Chapter 6 – The return to education

1. Introduction

The discussion of educational investment presented in the second chapter was based on the assumption that increasing education was associated with an increase in potential (permanent) income. While this remains true from an individual perspective (that is, families take this fact as a stylised fact, as in any partial equilibrium model), in this chapter we wonder why this may be true in the aggregate. We face competing explanations here, and will review them in turn, starting from the aggregate evidence supporting the view that human capital raises firms' productivity (in line with the human capital approach). We then proceed with a more recent view according to which education is associated with non-cognitive abilities. Next, we present the credentialist approach, where the education is just a signal of unobservable ability. The remaining part of the chapter is devoted to the problem of measuring the return to education, both from theoretical and econometric perspectives.

2. The productivity of human capital

While the positive correlation between education and earnings at individual level is one of the most established facts in economic literature, the existence of a causal relation between the two is not yet widely accepted. The strongest doubts come from the consideration that both earnings and schooling could both depend on additional factors, not observed by the researcher, thus constituting a patent case of spurious correlation. An incomplete list of unobserved factors that could affect both variables include parental education (children education is favoured by educated parents, offering wider cultural stimuli, helping with homework, selecting better schools, and so on – moreover, better educated parents are inserted in higher rank social networks, thus offering better job opportunities to their children), behavioural traits (self-consciousness and self-esteem are rewarded both during educational career and in the labour market), school quality based on peer effects (attending a school with brighter class mates raises educational attainment, meanwhile creating social networks that may be helpful when entering the labour market), discrimination (ethnic or gender discrimination can prevent access to the best positions in highly ranked colleges and in top paid jobs, sometimes leading to self-fulfilling expectations).[1]

However we now possess a large body of evidence that, despite all unobservable differences, education still plays a causal role in earnings determination, even if standard Mincerian regressions (where earnings are regressed onto education) do not account for more than one third of the observed variance. One of the most convincing pieces of evidence comes from studies on identical twins.[2] Twins are in general identical in terms of family background; in addition fraternal (monozygotic) twins are biologically identical. Under the debatable (and debated) assumption that intelligence is genetic, and that intelligence represents the bulk of unobservable ability, one can take the difference in educational attainment between twins and use it as a regressor measuring its effect onto earning differences. By analysing a sample of Australian twins, 60% of which are monozygotic, Miller et alt. (1997) compare the estimated return obtained from monozygotic twins with analogous estimates from dizygotic ones. Suppose that earnings are determined in their general form as follows

[1] There are several reviews of this literature on the return to education, and particularly on the causal effect of education on earnings: among the most recent ones see Card 1999 and 2001, Bowles; Gintis and Osborne 2001a; Harmon, Oosterbeek and Walker 2003. Earlier references in Willis 1986.

[2] See Ashenfelter and Rouse 1998 and 2000 for the analysis of an American sample of twins, and Miller, Mulvey and Martin 1997 for an Australian one. Additional references in Harmon, Oosterbeek and Walker 2003.

$$w_{ij} = \alpha_0 + \alpha_1 A_{ij} + \alpha_2 H_{ij} + \alpha_3 X_{ij} + \beta S_{ij} + \nu_{ij} \tag{6.1}$$

where w_{ij} are earnings of twin $j, j = 1,2$ in family i, A_{ij} is her (unobservable) ability, H_{ij} measures family background (education, income, socio-economic status), S_{ij} is her schooling and X_{ij} represents additional information (like gender, age, experience). ν_{ij} indicates the error term. By taking the difference between twins

$$w_{i1} - w_{i2} = \alpha_1(A_{i1} - A_{i2}) + \alpha_2(H_{i1} - H_{i2}) + \alpha_3(X_{i1} - X_{i2}) + \beta(S_{i1} - S_{i2}) + (\nu_{i1} - \nu_{i2}) \tag{6.2}$$

If the twins are monozygotic, then $(A_{i1} - A_{i2}) = 0$ and plausibly also $(X_{i1} - X_{i2}) = 0$. In addition, if they are reared together, also $(H_{i1} - H_{i2}) = 0$, then least square projection of $(w_{i1} - w_{i2})$ onto $(S_{i1} - S_{i2})$ yields an unbiased estimate of the true impact of education on earnings β. When the same procedure is applied to dizygotic twins, the estimates of β will be biased by the omission of the unobservable difference in ability, the extent of which can be assessed by comparing with the estimate obtained for monozygotic twins.[3] When we take into account the presence of unobservable family characteristics, as in the case of Ashenfelter and Rouse (1998),[4] and we introduce the additional assumption of potential correlation between family's ability and children's schooling, the return of education can be estimated from

$$\begin{cases} w_{i1} = \alpha_0 + \alpha_2 H_{i1} + \alpha_3 X_{i1} + \beta S_{i1} + \gamma \dfrac{S_{i1} + S_{i2}}{2} + \nu_{i1} \\ w_{i2} = \alpha_0 + \alpha_2 H_{i2} + \alpha_3 X_{i2} + \beta S_{i2} + \gamma \dfrac{S_{i1} + S_{i2}}{2} + \nu_{i2} \end{cases} \tag{6.3}$$

where this reduced form is characterised by correlated random effects.[5] The additional advantage of using twins data is that in most cases they allow the assessment of the measurement error for schooling, by comparing the one's recall of her schooling experience with the other twin's recall of the same experience.[6]

---

[3] Hence the conclusion from the traditional twins model is that the estimated return to schooling for males of 7.1% is comprised of 2.3% due to the 'true' returns to schooling, 4.2% due to the effect of family background and 0.7% due to the influence of genetic factors (Miller et al. 1997, p.130). Similar conclusions in Ashenfelter and Rouse 2000: "Although part of the correlation between income and schooling may be due to family background characteristics, the intrafamily correlation between income and schooling indicates that most of the relationship between income and schooling is due to something else". They report an ability bias in the order of one fourth of the estimated return of 8% per year, which compensates almost exactly for measurement errors in schooling.

[4] This is also consistent with the findings by Cameron and Heckman 2001: "It is the long-run factors that promote scholastic ability that explain most of the measured gap in schooling attainment and not the short run credit constraints faced by students of college going age that receive most of the attention in popular policy discussions." (Cameron and Heckman 2001, p.490).

[5] Ashenfelter and Rouse 1998 estimate a rate of return between 10% and 12% using individual estimates, and a rate between 7% and 10% using differences between twins. The estimated correlation between family ability and returns (γ coefficient) is found significant, but is sometimes positive and other times negative. They take the negative occurrence as evidence of the compensatory role of schooling in reducing the impact of natural ability.

[6] Mulvey et alt. 1997 describe that the correlation between self reported education and sibling's report is about 0.7, suggesting the presence of significant measurement errors, that lower by almost one percent point the return to education. See section 7 below for further discussion.

The main objection raised against twins analysis is the small sample size; in addition some of these samples were collected for other purposes (typically medical) and could in principle contain systematic biases. However two other pieces of evidence can be invoked in order to sustain that education has a causal impact on earnings independent of unobservable ability. The first one comes from instrumental variable estimates, while the second from the more recent literature on natural experiments. From an ideal point of view, in order to assess the existence of a causal role for education one would like to randomly allocate individuals to acquire additional education, and then measure income differences subsequently. However ethical and political considerations advise against preventing some children from education acquisition for scientific research purposes, and therefore researchers have to resort to either relative comparisons or to exogenous variations attributable to events that are beyond the control of individual subjects.

The approach based on instrumental variables crucially hinges on the ability of the researcher to identify a variable (or a group of variables) that are correlated with the educational choice (which is likely to be correlated with unobservable ability) but not with earnings. If this variable exists, then it is possible to obtain unbiased estimates of the return to education. In symbols, we replace previous equation (6.1) with a system of two equations

$$w_i = \alpha_0 + \alpha_1 X_i + \beta S_i + \nu_i \qquad (6.1)'$$

$$S_i = \gamma_0 + \gamma_1 X_i + \gamma_2 Z_i + \varepsilon_i \qquad (6.4)$$

where Z_i represents the potential instrument affecting education but not earnings.[7] The return to education can then be estimated by replacing the observed education S_i with the predicted value $\hat{S}_i$ obtained from equation (6.4), since there is no reason to expect this variable to be correlated with the error term in equation (6.1)'.

Several instruments have been proposed in the literature, and we will recall here only the most important ones.[8] A first candidate for the Z_i instrument has been parental education: it has a close correlation with the educational attainment of children, and in principle should be uncorrelated with their future earnings. However, where social networks are an important channel of access to the labour market, better educated parents may prove crucial in favouring the entrance to better jobs (which are obviously better paid as well). A second candidate that has been used is school (college) proximity, since students living nearby face lower costs of attendance (in terms of transportation and/or living away from home). This represents a better alternative, although one could always devise potential correlation with earnings (cities with colleges experience an extra supply of graduates, which may depress the local rate of return to college). A third alternative exploits the birth quarter of the individual, on the argument that people born at the beginning of the year reach compulsory leaving age earlier, and have therefore less incentives to proceed further in education. While this in principle represents purely exogenous variation, it has been challenged on the ground that the season of birth may be correlated with family background.[9]

---

7 However the list of individual characteristics X_i must include all potential variables that may affect both education and earnings (like cognitive ability, motivation, and so on). If this is not the case the least square estimate will remain biased, as the schooling coefficient β will capture some of the effects that would otherwise be attributed to the omitted variable. See Harmon, Oosterbeek and Walker 2003.

8 More complete surveys are in Card (1999) and Harmon, Oosterbeek and Walker (2003).

9 See Card 2001.

The last instrument points to compulsory education legislation as a source of exogenous variation. Effectively, several papers have utilised educational reforms as an instrument to estimate the return to education, finding a consistent pattern of higher estimated returns (in the order 20% greater return than ordinary least squares).[10] Even if also in this case one can raise doubts against this selection of instrument,[11] a more relevant problem arises when we question the assumption of constant return in the population. So far we have taken for granted the existence of a single measure of return, identical for all the population. When we take into account potential heterogeneity in the return to education, least square estimates (whether biased or unbiased) provide a measure of the average return to education in the population. In such a framework, the instrumental variable estimations are to be reinterpreted as a measure of the causal effect of education on the subgroup of population that was really affected by the educational reform.[12]

In all attempts to provide an unbiased estimate of the return to education (either using twins or using an exogenous source of variation) we find a consistent result: *differences in education explain differences in earnings even accounting for unobserved difference in abilities, and these outcomes can be taken as pieces of evidence of a productivity enhancing effect of schooling.*[13] Nevertheless, even leaving aside the problem of measuring the relative contribution of unobservable ability, the estimated return provides a measure of the private return of education, i.e. the increase in productivity converted into earnings for the holder of the additional education. Yet we should not forget that when one considers the optimal investment in education (as in chapter 2) we take the return to education for given, as in any partial equilibrium analysis. But when we aggregate individual choices, we should not forget that in a Walrasian perspective aggregate prices (and returns) reflect relative scarcity, and could be partially unrelated with real productivity.[14] Cyclical fluctuations of returns to education can be observed independently from technological shocks, and can be attributed to both supply and demand shifts that are unrelated to changes in individual productivity.[15] Acemoglu (1999) and (2003) provide an elegant model to account for different dynamics observed in wage inequality in United States and Europe: a skill biased

[10] Trostel, Walker and Wolley 2002 find similar order of magnitude comparing IV and OLS estimates in a sample of 25 countries over the period 1985-95, using parental and spouse education as instruments.

[11] "These results suggest that changes in the institutional structure of the educational system can affect the mapping between individual ability and educational outcomes, leading to a violation of assumptions such as independence or homoskedasticity needed for a conventional IV (instrumental variable) estimator to yield a consistent estimate of the average marginal return to education." (Card 2001, p.1140 – see also the evidence reported in table 1).

[12] "When the instrument is formed on the basis of membership of a treatment group the IV estimate of the return to schooling is the difference in expected log earnings between the control group and the treatment group, divided by the difference in expected schooling for the two groups. This implies that if all individuals in the population have the same marginal return the IV estimate is a consistent estimate of the average marginal rate of return. However, if the return to schooling is allowed to vary across individuals the IV estimate a weighed return, where the weights reflect the extent to which the subgroup is affected by the treatment or instrument. If only one subgroup is affected by the intervention, the IV estimator will yield the marginal rate of return for that subgroup." (Harmon, Hogan and Walker 2003, p.143). This is known in the literature as the LATE (local average treatment effect) estimator. It represents a possible explanation of the IV estimates exceeding those obtained under OLS: see the discussion on this issue in Card 2001. Harmon, Hogan and Walker 2003 consider explicitly the possibility of heterogeneity of returns, using a random coefficient estimator, without finding significant differences between these estimates and standard least square ones.

[13] In the same vein of natural experiments, a growing literature on randomised experiments in educational resources shows that earnings are affected by schooling experiences of agents. For a cursory but enlightening review of this literature see Krueger 2002.

[14] Freeman 1986 presents a nice example of the relevance of this problem with respect to the market for engineers in United States, applying the cobweb model of delayed delivery to the market: when returns to engineering degrees are high, there is a strong incentive to take that subject, but as long as new graduates enter the market, the relative return for this choice declines. The next generation of students abstains from engineering, and the return rises again.

