-517.

Becker, G. S. and Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of Labor Economics, 4*(3), 1-39.

Blau, P. M. and Duncan, O. D. (1967). *The American occupational structure.* New York: Wiley.

Boudon, R. (1974). *Education, Opportunity and Social Inequality.* New York: Wiley.

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

Ermisch, J. F. and Francesconi, M. (2001). Family Matters: Impacts of Family Background on Educational Attainments. *Economica, 68*, 137-156.

Gould, W. and Sribney, W. (1999). *Maximum likelihood estimation with Stata.* College Station: Stata Press.

Gouriéroux, C. (1989). *Econométrie des variables qualitatives* (2 ed.). Paris: Economica.

Greene, W. H. (2000). *Econometric analysis* (4 ed.). London: Prentice Hall International.

Haisken-DeNew, J. and Frick, J. (2000). *Desktop Companion to the German Socio-Economic Panel (GSOEP) Study.* Berlin.

Hanushek, E. A. (1992). The Trade-off between Child Quantity and Quality. *Journal of Political Economy, 100*(1), 84-117.

Haveman, R. and Wolfe, B. (1995). The Determinants of Children's Attainments: A Review of Methods and Findings. *The Journal of Economic Literature, 33*(4), 1829-1878.

Heckman, J. J. (1976). A life-cycle model of earnings, learning and consumption. *Journal of Political Economy, 84*(4), 11-44.

Kodde, D. A. (1988). Unemployment Expectations and Human Capital Formation. *The European Economic Review, 32*(8), 1645-1660.

Lauer, C. (2001). *Educational Attainment: A French-German Comparison* (Documentation No. 01-02). ZEW.

Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics.* Cambridge: Cambridge University Press.

Manski, C. F. and Wise, D. A. (1983). *College Choice in America.* Cambridge, Mass.: Harvard University Press.

Mare, R. D. (1980). Social Background and School Continuation Decisions. *The Journal of American Statistical Association, 75*(370), 295-305.

Mincer, J. (1974). *Schooling, experience, and earnings.* New York: National Bureau of Economic Research.

Müller, W. (1990). *Does Education Matter? Evidence from Cross-national Comparisons.* Stanford: Center for Advanced Study in the Behavioral Science.

OECD. (2000). *Education at a Glance: OECD Indicators* (2000 ed.). Paris: Organisation for Economic Co-operation and Development.

Ronning, G. (1991). *Mikroökonometrie.* Berlin: Springer Verlag.

Rosenzweig, M. R. and Wolpin, K. I. (1993). Intergenerational Support and the Life-Cycle Incomes of Young Men and their Parents: Human Capital Investments, Coresidence and Intergenerational Financial Transfers. *Journal of Labor Economics, 11*(1), 84-112.

Shavit, Y. and Blossfeld, H.-P. (1993). *Persistent Inequality - Changing Educational Attainment in Thirteen Countries.* Boulder: Westview Press.

Taubman, J. (1976). The determinants of earnings: genetics, family and other environments: a study of white male twins. *American Economic Review, 66*, 858-870.

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

Wilson, K. (2001). The Determinants of Educational Attainment: Modeling and Estimating the Human Capital Model and Education Production Functions. *Southern Economic Journal, 67*(3), 518-551.

Appendix

Table 7: Descriptive statistics for the sample of the simultaneous estimation

	Variable	Percent
Secondary education	Hauptschule	56.95
	Realschule	27.74
	Gymnasium	15.32
Final education	No vocational degree	24.07
	Apprenticeship	45.42
	Technical college or maturity	17.15
	Higher technical college	4.26
	University	9.09
Post-school education of:		
Hauptschule-leavers	No vocational degree	27.82
	Apprenticeship	60.60
	Technical college or maturity	10.28
	Higher education	1.31
Realschule-leavers	No vocational degree	8.05
	Apprenticeship	46.55
	Technical college or maturity	32.44
	Higher education	12.96
Gymnasium-leavers	No higher education	28.08
	Higher technical college	9.03
	University	62.89
Sex	Male	50.04
	Female	49.96
Birth cohort	1929-38	22.95
	1939-48	23.75
	1949-58	23.07
	1959-68	30.23
Primary school cohort	1939-50	29.48
	1951-55	10.58
	1956-60	11.69
	1961-65	10.83
	1966-70	13.21
	1971-75	15.00
	1976-78	9.22

to be continued...

	Variable	Percent
Secondary school cohort	1945-50	13.58
	1951-55	14.39
	1956-60	10.65
	1961-65	11.16
	1966-70	10.48
	1971-75	11.53
	1976-80	15.48
	1981-86	12.73
Mother's education	No vocational degree	43.98
	Apprenticeship	36.66
	Technical college or maturity	3.39
	Higher education	1.16
	Missing	14.80
Father's education	No vocational degree	12.39
	Apprenticeship	52.88
	Technical college or maturity	12.63
	Higher education	6.90
	Missing	15.20
Father's occupation	Worker	35.62
	Farmer	5.57
	Self-employed	7.95
	Senior manager	10.41
	Middle manager	13.34
	Employee	4.07
	Missing	23.03