[15] Murphy and Welch 1992 combine supply factors (changes in age and education structure of the population) and demand factors (increase in import competition from countries endowed with low skill workforces) to account for the relative wage of skilled workers in US during the 70's.

technological (like the introduction of computers in production) should raise the relative return to education, under the assumption that new technologies are complementary to education (since educated workers are more able in managing software driven activities). This represents a genuine increase in productivity that converts into earnings, and is actually observed in the United States. However, if labour market institutions compress wage distribution, as in Europe, firms will have an incentive to adopt technologies that raise the relative wage of the unskilled, thus reducing the return to education.[16] Alternatively, if the supply response is sufficiently quick, the greater availability of skilled labour may offset the effect of technological change, leaving the return to education unaffected.[17]

The fact that private returns to education incorporate market equilibrium effects that are independent of individual productivity works against the claim that human capital is productive *per se*. Nevertheless, further supportive elements can be found in the literature on education and growth.[18] While a consolidated practice among growth theorists consists of controlling for initial conditions with various measures of educational attainment in the population (see Barro 1997), more recently greater attention has been paid to the implications of what Krueger and Lindhal (2001) call the "macro-Mincer" model.[19] By replicating equation (6.1) in a different context, let us express the determinants of earnings of individual i working in country j at time t

$$w_{ijt} = \alpha_0 + \alpha_1 A_{ij} + \alpha_2 H_{ijt} + \alpha_3 X_{ij} + \beta S_{ijt} + \nu_{ijt}, \qquad i = 1,..,n;\ j = 1,...,m;\ t = 0,...,T \qquad (6.5)$$

By aggregating across individuals within a country we obtain

$$\begin{aligned} \overline{w}_{jt} &= \sum_{i=1}^{n} w_{ijt} = \alpha_0 + \left(\alpha_1 \overline{A}_j + \alpha_3 \overline{X}_j\right) + \alpha_2 \overline{H}_{jt} + \beta \overline{S}_{jt} + \nu_{jt} = \\ &= \eta_{0j} + \eta_1 \overline{H}_{jt} + \beta \overline{S}_{jt} + \nu_{jt}, \qquad j = 1,...,m;\ t = 0,...,T \end{aligned} \qquad (6.6)$$

Equation (6.6) suggests an additional implication of the productive role of human capital at the macro level: average[20] labour earnings must correlate with average educational attainment; this proposition can be estimated using cross-sectional data at country level, including a country fixed effect (η_{0j}) and other covariates to capture the country wealth or income ($\overline{H}_{jt}$). With additional assumptions regarding the feature of the aggregate production function (and specifically the presence of constant labour share), equation (6.6) can be easily converted into a growth equation, where the current level of output is taken as a function of the current stock of human capital, its dynamic equivalent being that the output growth rate is a function of the change in average educational attainment.[21] The empirical analysis has found

[16] "Put differently, the labour market institutions that push the wages of these workers up make their employers *the residual claimant* of the increase in productivity due to technology adoption, encouraging the adoption of technologies complementary to unskilled workers in Europe." (Acemoglu 2003, p.128f). The alternative implication is an increase in relative unemployment, also observed for unskilled workers in Europe.

[17] One implication of this perspective is the negative correlation of measured return to education and average educational attainment. Similarly, if workers with various levels of education are perfect substitutes, earnings inequality is also negatively associated with average educational attainment. Empirical evidence partially contradicts these propositions: see Bils and Klenow 1998 and Teulings and van Rens 2002.

[18] Recent surveys of this literature are in Krueger and Lindhal 2001 and in Sianesi and Van Reenen 2003.

[19] "An attractive feature of Mincer's model is that time spent in school (as opposed to degrees) is the key determinant of earnings, so data on years of schooling can be used to estimate a comparable return to education in countries with very different educational systems." (Kruger and Lindhal 2001, p.1103).

[20] If w represents log-income (as in the standard Mincer model), then $\overline{w}$ represents its geometric mean.

[21] Both propositions have been tested in the modern theory of growth. Lucas 1988 considers the stock of human capital of the representative agent as an input of the aggregate production function, assumes that human capital grows according to

that the initial stock of human capital shapes future growth paths, and that secondary education is more important than primary.[22] Krueger and Lindhal (2001) analyse several problems connected with the absence of significance of changes in human capital onto output growth, including measurement errors for human capital (reliability ratios[23] are obtained using information from different sources), length of time horizon, heterogeneity of impact across country, potential endogeneity bias and linearity assumption. Under some specifications the estimated impact of human capital on growth (coefficient β) is much higher than the private return to education, ranging from 18% to 30%.[24] If the return to education estimated at the aggregate level exceeds the corresponding estimate based on individual information, this represents a clue of the potential existence of externalities, i.e. beneficial spill-over effects deriving from individual choices.[25] However, aggregate return does not necessarily measure the social rate of return, since this latter concept includes other externalities that may not necessarily affect output growth, despite being important from a policy point of view: think of greater educational attainment being positively correlated with lower crime rates, reduced welfare dependence, better public health, better parenting, wider political participation and greater social cohesion (Blöndal, Field and Girouard 2002).

Overall, this section has put forward several pieces of evidence supporting the view that human capital raises individual productivity, possibly introducing positive externalities on aggregate output. In both micro and macro analyses human capital has been measured by years of education, thus implying that school attendance has beneficial effect *per se*. However the true reason why staying at school should increase productivity has not yet been well understood, and this subject will be discussed in the next section.

---

the share of time devoted to education and derives the steady state conditions under which human capital and output grow at the same rate. Romer 1990a includes the stock of human capital as a productive input, but the technical progress expands the variety of commodity at a rate that is positive function of the existing stock of human capital. Thus the two classes of models originated by these seminal models have different empirical implications: the growth of human capital (i.e. enrolment rates) in Lucas model should affect output growth, while the stock of human capital should affect growth in the Romer model. See Aghion and Howitt 1998 for details and Gemell 1996 fur further discussion.

[22] Romer 1990b uses literacy rates as a proxy for the stock of human capital, finding that its initial level (but not its change) affects subsequent growth. Gemell 1996 distinguishes between educated labour force and educated population, dividing education into three levels; he finds a positive impact of initial stock of human capital on output growth, and a positive correlation between investment in fixed capital (relative to output) and school enrolment rates. Primary education seems crucial for the poorest LDC, while secondary education affects output growth in intermediate LDC and tertiary education as an impact only for OECD countries.

[23] As in the case of imperfect recall across twins, consider the case of possessing two noisy measures of true schooling S^*, $S_1 = S^* + \varepsilon_1$ and $S_2 = S^* + \varepsilon_2$, where $\varepsilon_i, i = 1,2$ are measurement errors. If ε_1 and ε_2 are uncorrelated, the fraction of variability in S_1 due to measurement error can be estimated as $R_1 = \frac{Cov(S_1, S_2)}{Var(S_1)}$ (reliability ratio).

[24] "This is an enormous return to investment in schooling, equal to three or four times the private return to schooling estimated within most countries. The large coefficient on schooling suggests the existence of quite large externalities from educational changes (Lucas 1988) or simultaneous causality in which growth causes greater educational attainment." (Krueger and Lindhal 2001, p.1120). However their final conclusion is much more sceptical: "The macro-economic evidence of externalities in terms of technological progress from investments in higher education seems to us more fragile, resulting from imposing constant-coefficient and linearity restrictions that are rejected by the data." (p.1130).

[25] Sianesi and Van Reenen 2003 discuss the issue at length, but they also are rather cautious about the size of the externality effect: "We join Topel (1999) – 'the magnitude of the effect of education on growth is vastly too large to be interpreted as a causal force' – in finding too hard to view such huge effects as uniquely the result of economy-wide externalities generated by the increase in average educational attainment" (p.188).

3. Effort enhancing preferences

If education has a productivity enhancing effect, this can occur either because it provide students with know-how that will prove important once inserted in a productive process or because it reveals information on the students' abilities. In principle we can partially control for both of these channels, inserting in an earning function like (6.1) proxy measures of specific skills learnt at school (like specific subjects taken during high school or when in college) or test measures of general intelligence (like IQ test, AFQT scores and the like).[26] Nevertheless, Bowles, Gintis and Osborne (2001a) report a meta-analysis on empirical estimates of returns to education, where some measure of cognitive ability was included among the regressors. If the effect of schooling on earnings was due to acquired abilities, by explicitly accounting for them should lead the coefficient on education to statistical non-significance and/or it should raise the variance explained by the regression (which typically does not exceed 30% in Mincerian regressions like equation (6.1)). Surveying 25 studies reporting empirical estimates of the determinants of earnings in United States from the late 1950s to the early 1990s, they find that the introduction of a measure of cognitive performance into an equation using educational attainment to predict earnings reduces the coefficient of years of education by an average of 18 percent, without exhibiting any specific time trend.[27] By implication, they conclude that "This suggests that a substantial portion of the returns to schooling are generated by effects or correlates of schooling substantially unrelated to the cognitive capacities measured on the available tests" (Bowles, Gintis and Osborne 2001a, p.1149).

Finding a limited explanatory power for cognitive skills in accounting for earnings does not answer the underlying question of why firms are disposed to pay higher wages to individuals who spent a larger fraction of their time at schools. In addition to enhancing productivity on the job, firms may prefer educated individuals because they are self-selected according to their ability to identify with the authority, or even more simply because they believe that achieving educational degrees is a signal of non-cognitive ability. The first perspective does not necessarily require that acquired knowledge necessarily induces its owner to introduce better techniques of production, to discover new products and so on. The "human capital" concept is wide enough to include the life style of its owner.[28] If additional education is correlated with better health control and improved planning of fertility, education is valuable to a firm, since it is associated with a more continual job performance.[29] Human capital can therefore be alternatively described as the *quality of job performance* yielded by a person that can be modified through increased education.

---

[26] Altonji and Dunn 1996 analyse school quality, family background ad IQ tests as determinants of (possibly heterogeneous) returns to education, using representative samples of Americans. They find that parental background (and especially mother education) is by far the most important factor, with IQ remaining significant and school quality measures (student/teacher ratio, expenditure per student, teacher salaries) exhibiting inconsistent effects. Murnane, Willet and Levy 1995 divide up the general contribution of education to earnings using information on specific subjects taken during college, finding a substantive increase in the return to mathematics scores. Murnane et al. 2001 include among the determinants of wages academic ability, speed in problem solving and self-esteem, finding positive contribution by all three measures. Green and Riddel 2003 propose a model where earnings depend on cognitive and non-cognitive skills, which in turn are produced by education, experience and family background. By inference from sign and significance in a reduced form estimated on Canadian data (which use literacy score as measure for cognitive ability) they conclude that cognitive ability depends on years of education and parental background, but not on work experience.

[27] The contribution to explained variance is negligible, the mean value of ΔR^2 being 0.01, its median equal to 0.007 with a range of variation between –0.15 to 0.04.

[28] Classical exposition of the concept of "human capital" in Becker 1993, chpt.2, where the author discusses the implication of greater human capital on fertility choices.

[29] Weiss 1995 stresses the same point with different examples: better educated people exhibit less absenteeism, are less likely to smoke, to drink or to make use of drugs and are generally healthier. A firm minimising absenteeism cost will be available to pay a wage premium in order to attract a better-educated labour force.

This view has to be inserted in the capitalist framework of labour exchange. From this point of view the use of the term 'capital' is ambiguous, since it suggests the idea that any person owns some sort of capital, whether in monetary form or in immaterial form through education. Nevertheless, except for the case of self-employment, human capital does not represent a true productive input, since its owner if forced to sell her labour force in the market in order to get the return on her human capital.[30] But any seller of human capital in the capitalist labour market does not have any incentive to achieve her best performance, since the residual claimant of any increase of productivity is typically the firm. The non-observability of work effort pushes firms to adopt higher wages (known in the literature as *efficiency wages*) and/or random monitoring accompanied by the threat to fire in case of non compliance (known in the literature as *contingent renewal*).[31] Whenever acquiring education modifies individual preferences, such that the agent's (i.e. the worker's) objective function gets more in line with the principal's (i.e. the firm's), then the firm can reduce its surveillance costs and share with the worker these potential extra-revenues. In such a case we will observe a positive relationship between education and wages, due to the inefficiency reduction associated with imperfect observability of effort.

From an empirical point of view, this perspective is observationally equivalent to alternative ones, based on productive role of knowledge or ability signalling. And it is not easy to provide an empirical test of this claim. We have already mentioned the fact that education still retains significant explanatory power despite controlling for cognitive abilities in earnings regressions, and this could be taken as indirect evidence of the potential existence of behavioural traits that are imparted in schools and valuable for firms.[32] Bowles and Gintis (1976) claim that schools in capitalist economies are designed to transmit skills that are appropriate to the hierarchical role to be followed in the labour market: thus vocational schools reward the adherence to rules and the carrying out of assigned duties, whereas university and higher stages of education, destined to the offspring of the elite, enhance creativity and problem solving abilities.[33]

Let us now offer a formal representation of these ideas. We take as starting point the model of contingent renewal proposed by Bowles (1985), where effort is unobservable and the firm can either use wage incentive and/or surveillance to extract effort from workers.[34] In this framework we study the impact of changes in workers' preferences, which can be seen as behavioural traits imparted at school. Under a given set of parameters, the model predicts a positive correlation between education and wages. Let us define the production technology by

---

[30] See the critique reported in Bowles and Gintis 1975.

[31] There is a large literature on the adoption of efficiency wages under alternative assumptions: as a solution to the problem of adverse selection in hiring (Weiss 1990), to reduce hiring costs and/or the turnover among the employees (Layard, Nickell and Jackman 1991, chpt.3). For contingent renewal models two classical references are Shapiro and Stiglitz 1984 and Bowles 1985; a more recent one is Bowles, Gintis and Obsborne 2001b. The obvious implications in terms of inefficiency of contingent renewal and the effectiveness of redistribution of property rights are discussed in the introductory chapter by Bowles and Gintis in Bowles, Gintis and Gustaffson 1993.

[32] Bowles, Gintis and Obsborne (2001a and 2001b) define these traits *incentive-enhancing preferences*: ""Examples of such profitable individual traits are a low time discount rate, a predisposition to truth telling, identification with the objectives of the firm's owners and managers as opposed to the objectives of co-workers or customers, a high marginal utility of income, and a low disutility of effort." (Bowles, Gintis and Obsborne 2001a, p.1145).

[33] In a pioneering paper, Edwards 1977 interviewed groups of students and workers, using peer group rating to rank different personality traits. Starting from 32 predefined traits, he applied factor analysis obtaining three factors: *rule* (rule orientation), *depend* (predictability and dependability) and *internalise* (personal identification with enterprise/school goals). Interesting enough, these three factors exhibit high predictive ability with respect to wages (for workers) and to school scores (for students), thus not contradicting the view that schools teaches good attitudes that are valuable in the labour market.

[34] Analogous models are also reported in Bowles and Gintis 2000 and in Bowles, Gintis and Obsborne 2001b.

$$Y = f(L, K) = L^{\beta} K^{1-\beta}, \qquad \beta < 1 \tag{6.7}$$

where Y is output, K are all intermediate inputs, taken as given in the short run. While in standard neoclassical models L, the labour input, is a homogenous contractible commodity, in the present context we assume that the labour contract fixes the amount of hours L_p during which a worker becomes subordinate to firm managers' authority, but the effective productivity of these hours depends on worker's choice. If we consider that only two alternatives are available in each instant, either producing or shirking, then we indicate with λ the fraction of time devoted to production, and the labour input is defined accordingly as

$$L = L_p \cdot \lambda \tag{6.8}$$

While L_p can taken for given from the institutional context and cannot be modified in the short run, the labour input varies according to the worker behaviour. We assume that λ cannot be observed continuously by the firm (unless at very high costs), but monitoring techniques can be introduced at some cost, rendering shirking detection feasible in probabilistic terms.[35] When a worker is caught while not producing, she is fired, goes back in the pool of unemployed and looks for another job, which is found with a probability that depends on labour market conditions.

The worker chooses λ according to two considerations: the unpleasantness of effort exerted in production, and the cost associated to the risk of being fired in case of being caught when shirking. A rational worker equates the marginal utility of expected income to the marginal (disutility) cost of effort.[36] The expected income R is given by the contractual wage W times the probability of not being caught shirking (which has an upper limit of one for a worker who never shirks). In case of dismissal for not being productive, a worker enters the unemployment pool, where she faces two alternative events: either finding a new job (that in equilibrium offers the same contract and the very same wage) or remaining unemployed and living on the dole. We summarise all these events in the following definition

$$R = W \cdot (\lambda + (1-\lambda) \cdot (1-p)) + p \cdot (1-\lambda) \cdot [(1-u) \cdot W + u \cdot b] = W - (W-b) \cdot u \cdot p \cdot (1-\lambda) \tag{6.9}$$

where $p = p(m)$ is the probability of detection, which is function of the amount of surveillance m introduced by the firm, u is the unemployment rate (which for simplicity we assume equal to the unemployment probability)[37] and b is the unemployment subsidy. In order to get an analytic solution for the model, we introduce a specific functional form for the detection probability function p

$$p = \frac{m}{m+1}, \qquad p' > 0,\ p(0) = 0,\ p(\infty) = 1 \tag{6.10}$$

The worker's preferences are defined over income and effort. We parameterise individual preferences by the parameter α_i, $i = 1,...,n$, intended to capture the relative disutility of effort (or the relative

[35] Think for example of supervisors monitoring continuously n workers; each worker is inspected a fraction $1/n$ of her working time.

[36] This corresponds to her reaction (best reply) function for any given preannounced wage of the firm and any surveillance structure.

[37] A more general formulation replaces u with $\varphi(u), \varphi' > 0$. The second part of expression (6.9), $((W-b) \cdot u \cdot p \cdot (1-\lambda))$ is indicated by Bowles as the *cost of job loss* that a firm can inflict to a worker by raising the wage rate or the monitoring (or contributing to a higher unemployment rate as member of the capitalist class).

appreciation of income) for worker i. This parameter is taken as given by each worker when optimally selecting her level of effort, but can capture the impact of schooling experiences: we assume that more education lowers α_i, due to lower cost of effort and/or to higher utility associated to income.[38] Once again, we make use of an explicit Cobb-Douglas functional form for preferences

$$U(Y,\lambda) = (1-\alpha_i)\cdot \log(R) + \alpha_i \cdot \log(1-\lambda_i) \tag{6.11}$$

The worker chooses the optimal level of effort by maximising equation (6.11) after substituting the definition of R given in equation (6.9)

$$\lambda_i^* = 1 - \frac{\alpha_i W}{(W-b)\cdot u \cdot p} = 1 - \frac{\alpha_i W \cdot (1+m)}{(W-b)\cdot u \cdot m} = \lambda\left(\underset{+}{W}, \underset{-}{b}, \underset{+}{m}, \underset{+}{u}, \underset{-}{\alpha_i}\right) \tag{6.12}$$

The worker is more likely to produce the higher is the wage paid by the firm and/or the lower is the unemployment benefit, i.e. the higher is the cost of job loss. Similar impact is played by the unemployment rate: a higher unemployment rate, soliciting a greater effort as a consequence of the increased risk of permanence in unemployment, allows the firm to reduce either wage or monitoring. Bowles (1985) labels equation (6.12) as *labour extraction function*, because it displays alternative combinations of tools available to the firm: the 'carrot', accounted by the wage, and the 'stick', provided by surveillance.[39] We can therefore define isoquants associated to the effort production function (6.12) in the (W,m) space (let us call them *iso-effort schedules*), and the firm will select its optimal combination of stick and carrot, according to their relative costs. Which are the consequences of different values of the α parameter, for example associated to differences in the educational backgrounds ? A lower α implies a lower disutility of effort, allowing the firm to reduce monitoring m and/or wage W while eliciting the same level of effort.[40] The figure 1 shows alternative configurations of iso-effort schedules derived from equation (6.12).

The firm anticipates the worker best response and will select the optimal combination of wage W_i and surveillance m_i for worker i achieving the maximum profit in the short run

$$\max_{L_p, W_i, m_i} Y - (W_i + p_m \cdot m_i)\cdot L_p - p_k \cdot \overline{K} \tag{6.13}$$

under the constraint represented by equation (6.12) and the definition of Y given in equation (6.7). p_m is the unitary cost of per-capita surveillance. First order conditions for profit maximisation require that

[38] Bowles, Gintis and Obsborne 2001b show that a lower discount rate increases the value of the job rent that is shared between firm and worker when the worker is not shirking.

[39] Since the possibility of monitoring can be job specific, we have potentially here a theory of wage differentials. In addition, the model could also account for wage discrimination, whenever the firm is offering different combinations of "wage cum surveillance" to otherwise identical workers (think of white and blue collars), in order to prevent the formation of a common front of wage claims. See the discussion in Bowles 1985.

[40] Similarly: "We say a parameter *b* in the employee's utility function is incentive-enhancing if an increase in *b* shifts the employee's best-response function upward, an increase in incentive enhancing preferences leading an employee to work harder at every wage rate, holding all else constant." (Bowles, Gintis and Osborne 2001b, p.156). In our framework this corresponds to an inner shift of the iso-effort schedule.

$$\frac{\partial \Pi}{\partial L_p} = \lambda\beta\left(L_p\lambda\right)^{\beta-1}\overline{K}^{1-\beta} = W + p_m m$$
$$\frac{\partial \Pi}{\partial W} = L_p\beta\left(L_p\lambda\right)^{\beta-1}\overline{K}^{1-\beta}\frac{\partial \lambda}{\partial W} = L_p \qquad (6.14)$$
$$\frac{\partial \Pi}{\partial m} = L_p\beta\left(L_p\lambda\right)^{\beta-1}\overline{K}^{1-\beta}\frac{\partial \lambda}{\partial m} = p_m L_p$$

where for notational simplicity the subscript i has been ignored. The first condition is the standard equality between the marginal productivity of labour (measured in efficiency units) and its marginal cost (including monitoring). Dividing the second condition by the first one we get

$$\frac{\partial \lambda}{\partial W}\cdot\frac{W + p_m m}{\lambda} = 1 \qquad (6.15)$$

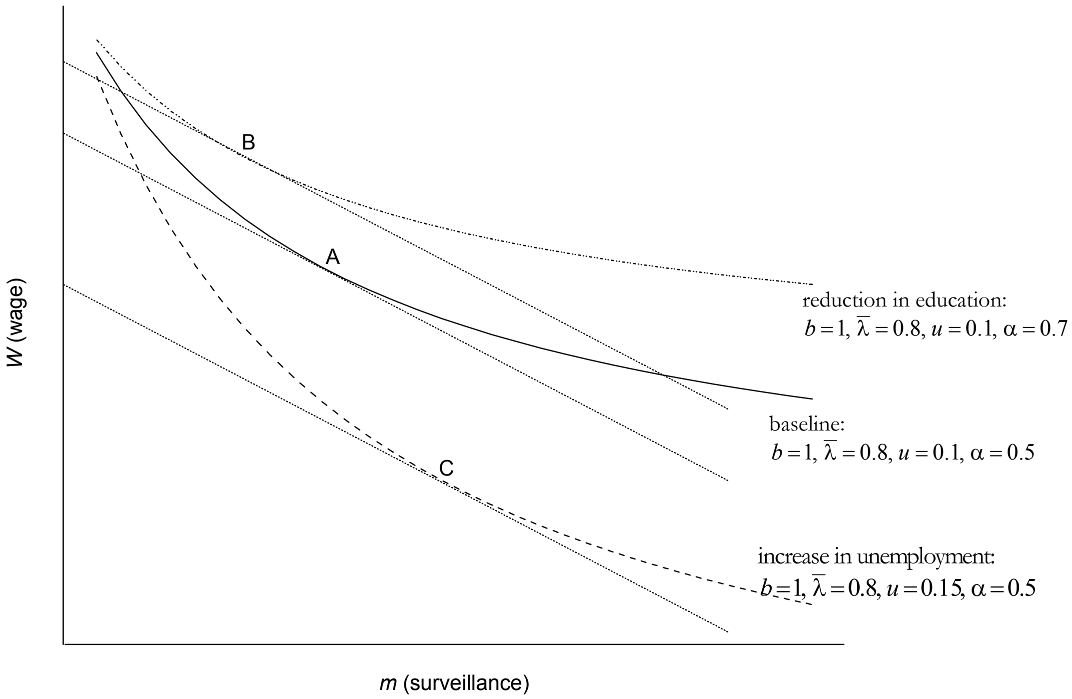
which is the standard Solow's condition stating that any efficiency wage must equate at the margin the rate of change of effort with the rate of change of wage. By contrast, if we divide the third condition in (6.14) by the second we get

$$\frac{\frac{\partial \lambda}{\partial m}}{\frac{\partial \lambda}{\partial W}} = p_m \qquad (6.16)$$

Condition (6.16) suggests that the firm will select the lowest iso-cost schedule that is tangent to the iso-effort curve, since the left hand side of (6.16) is the (inverse of the) slope of the schedule implied by equation (6.12) and the right hand side is (the inverse of the) slope of a per-capita cost function like $C = W + p_m m$. It tells us that the cost of monitoring has to equate the marginal rate of substitution in the effort extraction function between the cost of job loss and the probability of being detected. This corresponds to point **A** in figure 1. If we now wonder about the consequences of better educated labour force in this context, we have to recall that less education implies higher cost of effort, thus rising the iso-effort locus: in order to extract the same amount of effort for a less educated worker it is necessary to offer a higher wage and/or to introduce greater surveillance (point **B** in figure 1). Thus in general equilibrium, better educated workers will be in greater demand, since they are characterised by lower cost of incentives, while less educated workers will suffer greater unemployment. Higher unemployment creates an opposite situation to an increase in α, since it lowers the need of incentives to elicit the same amount of effort (see point **C** in figure 1). The overall effect is ambiguous, depending on labour market equilibrium, which has to be stratified for educational attainment, since the supply is inelastic in the short run. If unemployment effects are stronger than relative demand for better educated workers, we will obtain a positive correlation between education and wages, which does not depend on productivity of human capital but relies on behavioural traits supposedly induced by education (what can be called as *affective capital*).[41]

[41] Even in partial equilibrium analysis we can obtain a positive correlation between wage and education, i.e. a negative derivative of the optimal wage W with respect to α. By computing the first order derivatives from the λ function in equation (6.12) and plugging them into equations (6.15) and (6.16), we find that $sign\left(\frac{\partial W}{\partial \alpha}\right) = sign\left(W(2m-1) - b\right)$ which is always negative for $m < \frac{1}{2}$, which occurs whenever monitoring becomes extremely expensive.

Figure 1 – Alternative isoquants from the optimal effort function



4. Education as a signal or as a screening device

A complementary explanation for the returns to education considers the role of *information revealing* associated to schooling experience. This is consistent with the empirical observation that a large share of class activity in schools is devoted to testing students and marking their performance, and with the fact that previous academic career reveals to be a strong predictor of current one. The easiest way to frame this argument is to think of personal abilities (that might include affective capital, as in the previous section, or simply talent, already discussed in the second chapter); the ability endowment is private information of students. Ability is valuable for firms, since productivity rises with ability. We know from the literature (see Riley 2001 for a survey) that imperfect information can lead to adverse selection phenomena up to complete disappearance of the market. However agents may device alternative strategies to overcome these imperfections. One of these strategies is the emission of signals that is (imperfectly) correlated with the hidden information. Another is to adopt wage policies that induce individuals to undertake costly screening devices, eventually revealing the hidden information. The common trait of this approach is that education and wages are positively correlated, but it is a classical case of spurious correlation, since both correlate with unobservable ability. The other common trait is the absence of any direct impact of education onto productivity, which portrays expenditure on education as wasted resources.

Let us review these approaches more formally. We assume that individual talent A_i constitutes a productive factor for firms (i.e. it enters positively the output production function), but it is not directly observable. If it could be observed, then in equilibrium each profit maximising firm would pay each unit of talent its marginal product; as a consequence, better-endowed individuals would earn higher wages. Otherwise, each firm (supposed to be risk neutral for simplicity) will offer the same wage to all workers, based of the productivity of a worker endowed with the expected talent in the population.

The absence of perfect information over individual ability introduces potential conflicts of interest among the agents, thus rendering self-declarations non trustable. On one side there is an inherent conflict between firm and worker, since the latter has the incentive to self-declare the highest endowment of talent, given the impossibility of checking it; for symmetric reason, the former (firm) has the incentive to deny the presence of talent, since its declaration cannot be disproved. On the other side there is a conflict between workers with different endowments of talent. Let us suppose the existence of only two types of workers, the "talented", with talent endowment equal to A_2, and the "not talented", with talent endowment equal to A_1, $A_1 < A_2$.[42] The first group represents the fraction n of the labour force, while the second reaches the complementary fraction $(1-n)$. From past experiences, we suppose that firms know the talent distribution in the population, that is the parameters A_1, A_2 and n. In the absence of any further information, the best strategy for the firm is presuming that each worker is randomly extracted from the pool of job seekers and that her talent endowment corresponds to the expected value in the population (equal to sample mean $\overline{A} = nA_2 + (1-n)A_1$). As a consequence, the firm will offer an identical wage to all job seeker; if φ represents the marginal revenue of talent, firms competition will push the unique wage rate to the point indicated by the following condition

$$\overline{W} = \varphi\overline{A} = \varphi\left[nA_2 + (1-n)A_1\right] \tag{6.17}$$

According to equation (6.17), talented workers receive less than their actual productivity; if talent were freely observable, they would get $W_2 = \varphi A_2 > \overline{W}$. Conversely, less talented workers take advantage of the absence of information, since under perfect information they would get $W_1 = \varphi A_1 < \overline{W}$. Thus the absence of (or the costly) observability of talents leads to a compression of the wage distribution (in the limiting case yielding the disappearance of wage differentials), with implicit subsidisation from talented workers towards non-talented ones. If the former group could obtain the recognition of their true endowment from the firm, they would not hesitate in pursuing it, thus raising their own wage and implicitly lowering the wage of the latter group.

In such a context two alternative interpretations of educational attainments have been proposed, as responses to asymmetrical information. The first interpretation considers achieving education as a signal that workers emit towards firms, in order to reveal their true endowment of talent.[43] Suppose that a firm offers a wage schedule where the offered wage rises more than proportionally with the amount of obtained education

$$W_i = \beta(S_i)\cdot S_i,\ \ \beta' > 0,\ \ i = 1,2 \tag{6.18}$$

[42] At this point of the discussion it is totally irrelevant whether "talent" indicates biological abilities or the effects of family backgrounds. What is relevant is that agents cannot modify their talent endowments, and that firms cannot freely observe it. In this respect, see the discussion in Goldberger and Manski 1995.

[43] See the seminal paper by Spence 1973, where the equilibrium is defined by self-fulfilling conjectures of the firm: "Thus, in these terms an equilibrium can be thought of as a set of employer beliefs that generate offered wage schedules, applicant signalling decisions, hiring, and ultimately new market data over time that are consistent with the initial beliefs". (p.360). However his claim of the existence of a multiplicity of equilibria, some of which of the "pooling" type (namely, an identical wage paid to all workers), have been criticised as a general equilibrium solution, since it is not robust to deviations by competing firms. Thus there would be a unique separating equilibrium, as described by Riley 2001, p.438-42.

where β can be thought as the marginal return to education, with the counterintuitive property that $\beta(S_2) > \beta(S_1)$ for any $S_2 > S_1$.[44] If firm's conjecture are confirmed by individuals choosing different amounts of education for different talent endowments (thus conforming to their future employer expectations), then acquiring education constitutes an effective signal able to overcome the inefficiencies introduced created by imperfect information.

However, the necessary (but not sufficient) condition for a potential signal to become an actual one is that signalling costs are negatively correlated with an individual's unknown talent.[45] In other words, we are requiring that the unitary cost of education for talented individuals be lower than the corresponding cost for less talented ones. In addition, we require that the offered wage schedule will encourage talented individuals in acquiring education (yielding a net gain for them), in the meanwhile discouraging less talented people from imitating talented ones in acquiring education (thus associating a net loss to this choice). More formally, a separating equilibrium requires that the amount of signalling emitted by each type of individual being a dominant strategy against all the other strategies available. Let us assume that individual preferences are simply described by

$$V_i = W_i(S_i) - \gamma(S_i, A_i), \quad \gamma_S > 0, \quad \gamma_A < 0, \quad i = 1,2 \tag{6.19}$$

where $\gamma(S_i, A_i)$ represents direct and indirect cost of achieving the S_i amount of schooling for an individual with talent A_i. Cost of attendance is obviously increasing in schooling and decreasing in talent. This assumption implies that for any amount of education talented individuals face a lower cost than less talented ones.[46]

Given the wage offer described by equation (6.18), less talented people do not find profitable imitating the behaviour of talented people whenever $W_2 - \gamma(S_2, A_1) \leq W_1 - \gamma(S_1, A_1)$, where the first term indicates the utility of emitting the highest signal S_2 for a less talented individual (denoted by her talent endowment A_1), while the second term corresponds to the utility associated to emitting the signal S_1. In order to obtain a separating equilibrium, we also need a participation constraint for talented individuals, who have to find convenient to emit the signal under the wage offer (6.18). This corresponds to the condition $W_2 - \gamma(S_2, A_2) > W_1 - \gamma(S_1, A_2)$. By combining the two previous inequalities, we get

$$\gamma(S_2, A_1) - \gamma(S_1, A_1) \geq W_2 - W_1 > \gamma(S_2, A_2) - \gamma(S_1, A_2) \tag{6.20}$$

Condition (6.20) shows that the wage differential proposed by the firm must lie in an intermediate position between the cost differentials faced by a less talented individual when considering the possibility of emitting the signal S_2, and the corresponding cost differential faced by a talented

---

44 This wage offer can be rationalised by the conjecture that talented individuals will obtain more education, and at the same time they are more productive because better endowed with talent. See also Weiss 1995 for a discussion of these implications. Belzil and Hansen 2002 find empirical evidence of convex relationship between education and earnings that they interpret as evidence of "ability bias".

45 An implicit assumption of the literature about signalling is the absence of any other obstacles in acquiring education. On the contrary, if financial markets are imperfect and individuals from poor families are liquidity constrained, then the signal becomes noisy, its informative content vanishes, and the recipient is unable to recognise the true message: the absence of education could indicate either low talent endowment or poor family origins. See the discussions of equilibria under these circumstances in Giannini 2001. Bedard 2001 includes the possibility of financially constrained agents, but she assumes that firms can observe those who are prevented from acquiring education.

46 With reference to figure 2, this assumption corresponds to the fact that indifference curves for talented individuals are flatter that corresponding curves for less talented people (often referred in the literature as "single-crossing property").

individual. When the left hand side of inequality (6.20) is violated, the less talented individual will find it convenient to imitate the talented individual in achieving the amount S_2 of education; on the contrary, when the right hand side is violated, the talented individual will not find it convenient to emit the signal, and they will become undistinguishable from the less talented ones. This situation can be visualised with the help of figure 2, where we have drawn two indifference curves corresponding to the two types of individuals, with flatter indifference curves being associated to more talented people. The two straight lines exiting from the origin describe the wage offer by the firm; the first line applies to individuals choosing the amount S_1 of education, while the second schedule is referred to those choosing S_2 of education. Conditional to the firm's wage offer, less talented individuals will maximise their utility by selecting the amount S_1 of education (point **A**), while talented individuals will achieve their maximum choosing S_2 of education (point **B**).[47] This corresponds to a separating equilibrium, where no agent has incentive to deviate: if talented individual were to select education S_1^*, they will end up on a lower indifference curve; similarly, if less talented individuals were to choose S_2^* they will not improve their welfare. Finally, a firm offering a wage exceeding $W_i = \beta(S_i) \cdot S_i$ will make losses, despite being able to attract only talented individuals.[48] We have therefore shown that an appropriate wage announcement of the firms is able to induce individuals to self-sort according to their unobservable characteristics, thus revealing their hidden information. The workers' choices confirm firms' conjectures, thus satisfying the equilibrium requirements.

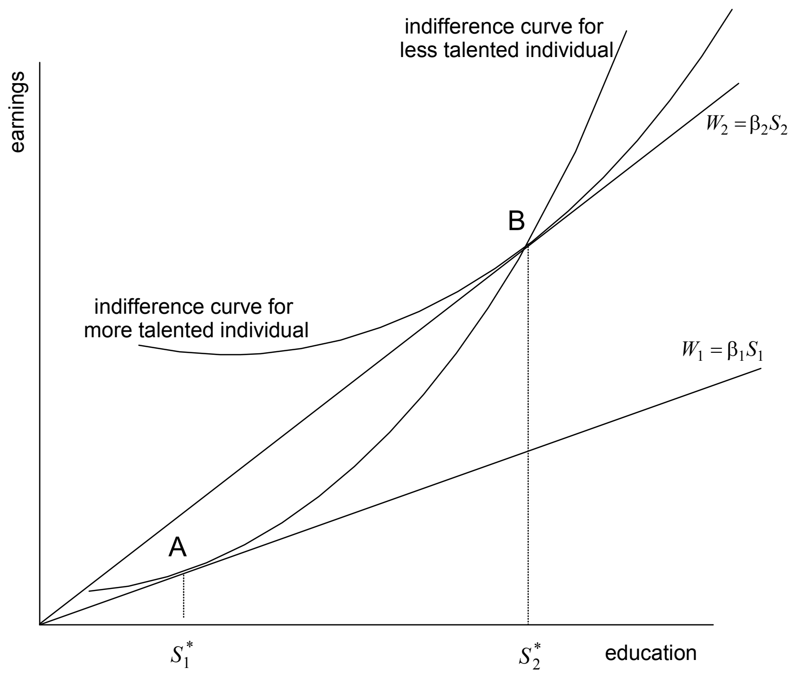
It is important to remember that in this framework acquiring education does not increase a worker's productivity per se, nevertheless education is positively correlated with education thanks to its information revealing property. Signalling models are characterised by the informed side of the market (the workers in the present case) acting as first mover, with the uninformed side, the firm, forming (and/or updating) their expectations on the relationship between education and productivity, and consequently offering wage contracts analogous to those reported in equation (6.18).[49] Despite the revelation of hidden information through educational choice, this equilibrium is a constrained Pareto-efficient equilibrium, since resources are wasted to emit the signal.

[47] In principle, less talented individuals are indifferent between points **A** and **B**. However, a firm could always leave out less talented individuals by shifting to the right of point **B**, along the $W = \beta_2 S$ line, since the utility reduction for less talented individuals are first-order magnitude, whereas they are only second order for talented individuals. This also explains why it is possible to get a continuum of separating equilibria (see Riley 2001 for discussion).

[48] While the original model proposed by Spence 1973 allowed for the existence of pooling equilibria, Riley 2001 claims that this result is only possible as a partial equilibrium case, because under firm competition it always pay to deviate in order to attract the best workers, unless firms are already offering a wage associated with a zero profit condition.

[49] Even if formally a simultaneous game, it could be modelled as a sequential game (as in figure 1 of Spence 1973, p.359).

Figure 2 – Signalling equilibrium



A different situation emerges when we abandon the assumption of talent-based differences in the cost of educational achievement, and we wonder what types of equilibria emerge in this framework. Let us assume the existence of a screening device, able to reveal the unobservable information. Accessing to the screening has a fixed cost equal to γ for any type of agent; afterwards information concerning the examined agent becomes freely available to anyone. Some aspects of the schooling experience, and particularly some crucial turning points (like completing compulsory education, graduating from high school, taking tests for college admission) may conform to this view.[50] As in the previous case we assume the existence of only two types of agents, talented (type A_2) and non-talented (type A_1). Their existence and their distribution in the population is common knowledge.

Given the availability of an information revealing tool, firms announce a different wage policy: they will pay a wage $W_2 = \varphi A_2$ to any worker accepting the screening and shown to be talented, and a wage $W_1 = \varphi A_1$, $W_1 < W_2$ otherwise. If no worker agrees to be screened, the firm operates as under perfect ignorance about talent features of job applicants, and therefore will pay an identical wage to everyone, based on the mean ability in the population (as described by equation (6.17)). If the firms are paying an identical wage to all workers, more talented people have an incentive to stand out from the others, in order to appropriate the implicit rents associated to their endowment. However they will not find it convenient to undertake the screening if

$$\gamma > W_2 - W_1 = \varphi(A_2 - A_1) \qquad (6.21)$$

[50] Stiglitz 1975 advances an interpretation of the entire educational system as a screening mechanism working on the benefit of potential employers. That paper inspires most of the following discussion.

Whenever the cost of screening is high enough, there is no incentive to use it because the net income for a talented worker $(W_2 - \gamma)$ exceeds the income she would be granted even under the worst situation (like being misclassified as less talented). When inequality (6.21) is satisfied, we observe a single wage paid to all workers, despite their different (but unobservable) quality. This situation is indicated as *pooling equilibrium*.

We now consider an opposite situation, where the screening cost is low enough to satisfy

$$\gamma < W_2 - \overline{W} = \varphi A_2 - \varphi[nA_2 + (1-n)A_1] = \varphi(1-n)(A_2 - A_1) \tag{6.22}$$

In such a case the talented people find it convenient to afford the cost of screening in order to signal their ability endowment to the firm. Whenever some workers undertake the screening, the firm reduces the wage to W_1 to the remaining workers, under the presumption that they are less talented. This will convince some reluctant and talented worker to undertake the screening. For a sufficiently low cost of screening, the information is revealed and a separating equilibrium emerges.

What does happen in the intermediate case, when the screening cost lies in the interval

$$W_2 - \overline{W} < \gamma < W_2 - W_1 \tag{6.23}$$

or alternatively

$$\overline{W} > W_2 - \gamma > W_1 \tag{6.24}$$

In this case the talented workers do not have an incentive to submit themselves to the screening, for their net income is reduced. Nevertheless, the firms can force them to take it under the threat of considering all workers less talented, therefore reducing the general wage from $\overline{W}$ to W_1. Once again we obtain a separating equilibrium, which however is Pareto inferior to the pre-existing pooling equilibrium. In fact all workers experience a wage reduction, because talented workers obtain $W_2 - \gamma < \overline{W}$ and less talented workers receive $W_1 < \overline{W}$.[51] This is due to the specific assumption we made at the beginning of the section, where we considered the resources spent to acquire education as wasted resources.

We can sum up different cases by stating that under asymmetric information the possibility of screening (to be interpreted as either a cost of schooling or a cost of on-the-job training and selection) is associated to multiple equilibria, parameterised over the cost of screening γ. However, unlike previous results, the demand for education does not vary continuously with its price. This is due to the fact that the possibility of screening improves the quality of matching (unobservable) talents to job opportunities at the cost of greater earnings inequality. In fact there is an underlying conflict about the value of information. Talented individuals (type-2 agents) have the interest to see the recognition of their endowment, in order to raise their market value. For different reasons, a single firm also has the same interest, conditional on being able to conceal the information: if better workers can be identified without public recognition of the sighting, they could be allocated to more appropriate jobs, securing the extra profit associated to the productivity difference $(W_2 - \overline{W})$. However, as soon as the

[51] However the private return to screening for talented workers is still positive because the right hand side of inequality (6.23) can be rearranged as $\frac{W_2 - W_1}{\gamma} > 1$. By contrast, the social rate of return (which is a weighed average of private rate of returns, including firms' profits) is negative.

information is disclosed to the public, competition among firms drives the extra profit to zero, and talented workers obtain a fair reward for their talent endowment.

Things are slightly different under the case of imperfect symmetric information, namely when both workers and firms cannot freely observe the talent endowment. If workers are risk adverse, they will never accept being screened, since they do not want to incur the risk of wage reduction once identified as "less endowed". In fact, indicating with $U(W), U' > 0, U'' < 0$ the risk-adverse worker preferences, we know that

$$U(\overline{W}) = U(nW_2 + (1-n)W_1) > nU(W_2 - \gamma) + (1-n)U(W_1 - \gamma) \qquad (6.25)$$

Inequality (6.25) tells us that individuals always refuse to be screened even in the case of zero cost of screening γ. They may be induced to afford the screening only if they are they are subsidised to do it (as in the case of $\gamma < 0$). If firms have to meet the expense of screening, the assumptions we introduce over the shape of their objective function become crucial.[52]

There are several objections that can be raised against this approach. The most immediate one is that we do not observe an empirical equivalent of what we termed as "screening". If the term has to be taken as equivalent of "educational career", then it typically provides an imprecise assessment of individual abilities. No employer will base its offered wage on the grades obtain in a specific subject! In addition, ability is a multi-dimensional concept, and different school subjects typically call into play different types of abilities (creativity, logic, expressivity, and so on). If screening has to be taken as "admission test" for job applicants, it seems implausible that firms should base their entire wage policy on these results, when they may hire a worker on a temporary base and test her directly on the job. The same type of objection applies to the case of screening as equivalent to "possessing a school diploma": why should a firm rely on an educational certificate released by an often unknown college, once it has the opportunity to directly test the worker ?.

* * *

With respect to the empirical validity of the signalling hypothesis, and particularly on the relative importance of the signalling versus human capital explanations, we find in the literature several attempts to discriminate between the two. Starting from the idea that in a signalling perspective education is worthless whenever no screening is required, earlier attempts focused on sectoral differences in the return to education (Layard and Psacharopulos 1974). Similarly, significant differences in returns to education between employees and self-employed is compatible with the idea that self-employment does not require any screening at the entrance to the job.[53] However, as clearly recognised by Riley (2001), in cross-sectional analysis the human capital and the signalling/screening approaches are observationally identical, since both are based on the assumption that individuals optimally select the amount of education that will maximise their expected utility, and that firms reward education because it is associated with greater productivity. In addition, empirical tests based on sample splitting are not robust against sample selection bias.

---

[52] When firms are risk neutral and there are decreasing returns to ability in production, they have an incentive to pay for screening, since $\overline{W} > nW_2 + (1-n)W_1$ where $W_i = f(A_i), f' > 0, f'' < 0$. However this depends on the relative cost of screening, which has not to exceed the expected gain, i.e. $\gamma < \overline{W} - [nW_2 + (1-n)W_1]$. Under the linearity assumption made in the text $W_i = \varphi A_i, i = 1,2$, this obviously does never occur.

[53] See for example Brown and Sessions 1999, where they exploit the employee/self-employment divide on Italian data, finding evidence supportive of the "weak screening hypothesis", that is "...whilst the primary role of schooling is to signal, it may also augment inherent productivity."(p.397). However their results do not take into account that data on earnings and worked hours for self-employed are much less reliable than corresponding data for dependent employment.

A different approach has been followed among others by Riley (1979) and Groot and Oosterbeek (1994). The first author considered the fact that within a sorting framework extra information about worker productivity (through admission tests or job tenure) reduces the importance of education as a signal. His results are consistent with a sorting model based on unobservable ability, but also with self-sorting based on different degrees of risk aversion. In addition he finds little differences between acquiring and not acquiring qualifications (absence of ship skin effects), providing additional support to the human capital hypothesis. The second paper focussed on the different informative content possessed by years of education when distinguishing repetition, skipped or dropout years. They exploit the idea that repeating one year should have a non-negative effect on wages in the human capital approach, and a negative effect according to the screening model. The opposite situation applies to skipping years. Using Dutch data, they find absence of impact from repeated years and negative impact of skipping years, both results pointing towards the human capital approach.

A more ingenious approach has followed a route similar to the natural experiment approach. Lang and Kropp (1986) started from the assumption of ability differentials giving rise to a separating equilibrium based on a continuum of ability classes, and made the following observation: "In any sorting model the wage associated with a given level of education depends on the innate ability of the individuals who obtain that level of education….Suppose then that for some reason a group of (lower ability) workers who previously would have left school after education $s-1$ remain in school through s. Since the average ability among workers with schooling s declines, the wage associated with s declines. Consequently the benefit of obtaining $s+1$ increases".[54] Building on this consideration, they proposed a discriminatory test between the human capital and the screening approaches based on relative enrolment rates: according to the former approach, educational reforms affecting the school leaving age should affect only students who were previously prevented from attendance, leaving all the others unaffected; vice versa, according to the latter approach the entire distribution of enrolment is modified, because the informational content of different achievements is modified.[55] On similar lines, Bedard (2001) presents a model of sorting where a fraction of agents is constrained from achieving education by reasons that are uncorrelated with ability (think of family income). In a three-fold partition (dropping out, graduating from college and accessing the university), college proximity corresponds to an exogenous variation that reduces the mean ability of the intermediate group, since it eases the access to college. Because of the reduction of the signalling value of the credentials of the intermediate group (graduate from high schools), an increase in dropping-out rates is expected (i.e. members of the first group), which could not be accounted for by the standard human capital theory. She finds significant evidence of sensitivity of ordered probit thresholds to college proximity in NLS data for individuals who completed their schooling experience in the late 1960's and early 1970's.

Overall, we may conclude this section by stating that education may have a return in the labour market as signal for unobservable components. However the empirical evidence in support of this view is not entirely convincing, and we should limit ourselves to saying that educational credentials convey information to potential employers who do not have the time/possibility of assessing all self-declared competences. But the information concerns much more the amount of knowledge incorporated by the job applicants than her unobservable skills.

[54] Lang and Kropp 1986, p.613.

[55] Lang and Kropp 1986 find that compulsory attendance laws (CAL) increased the educational attainment of individuals who are not directly affected by the reforms using American aggregate Census data. On the contrary, Chevalier et al. 2003 find no impact of increasing compulsory leaving age in Britain and Wales out of lowest educational attainers, thus arguing against the screening hypothesis.

5. *On the job training*

The return to education has to be assessed not only with respect to differences in average earnings, but also in terms of life-long earnings. Figure 3 in chapter 1 showed a typical age-earning profile, which is characterised by a hump-shaped contour, with different profiles according to obtained educational attainment.[56] The rising portion comes to a standstill during the first half of the forties, and then begins to decline afterward, but the turning point is postponed for higher educational attainment (and in the case of tertiary education it is estimated to occur after the retirement age). As a consequence, the return to education are not restricted to the entrance in the labour market, but appears along the entire working life of individuals.

From a theoretical viewpoint, it is challenging to provide an explanation of this dynamics. In the appendix to chapter 2 we have already advanced an explanation based on the human capital investment theory. If we accept the idea that human capital can be acquired at any point of life (but the incentives decline with the shortening of the expected duration of the life), and that it decays at a constant rate along the entire life span, we obtain the hump-shape: during the initial portion of the life individuals devote all the available time to human capital investment (conditional on talent endowment and family background); when entering the labour market, the fraction of time devoted to human capital formation shrinks but not disappears, thus contributing to explain the rising part; but then the obsolescence of acquired knowledge sooner or later becomes dominant, and the down sloping part of the schedule appears.

Alternative explanations have been put forward for the upward trend in the earnings profile. The initial one invokes the existence of *learning by doing* (Arrow 1962). If individual productivity is positively correlated with tenure because each worker improves its performance through the practice in a job, then competitive markets predict a positive correlation between age and earnings. A second explanation concerns the screening approach to the return to education: if the unobserved talent is gradually revealed by the observed on-the-job performance of the worker, firms may be willing to pay higher wages for more able workers, in order to retain them.[57] A third explanation combines the problem of revealing the hidden characteristics of the worker (*talent allocation*) with the need of providing the right incentive to put effort in a context of imperfectly observable action (*objective alignment*). Promotions are typically a low-frequency high-impact wage change that may be used by firm management to motivate and to retain their personnel.[58] However all these explanations have problems with explaining the declining segment, unless demotion and lay-off are introduced as possibilities.

[56] The impact of educational attainment on the slope of the age-earning profile in Europe (using the European Community Household Panel) is analysed in Brunello and Comi 2003.

[57] In a pooling equilibrium, a related prediction would be the decline of wages for workers below the average talent endowment. However, this could not be observed in institutionalised contexts, where minimum wage legislation and workers' union may prevent this drop.

[58] In the language of the model exposed in section 3, a rising wage represents a greater cost of job loss, thus becoming a powerful deterrent to shrinking and/or quitting. See also the discussion of the promotion policies in Lazear 1995, chapter 5.

A fourth explanation, that we will discuss at some length in this section, extends the paradigm of human capital to the entire working life. If we consider that human capital can also be accumulated through job training (which is a mixture of formal schooling, work experiencing and interaction with senior colleagues), and we allow for positive correlation between human capital, productivity and wage, we obtain the upward portion of wage profile. When we then consider the decreasing marginal productivity of job training, sooner or later the accumulation of human capital ceases and the earning profile becomes flat. Eventually, obsolescence is responsible of the decline in earnings in the final part of the working life.

Differently from schooling decisions, the job training decision involves two agents, the worker and the firm. Not only the cost of the training can be borne by one or the other, or shared among them, but also the very same content of the training can be jointly determined. In sectors characterised by rapid technological change, competitiveness requires a rapid adaptation to new technologies, and the employed labour force is requested of quick learning of how to operate with improved equipment. Both firm and worker have a common interest in this state of affairs, the former to enhance its competitiveness, the latter to lengthen the expected duration of employment through being indispensable to the production process. This attitude requires on both sides the expectation of a long-lasting labour relationship: if either one or the other side expects the labour contract to terminate soon (because of low wages, bad jobs, absence of firing costs), the human capital investment associated to job training is difficult to materialise.

Table 1 reports some measures of the diffusion of training in European firms. Apart from wide variation across countries, which is correlated with the country technological level (and therefore with the level of GDP per capita[59]), it can be noted that the countries offering more opportunities of training are also the countries where the percentage of participating workers is higher: (unweighed) sample correlation between "offer" and "participation" is 0.64. At the same time, most of country differences are related to the structure of the vocational system: countries where vocational schooling is less pervasive[60] are countries where the participation and the intensity of job training are also low.

This aggregate evidence could however mix different sources of variation, including gender composition of the labour force, sectoral composition of employment, educational attainment, and so on. The following table 2 (taken from Bassanini and Brunello 2003) reports information extracted from the 1996 wave of the European Community Household Panel concerning male employees, aged between 30 and 60, working full-time in the non-agricultural private sector. From this table one may easily recognise that training opportunities are not evenly distributed among workers: former schooling seems complementary to subsequent training, which is also more frequent among high rank occupations.

---

59 The unweighed sample correlation of GDP per capita (measured at PPP) with offer and participation are respectively 0.55 and 0.48.

60 According to OECD 2001, table C2.1, the share of secondary enrolment in prevocational and vocational oriented schools is lowest in Ireland (20.6), Portugal (25.0), Greece (25.8) and Spain (31.2), and is highest in Austria (77.9), United Kingdom (66.7) and Netherlands (66.1).

Table 1 – Selected indicators of training – Europe (1999)

	1 offer	2 participation	3 intensity	4 cost	5 gdp capita
Austria	72	35	29	1.3	111.3
Belgium	70	54	31	1.6	106.5
Denmark	96	55	41	3.0	118.8
Germany	75	36	27	1.5	106.4
Spain	36	44	42	1.5	82.2
Greece	18	34	39	0.9	66.1
France	76	51	36	2.4	99.7
Finland	82	54	36	2.4	100.7
Italy	24	47	32	1.7	103.4
Ireland	79	52	40	2.4	112.3
Luxembourg	71	48	39	1.9	188.8
Netherlands	88	44	37	2.8	114.6
Portugal	22	45	38	1.2	72.3
Sweden	91	63	31	2.8	105.2
United Kingdom	87	51	26	3.6	100.6
Norway	86	53	33	2.3	129.0
Bulgaria	28	28	35	1.0	28.3
Czech Republic	69	49	25	1.9	59.1
Estonia	63	28	31	1.8	38.7
Hungary	37	26	38	1.2	49.0
Latvia	53	25	34	1.1	28.6
Lithuania	43	20	41	0.8	34.5
Poland	39	33	28	0.8	39.0
Romania	11	20	42	0.5	23.8
Slovenia	48	46	24	1.3	68.5

Notes:
1 – Offer: Enterprises offering continuing training as a proportion of all enterprises (%)
2 – Participation: participation rate in enterprises offering training courses (%)
3 – Intensity: participation hours per participant
4 – Cost: costs of training courses as a proportion of the total labour costs of all enterprises
5 – Gdp: gross domestic product per capita in Purchasing Power Standard (Europe15=100):
Source: Eurostat 2000, Continuing vocational training survey (CVTS2)

Table 2 – Incidence of training – selected European countries (1996)

	Percentage of individuals receiving on the job training
Country	
Austria	21.1
Belgium	19.9
France	15.5
Germany	19.9
Italy	6.7
Spain	11.8
United Kingdom	39.1
Sectors	
Mining, manuf. and utilities	15.0
Services	21.4
Individual characteristics	
Less than upper sec. education	7.7
Upper secondary education	18.4
More than upper sec. education	33.7
High-skilled occupations	31.0
Medium-skilled occupations	12.4
Low-skilled occupations	6.3
Total	***17.3***

From a theoretical point of view, whenever we observe occurrence of training a crucial question can be asked about who is paying the cost. The cost of training incidence reported in table 1 includes direct costs (teachers, enrolment, materials) and personnel absence costs, with relative shares ranging from 78% in direct costs for United Kingdom to 62% of absence costs for Spain. Direct costs can be paid either by the firm or by the worker, whereas when the training is imparted during workers hours, the opportunity costs is entirely paid by the firm. A useful distinction introduced by Gary Becker is between *perfectly general training* (general skills that are useful with other employers and that increase labour productivity by the same amount with all employers) and *perfectly specific training* (skills that increase the productivity of the worker only in her current job).[61] Examples of the first type are the abilities of driving a car or using a computer, whereas examples of the second group are the ability to operate a specific good-lift or to use software developed for the company needs. An increase in generic human capital raises the market value of a worker, whereas an increase in firm specific human capital leaves her value unaffected for all firms except the employing one.

It is easy to prove that in a competitive labour market, where the ongoing wage reflects the average labour productivity, no firm will accept to pay the cost of training for general human capital, since it will be forced by competitive forces to pay a wage corresponding to the augmented productivity of the worker, without being able to recover training costs. Should it try to reduce the wage below the labour productivity, the trained worker would be hired by competitors, and the training investment would be sunk. Seen from the point of view of competitors, the incentives operate in a complementary way: the best strategy is waiting for a worker to be trained by another firm, and then 'steal' her (a common practice, often indicated as *cherry picking*). The aggregate outcome is that no firm will be available to invest in training for general human capital, and the workers will be forced to afford the cost of training by themselves. This does not necessarily require a direct monetary payment, since there are alternative ways to charge workers with the cost of their training: for example think of apprenticeship contracts providing for a reduced wage during the training period. Whenever financial markets are imperfect and workers are liquidity constrained (or risk adverse), the likely outcome is the underinvestment in training for general human capital. On the contrary, if we consider the investment in firm specific human capital, a worker will be reluctant to pay the cost, since once out of the firm this investment will be wasted. In such a case, this investment will materialise if and only of the firm will afford it.

Despite these sharp theoretical predictions, in the real world we observe that firms quite often afford the costs of general training of their workers. A typical example is given by the training taking place during working hours. To account for these phenomena, there are two alternatives: either we introduce imperfect information on workers ability (and therefore the training acts as a screening device, where the expected gain is given by improved information), or we abandon the paradigm of perfect competition in the labour market (if a firm has monopsonistic power, it will set the wage below the productivity, thus recovering training costs).[62]

While useful from a theoretical point of view, the distinction between general and specific human capital is also empirically vague. Take the case of skills that are sector specific, which are neither firm specific nor completely general. In addition, the generality of skills depend on the work organisation within a firm: multitasking assignment favours the development of general skills, whereas a rigid organisation tends to transform any task in firm specific. The generality of training also depends on the

[61] See Becker 1993, pg.36-40. See also the discussion in Acemoglu and Pischke 1999.

[62] Suppose for example that firms have an informational advantage over the true workers' productivity obtained from training investments. If they do not reveal it (and they have no incentive in doing so), they will pay a wage below productivity, and therefore will be willing to afford the cost of training, in order to cash this gain: see Agemoglu and Pischke 1999. In the same vein, Bassanini and Brunello 2003 show that compression in the wage distribution favours the occurrence of workers' training.

external vocational system: the provision of official certification (as in the well-known German system) favours the adaptability of skills outside the firm where the training has taken place.[63] If a syndicate of firms offers the training, this reduces the competition between firms and increases the generality of skills; in addition, it also renders more likely the training in small-size firms, which may be quite unwilling to invest in general human capital.

Table 3 (adapted from Lynch 1994) shows the cross-country variety of institutional solutions to the training problem. National systems differ in terms of where training takes place (whether within firms or outside them, in external schools), who is financing it (public authority, firms, workers, or any combination of them), and the potential existence of certification. A crucial factor is also given by the average educational attainment in the population: since training is more likely when educational attainment is higher, an educated labour force constitutes a favourable precondition for the development of general and firm specific skills.[64] Eventually, the way in which the secondary school system is organised matters: some educational systems are characterised by early tracking (Germany, Italy), while others provide comprehensive education (United Kingdom, United States). If tracking reduces workers' adaptability, then some national systems may be more favourable to general human capital formation than others.[65]

Table 3 – National training systems

Training system	*Countries*	*Essential features*
Apprenticeship	Germany, United Kingdom (before 1980), Netherlands	- codetermination (firms, trade unions, local authorities) - cost sharing - skill certification - good results in schooling activities are taken into account when firms hire apprentices.
Within firm training (low turn-over)	Japan	- life employment reduces job turn-over - firms provide both general and specific training - on the job-training - skills are homogenous at school exit
Vocational training in public schools	Sweden, Norway, United Kingdom (after 1980)	- funding from the central government - training contents are defined at school level
Training financed by specific taxes on firms	France, Australia	- the burden is spread across a wide range of taxpayers - no guarantee that low-skill workers and/or employee in small-size firms will be trained
Vocational training in public schools and/or on the job training	United States, Canada	- greater variety in training possibility - lack of nation-wide certification - most of the training is firm-specific

Source: adapted from Table 2 in Lynch 1994

6. Measuring the return to education

In the previous section we discussed alternative explanations for the positive correlation between acquired education and earnings. If we are interested in measuring the extent of monetary returns (for

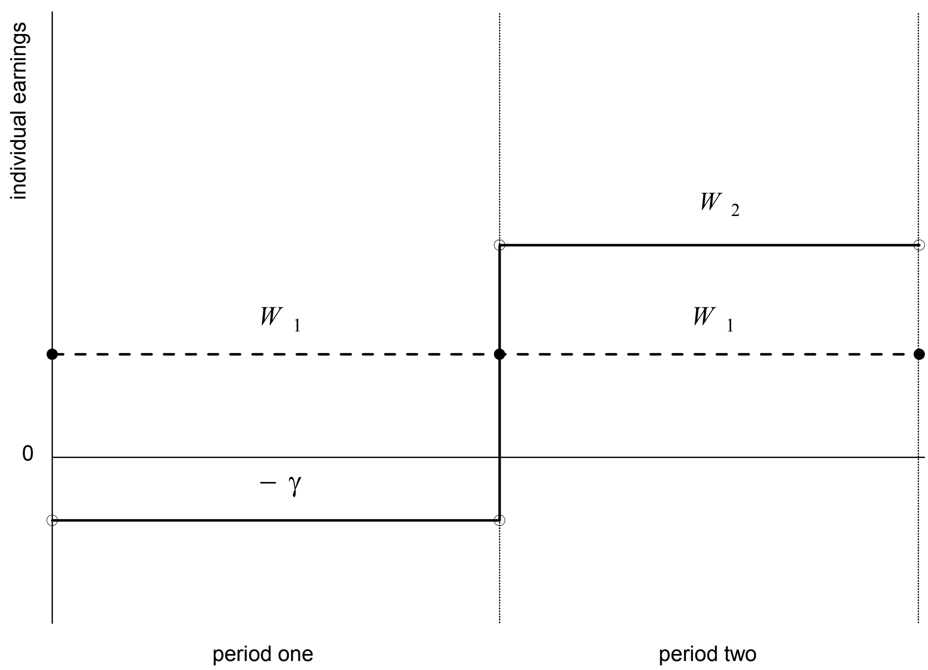
[63] Soskice in Lynch 1994 stresses this point. Acemoglu and Pischke 2000 present a model where certification provided by an external authority is capable to partially solve the asymmetry of information between the training firm and its competitors who cannot observe the effective outcome of training on workers' productivity. An equilibrium with training is Pareto superior to an equilibrium without training, and in a symmetrical framework the firms, who possess an informational advantage over workers' productivity, pay the cost of training.

[64] See Brunello 2003 for an empirical analysis of this link. The idea that an educated labour force is favoured by employers because of it is more easily trained was originally proposed by Thurow 1975.

[65] See Brunello and Giannini 2003.

example to make comparisons among individuals from different age cohorts, birth regions or countries), we resort to a general principle in financial economics, stating that the internal rate of return of any investment project corresponds to the discount rate that equalises the (discounted) flow of revenues to the (discounted) flow of outlays.[66] In order to explain this general principle, we consider the simplest case of a person living just two periods and facing the following alternative: either not going to school and working in both periods as unskilled worker, obtaining a wage W_1 (assumed constant in both periods); or attending school with a direct cost γ in the first period, and working as skilled worker in the second period for a wage W_2.[67] It is obvious that there exists a monetary incentive to acquire education if and only if $W_2 > W_1$. The different wage profiles associated to these two alternatives are reported in figure 3, where the dashed line (denoted by the ● bullet) corresponds to the unskilled profile, whereas the solid line (denoted by the ○ bullet) describes the skilled profile.

Figure 3 – Alternative earnings profiles



If we now wonder what is the internal rate of return associated to the choice of attending school in the first period, we have to take into account costs and benefits. The former is given by the algebraic sum of direct and indirect costs $(\gamma + W_1)$, the latter is given by the skilled-unskilled wage premium $\Delta W = W_2 - W_1$. Since the costs are incurred in the first period, while the benefits accrue in the second period, we discount the benefits with the discount rate $(1+\beta)$. By equalising costs and benefits we get

[66] An alternative way to express the same concept is to identify the discount rate that sets to zero the net present value of the investment, where the net present value is defined as the discounted flow of all differences between inflows and outflows.

[67] We neglect the possibility of different employment probabilities associated to different educational attainments, but the wages could easily be reinterpreted as expected wages.

$$\gamma + W_1 = \frac{W_2 - W_1}{1+\beta} \tag{6.26}$$

or alternatively we introduce the *net present value* (NPV) of the investment project associated to attending school defined as

$$NPV = \underbrace{-(\gamma + W_1)}_{\text{costs}} + \underbrace{\frac{W_2 - W_1}{1+\beta}}_{\text{benefits}} \tag{6.27}$$

The internal rate of return is given by the value of β that satisfies equality (6.26) (or equivalently that sets $NPV = 0$), that is

$$\beta = \frac{W_2 - W_1}{\gamma + W_1} - 1 \tag{6.28}$$

The equation (6.28) shows that the return to education rises with the benefits (the numerator) and declines with the costs (the denominator). The general principle can be obviously extended to a multiperiod context. If we consider the case of a given school order lasting m years, a direct cost of school attendance γ_t, wage rates associated to the choice of not attending or attending s years of schooling (respectively W_t and W_t^s) and a working life lasting n years, then the return to education associated to school attendance β satisfies

$$\sum_{t=1}^{s}\left[\frac{\gamma_t + W_t}{(1+\beta)^{t-1}}\right] = \sum_{t=s+1}^{n}\left[\frac{W_t^s - W_t}{(1+\beta)^{t-1}}\right] \tag{6.29}$$

Equation (6.29) is the multiperiod equivalent of equation (6.26).[68] Since it only includes costs and benefits born by an individual undertaking the choice of school attendance, β represents a *private rate of return* that does not take into account costs and benefits accruing to the society as a whole.[69] If we now add to the left hand side of equation (6.29) the cost E_t born by public budget for school attendance, and to the right hand side the potential benefits associated to better-educated labour force, we obtain a measure a social rate of return β_{soc}

$$\sum_{t=1}^{s}\left[\frac{\gamma_t + W_t + E_t}{(1+\beta_{soc})^{t-1}}\right] = \sum_{t=s+1}^{n}\left[\frac{W_t^s - W_t}{(1+\beta_{soc})^{t-1}}\right] + externality \tag{6.30}$$

---

[68] However equation (6.29) implies an additional complication, since it describes a polynomial of $(n-1)$-th degree in $(1+\beta)$, thus admitting $(n-1)$ roots. The problem of roots indeterminacy in returns is well known in the financial literature, which has not yet provided a convincing explanation of the common practice of selecting the positive real value that is closest to unity.

[69] A further simplifying assumption implicit in equation (6.29) is given by the fact that each individual takes the wage rates W_t and W_t^s as given when considering the choice of school attendance. While this is reasonable on an individual basis, it does not hold in the aggregate, because whenever all individuals find it convenient (and select) to attend s years of schooling, the expected gain $(W_t^s - W_t)$ declines, with a potential reversal of intended choices.

It is evident that $(1+\beta_{soc})<(1+\beta_{priv})$ whenever there exists public subsidisation of education (i.e. $E_t > 0$) and externalities are hard to grasp. However, if for example growth externalities are significant (because a better educated labour force is associated with higher participation rates, lower unemployment rates and/or allows for the introduction of more productive technologies), the social rate of return may exceed the private one.

The calculations implied by equations (6.29) and (6.30) (often indicated as *full* or *integral method*) are cumbersome, since they require information about educational expenditures and wages over the entire life span of individuals. For this reason, whenever possible economists have resorted to alternative methods to measure the private return to education. The most common strategy is based on estimates of the determinants of individual earnings (typically indicated as *Mincerian function* from Mincer (1974) who originally proposed it). Table 4 shows private and social rates of return, computed under alternative methods. It is easy to recognise that private returns constantly exceed social ones, with a gap that expands at higher school orders. In addition, in accordance with the decreasing marginal productivity assumption, these returns tend to decline with the increase in the level of the educational attainment.

Table 4 – Returns to education – private and social – regional averages of published estimates

	integral method						earning function	
	private return to education			social return to education			avrg years education	private return
	primary	secondary	tertiary	primary	secondary	tertiary		
Sub-Saharan Africa	41.3	26.6	27.8	24.3	18.2	11.2	5.9	13.4
Asia	39.0	18.9	19.9	19.9	13.3	11.7	8.4	9.6
Europe/Middle east and North Africa	17.4	15.9	21.7	15.5	11.2	10.6	8.5	8.2
Latin America and the Caribbean	26.2	16.8	19.7	17.9	12.8	12.3	7.9	12.4
OECD countries	21.7	12.4	12.3	14.4	10.2	8.7	10.9	6.8
World	*29.1*	*18.1*	*20.3*	*18.4*	*13.1*	*10.9*	*8.4*	*10.1*

Source: Tables 1 e 4 in Psacharopulos 1994

In order to show the implicit assumptions underlying the earning function approach, let us go back to equation (6.29) and introduce the following hypotheses:

i) direct private costs for acquiring education are mainly given by foregone incomes (alternatively, labour incomes obtained during the educational career just match the direct costs of attendance - $\gamma_t = 0,\ t = 1,...,s$).

ii) the working life has to be identical for everyone, irrespective of the number of years of school attendance.[70]

In such a case, equation (6.29) can be re-expressed as

$$\sum_{t=1}^{s}\left[\frac{W_t}{(1+\beta)^{t-1}}\right]=\sum_{t=s+1}^{n+s}\left[\frac{W_t^s - W_t}{(1+\beta)^{t-1}}\right] \tag{6.31}$$

or rearranging terms

$$\sum_{t=1}^{n}\left[\frac{W_t}{(1+\beta)^{t-1}}\right]=\sum_{t=s+1}^{n+s}\left[\frac{W_t^s}{(1+\beta)^{t-1}}\right] \tag{6.32}$$

[70] This is equivalent to saying that if an uneducated worker stops working at the age of n, a worker who has attended s years of school can work up to the age of $(n+s)$. Many pension schemes are consistent with this assumption, under the rationale that individuals who spent part of their lives investing in their education should be given time enough to recover their investment outlays.

Equation (6.32) compares the earnings profile of an uneducated worker (left hand side) with the profile of an educated worker (right hand side); it suggests that in order to observe both choices in a population of otherwise identical individuals, at the margin they must offer the same returns.

Adding the further hypothesis of
iii) constant wages across the life cycle[71], equation (6.32) can be rewritten as

$$W = \frac{W^s}{(1+\beta)^s} \tag{6.33}$$

Taking the logarithms of both sides and rearranging terms we get

$$\log\left(W^s\right) = \log(W) + s \cdot \log(1+\beta) \cong \log(W) + s \cdot \beta \tag{6.34}$$

Recalling that log differences approximates percent variations, equation (6.34) indicates that percent wage differences between two individuals who attended a different number of years of schooling must be proportional to the gap in years of schooling, and the proportionality factor corresponds to the internal rate of return.[72] Equation (6.34) is known in the literature as the base of the *Mincerian equation*, and it allows the direct estimate of the rate of return to education using data from a representative sample of working population (once work experience is taken into account). If we can identify the factors that systematically affect individual earnings independently from acquired education (for example age, gender, ethnicity, marital status, local labour market conditions, residential area, and so on), and we generically indicate them with the vector $\mathbf{Z}_i$, equation (6.34) can be replaced by

$$\log\left(W_i^s\right) = \boldsymbol{\alpha}'\mathbf{Z}_i + \beta \cdot s_i + \varepsilon_i \tag{6.35}$$

Equation (6.35) indicates that the labour income of individual i depends on systematic factors $\mathbf{Z}_i$ affecting sample population, on educational attainment s_i and on an idiosyncratic component ε_i which cannot be predicted by an external observer (like chance or unobservable ability).[73] Equation (6.35) is consistent with equation (6.34) since two generic and otherwise identical individuals i and j (such that $\mathbf{Z}_i = \mathbf{Z}_j$) exhibit an earnings gap that is systematically proportional to the difference in their educational attainments

$$\log(W_i) - \log(W_j) = \beta \cdot s_i - \beta \cdot s_j + \varepsilon_i - \varepsilon_j = \beta \cdot (s_i - s_j) + (\varepsilon_i - \varepsilon_j) \tag{6.36}$$

---

[71] At the cost of some additional analytical complication, we could consider the case of wages growing at constant rates.

[72] For readers who are more familiar with continuous time analysis, equation (6.32) can be rewritten as

$$\int_0^n W(t) \cdot e^{-\beta t} dt = \int_s^{n+s} W^s(t) \cdot e^{-\beta t} dt$$

When wages are constant over the life cycle, the two integral can be solved, yielding

$$W\left[-\frac{e^{-\beta n} - 1}{\beta}\right] = W^s\left[-\frac{e^{-\beta(n+s)} - e^{-\beta s}}{\beta}\right] \qquad \Leftrightarrow \qquad \frac{W^s}{W} = e^{\beta s}$$

which leads to equation (6.34) without approximations.

[73] The unsystematic component ε_i must have zero mean $E[\varepsilon] = 0$. If that were not the case, any systematic component should be included in the $\mathbf{Z}_i$ vector, which already includes a constant.

By taking sample expectations on both sides of equation (6.36) one may retrieve a sample estimate of the return to education $\hat{\beta}$, but this invokes two further assumptions. The first one is just a corollary of the previous statement of "otherwise identical" individuals: if they are identical they face similar work opportunities, they possess identical abilities and therefore experience the same return to education.[74] When this is not the case (i.e. when $\beta_i \neq \beta_j$), we can only obtain estimates of the returns for population subgroups, which are specifically affected by exogenous variations (see next section). The second assumption derives from extending the bilateral comparison between two generic educational attainments, s_i and s_j, to all potential comparisons, while maintaining the same return rate. This is equivalent to postulating a (semi-log) linear relationship between earnings and years of education, which does not seem contradicted by empirical evidence.[75] In a similar way, one may want to consider the role of work experience and/or job tenure as specific determinants of earnings. In order to take into account the inverted U-shaped wage-life profile, it has become common practice to estimate equation (6.35) in the following form

$$\log\left(W_i^s\right) = \boldsymbol{\alpha}'\mathbf{Z}_i + \beta \cdot s_i + \theta_0 \cdot e_i + \theta_1 \cdot e_i^2 + \varepsilon_i \tag{6.37}$$

where e_i measures the work experience of individual i. When this piece of information is absent in a dataset, economists have resorted to the concept of potential experience (equal to age less schooling + starting age of schooling) or directly to age, to avoid potential collinearity with schooling. Independently from the way in which experience is measured, $\left(\theta_0 + 2\theta_1\bar{e}\right)$ represents the marginal return to experience (at sample mean $\bar{e}$) and $\frac{\theta_0}{2\theta_1}$ the age at which the average individual achieve her maximum earning along her working life. The way in which experience is measured has a direct impact on the measured return to education, as it can be assessed by looking at table 5 reporting least square estimates of the return to education β under alternative definitions of work experience (potential experience, actual experience, when available, and age). In all cases the variable used for experience is considered in levels and squares.

We now move to the econometric problems posed by the attempt to estimate the return to education from equation (6.35) in a representative sample of a working population.

---

[74] There is an even stronger presumption in the application of condition (6.29), because each individual has to compare two alternative wage profiles, as if she had been able to conduct two alternative lives. Since counterfactuals cannot be observed in reality, agents make inferences from observing life experiences of similar individuals.

[75] See for example figure 1 Krueger and Lindahl 2001.

Tab.5 – Return to education in Europe, by countries – mid 1990's – OLS

	men			women		
	exp.pot.	exp.eff.	age	exp.pot.	exp.eff.	age
Austria (95)	0.069	==	0.059	0.067	==	0.058
Denmark (95)	0.064	0.061	0.056	0.049	0.043	0.044
West Germany (95)	0.079	0.077	0.067	0.098	0.095	0.087
Netherlands (96)	0.063	0.057	0.045	0.051	0.042	0.037
Portugal (94)(95)	0.097	0.100	0.079	0.097	0.104	0.077
Sweden (91)	0.041	0.041	0.033	0.038	0.037	0.033
France (95)	0.075	==	0.057	0.081	==	0.065
United Kingdom (94-96)	0.094	0.096	0.079	0.115	0.122	0.108
Ireland (94)	0.077	0.068	0.050	0.105	0.100	0.089
Italy (95)	0.062	0.058	0.047	0.077	0.070	0.061
Norway	0.046	0.045	0.037	0.050	0.047	0.044
Finland (93)	0.086	0.085	0.072	0.088	0.087	0.082
Spain (94)	0.072	0.069	0.055	0.084	0.079	0.063
Switzerland (95)	0.089	0.088	0.075	0.092	0.086	0.082
Greece (94)	0.063	==	0.040	0.086	==	0.064
Mean	*0.072*	*0.070*	*0.057*	*0.079*	*0.076*	*0.066*

Note: in brackets the sample year from which the estimates are obtained.
Source: table 4 in Brunello, Comi and Lucifora 2001

7. Estimating the return to education

A direct measure of the private return to education can easily be obtained by an ordinary least square estimator (OLS) applied to a sample of working individuals

$$\hat{\beta}_{OLS} = \frac{Cov(\log(W),s)}{Var(s)} \tag{6.38}$$

However, the pioneering paper by Griliches (1977) quickly warned that OLS estimates could be biased and inconsistent, and therefore unreliable for policy advice. Griliches disputed the - at that time - general opinion of an upward bias in OLS estimates, arguing for the equivalent possibility of a downward bias. More recently, Card (1999) has expressed the opinion that downward and upward biases could almost offset each other, thus restoring some trust in OLS estimates.

Let us review the sources of potential bias, which we generically indicate as problems of regressor endogeneity.[76] We summarise the problem at hand with the following two equations

$$s_i = \boldsymbol{\mu}'\mathbf{Z}_i + \eta_i \tag{6.39}$$
$$\log(W_i) = \boldsymbol{\alpha}'\mathbf{Z}_i + \beta \cdot s_i + \varepsilon_i \tag{6.40}$$

where $\mathbf{Z}_i$ is a vector of individual characteristics of individual i, s_i is the amount of education (measured by years of school attendance), W_i is individual earnings, η_i and ε_i are error components. We assume that $E(\mathbf{Z}_i\eta_i) = E(\mathbf{Z}_i\varepsilon_i) = 0$, that is all individual characteristics but education are orthogonal with the residuals of equations (6.39) and (6.40). The problem of endogeneity in the case of the schooling variable s_i derives from its correlation with the residual η_i, that is from $Cov(\varepsilon_i, \eta_i) \neq 0$. Whenever $Cov(\varepsilon_i, \eta_i) \neq 0$ applies, the ordinary least square estimate of the return to education β are biased and inconsistent. Endogeneity comes from three potential sources: omitted variables, measurement errors and heterogeneity of returns in the population (Griliches 1977, Card 1999).

[76] The current exposition is based on Card 1995.

The case of omitted variable potentially applies whenever the researcher is unable to control for family background and/or for individual ability, because both groups of variables could raise earnings independently from human capital variables (education, experience) and from other controls in the wage function. A typical example is given by unobservable ability: more talented persons achieve more education because it is easier for them to do so, and at the same time they are more productive when working. If we indicate individual ability with A_i, its omission implies that the error component in equation (6.40) consists of

$$\varepsilon_i = \delta A_i + \sigma_i \tag{6.41}$$

But if ability also affects educational attainment we have

$$\eta_i = \lambda A_i + \omega_i \tag{6.42}$$

It is obvious from (6.41) and (6.42) that $Cov(\varepsilon_i, \eta_i) \neq 0$, thus providing biased estimates of the β coefficient; in fact from (6.38) we see

$$\begin{aligned}\hat{\beta}_{OLS} &= \frac{Cov(\log(W), s)}{Var(s)} = \frac{Cov(\boldsymbol{\alpha}'\mathbf{Z} + \beta s + \varepsilon, s)}{Var(s)} = 0 + \beta \frac{Cov(s,s)}{Var(s)} + \frac{Cov(\varepsilon, \boldsymbol{\mu}'\mathbf{Z} + \eta)}{Var(s)} = \\ &= \beta + \frac{Cov(\delta A + \sigma, \lambda A + \omega)}{Var(s)} = \beta + \delta\lambda \frac{Var(A)}{Var(s)} \neq \beta\end{aligned} \tag{6.43}$$

It is easy to see that the sign of the bias depends on the signs of δ and λ. The first one is expected to be positive (abler individuals are more productive, and therefore better rewarded), while the latter has uncertain sign. On one side it could be positive, since more intelligent and disciplined persons also perform as more brilliant students, thus achieving longer schooling. In such a case the OLS estimate will be upward biased. However, we could also have the case of a negative λ coefficient, whenever better endowed individuals face a higher opportunity cost in attending school, and they may end up leaving school earlier. In such a case, the OLS estimate will be downward biased. Similarly, ambiguous conclusions arise when we take into account that educational investment decisions are taken by parents. On one side, efficiency considerations suggest that parents want to invest more in more talented children (λ positive); on the other, equity considerations could produce the opposite result, whenever parents tend to compensate with financial resources the differences in ability endowments among their children (λ negative). Overall, in case of omitted variables related to ability and/or parental investment, a priori we are unable to assess the direction of the bias in the OLS estimate of the return to education.

Measurement error is a second source of endogeneity. It implies that the measure for educational attainment s that we observe is equal to the true measure s^* except for an error component $\chi \sim (0, \sigma_\chi^2)$, that is

$$s_i = s_i^* + \chi_i \tag{6.44}$$

If we have $\sigma_\chi^2 > 0$, the OLS estimate of β is biased and inconsistent. In fact

$$\hat{\beta}_{OLS} = \frac{Cov(\log(W), s)}{Var(s)} = \frac{Cov(\boldsymbol{\alpha}'\mathbf{Z} + \beta s - \beta\chi + \varepsilon, s)}{Var(s)} = \\ = 0 + \beta\frac{Cov(s,s)}{Var(s)} - \beta\frac{Cov(\chi, s^* + \chi)}{Var(s^* + \chi)} + 0 = \beta\left(1 - \frac{Var(\chi)}{Var(s^*) + Var(\chi)}\right) < \beta \tag{6.45}$$

The factor $\frac{Var(s^*)}{Var(s^*) + Var(\chi)}$ is indicated in the literature as the *reliability factor* of observed schooling.[77] In such a case OLS estimates will be downward biased, the extent of the bias depending on the magnitude of variance of the measurement error.[78]

Eventually, the third source of bias derives from heterogeneity of the coefficient to be estimated in the population. According to Card (1994) there are two potential sources of such heterogeneity. The first one originates from the fact that differences in abilities reflect into differences in productivities, such that abler individuals face a higher return schedule for identical amount of education (*ability bias*). The second comes from the fact that under financial market imperfections differences in family backgrounds imply different marginal costs in acquiring education, so that children from poorer families face a higher cost for education (*cost bias*). The consequence of both distortions is that the subset of population with low educational attainment will be composed of individuals with lower returns (less able) and by individuals facing higher costs (poorer background). Since the underlying model implies that each individual will optimally select the amount of education that will equate her expected return to her marginal cost, the population estimate of the return of education will depend on subgroup composition. If the group of less able individuals prevails, we observe a positive correlation between education and error component ε in the wage function, and therefore the OLS estimate will be upward biased. Otherwise, when the group of individuals from poorer families prevails, the opposite situation will occur, and we will observe a downward bias.

An application of the last case deals with the problem of sample distortion due to participation decision. When we regress labour income onto educational attainment, we are typically including the population sub sample holding a paid job. But what about when participation in the labour market is affected by sample selection ? A typical case is given by the female component in the labour supply. If only abler women enter the labour market, the estimate of the return to education is upward biased with respect to an average effect for the entire population. An econometric technique to deal with this distortion has been proposed by Heckman (1979), and is based of the idea of modelling the selection process.[79] The crucial issue following this approach is the possibility of modelling the participation decision in (at least partially) independent way from the determinants of the return to education. In the absence of this identifying restriction, the decision to participate is indistinguishable from an econometric point of view, the sample selection problem cannot be taken into account properly, and least square estimates of the return to education cannot be used to make inference to the entire population.

---

77 It corresponds to the linear projection of true schooling on observed schooling under the additional assumption of $Cov(s^*, \chi) = 0$. See Krueger and Lindhal 2001 and Card 2001.

78 A direct assessment of the measurement error is obtained from the comparison between self reported attainment and administrative records. Alternatively, it can be obtained from interviewing relatives, for example asking each twin about the educational attainment of the other. In such a case the correlation between self-reported and sibling obtained measure is in the order of 0.70 (see Ashenfelter and Rouse 1998 e Miller, Mulvey and Martin 1997). This is consistent with the following statement: "Research in the US over the past three decades has concluded that the reliability of self-reported schooling is 85-90 percent (Angrist and Krueger 1999, table 9), implying that the downward biases in the order of 10-15 percent – enough to offset a modest upward ability bias". (Card 2001, p.1134)

79 See Angrist and Krueger 1999 for a review of applications of the Heckman two-step procedure.

References

Acemoglu, D. 1999. Changes in unemployment and wage inequality: an alternative theory and some evidence. *American Economic Review* 89: 1259-1278.

Acemoglu, D. 2003. Cross-country inequality trends. *Economic Journal* 113(485): 121f-149f.

Acemoglu, D. and J.Pischke. 1999. Beyond Becker: training in imperfect labor market. *Economic Journal* 109: F112-F142.

Acemoglu, D. and J.Pischke. 2000. Certification of training and training outcomes. *European Economic Review* 44: 917-927.

Aghion, P. and P.Howitt. 1998. *Endogenous growth theory*. MIT Press, Cambridge MA.

Altonji, J. and T.Dunn. 1996. The Effects of Family Characteristics on the Returns to Schooling. *Review of Economics and Statistics* 78(4): 692–704.

Angrist, J. and A.Krueger. 1999. Empirical strategies in labor economics. in Ashenfelter, O. and D.Card (eds). *Handbook of Labor Economics* – vol.3. Amsterdam: North Holland.

Bowles, S. and H.Gintis. 2000. Does schooling raise earnings by making people smarter ? in Arrow, K., S.Bowles and S.Durlauf. (eds) 2000. *Meritocracy and economic inequality*. Princeton University Press.

Arrow, K. 1962. The economic implications of learning by doing. *Review of Economic Studies* 29: 155-173.

Ashenfelter, O. and C.Rouse. 1998. Income, schooling and ability: evidence from a new sample of identical twins. *Quarterly Journal of Economics* 113:253-284

Ashenfelter, O. and C.Rouse. 2000. Schooling, intelligence and income in America. in Arrow, K., S.Bowles and S.Durlauf. (eds) 2000. *Meritocracy and economic inequality*. Princeton University Press.

Barro, R. 1997. *Determinants of economic growth - a cross empirical study*. MIT Press, Cambridge Mass.

Bassanini, A. and G.Brunello. 2003. Is Training More Frequent when Wage Compression is Higher ? Evidence from the European Community Household Panel. IZA Discussion Paper n.839.

Becker, G. 1993. *Human capital: a theoretical and empirical analysis, with special reference to education*. University of Chicago Press, Chicago (first edition: 1964).

Bedard, K. 2001. Human capital versus signaling models: university access and high school dropouts. *Journal of Political economy* 109(4): 749-775.

Belzil, C., and J.Hansen. 2002. Unobserved ability and the return to schooling. IZA Discussion Paper n.508.

Bils, M. and P.Klenow. 1998. Does schooling cause growth or the other way round ? NBER working paper n.6393

Blöndal, S., S.Field and N.Girouard. 2002. Investment in Human Capital Through Post-Compulsory Education and Training. OECD Economic Department Working Paper n.333.

Bowles, S. 1985. The production process in a competitive economy: Walrasian, Neo-Hobbesian, and Marxian models. *American Economic Review* 75(1): 16-36.

Bowles, S. and H.Gintis. 1975. The problem with human capital theory - a Marxian critique. *American Economic Review* 65(P&P): 74-82.

Bowles, S. and H.Gintis. 1976. *Schooling in capitalist America*, Routledge & Kegan, London

Bowles, S., H.Gintis and B.Gustafsson. (eds) 1993. *Markets and democracy: participation, accountability and efficiency*. Cambridge University Press

Bowles, S., H.Gintis and M.Osborne. 2001a. The determinants of earnings: a behavioural approach. *Journal of Economic Literature* 39: 1137-76.

Bowles, S., H.Gintis and M.Osborne. 2001b. Incentive enhancing preferences: personality, behavior and earnings. *American Economic Review* 91(May): 155-8.

Brown, S., and J.Sessions. 1999. Education and economic status: a test of the strong screening hypothesis in Italy. *Economics of Education Review* 18: 397-404

Brunello, G. 2003. On the complementarity between education and training in Europe. in D.Checchi and C.Lucifora (eds). 2003. *Educationa, training and labour market outcomes in Europe*. Palgrave

Brunello, G., S.Comi and C.Lucifora. 2001. The Returns to Education in Italy: a New Look at the Evidence. in Harmon, C., I.Walker and N.W.Nielsen (eds.). *The Returns to Education in Europe*. Edward Elgar, London, 2001.

Brunello, G and S.Comi. 2003. Education and Earnings Growth. Evidence from 11 European Countries. forthcoming in *Economics of Education Review*

Brunello, G. and M.Giannini. 2003. Stratified or comprehensive? The economic efficiency of school design. forthcoming in *Scottish Journal of Political Economy.*

Cameron, S., and J.Heckman. 2001. The Dynamics of Educational Attainment for Black, Hispanic, and White Males. *Journal of Political Economy.* 109(3): 455-99.

Card, D. 1995. Earnings, schooling and ability revisited. *Research in labor economics*. Volume 14. Polachek, Solomon W., ed., Greenwich, Conn. and London: JAI Press. p 23-48. 1995.

Card, D. 1999. The causal effect of education on earnings. in Ashenfelter, O. and D.Card (eds). *Handbook of Labor Economics* – vol.3. Amsterdam: North Holland.

Card, D. 2001. Estimating the return to schooling: progress on some persistent econometric problem. *Econometrica* 69: 1127-60.

Chevalier, A., C.Harmon, I.Walker and Y.Zhu. 2003. Does education raise productivity or just reflect it ? CEPR Discussion Paper n.3993, July

Freeman, R. 1986. Demand for education, in Ashenfelter, O. and R.Layard (eds). 1986. *Handbook of labour economics*, Amsterdam: North Holland

Gemmell, N. 1996. Evaluating the impacts of human capital stocks and accumulation on economic growth: some new evidence. *Oxford Bulletin of Economics and Statistics* 58(1): 9-28.

Giannini, M. 2001. Education and Job Market Signalling: How Robust Is the Nexus?. *Economia Politica.* 18(1): 41-54.

Goldberger, A., and C.Manski. 1995. Review article: The Bell Curve by Herrnstein and Murray. *Journal of Economic Literature* 33: 762-76.

Green, D. and W.Riddell. 2003. Literacy and earnings: an investigation of the interaction of cognitive and unobserved skills in earnings generation. *Labour Economics* 100(2): 165-84

Griliches, Z. 1977. Estimating the Returns to Schooling: Some Econometric Problems. *Econometrica*, 45(1): 1-22.

Groot, W., and H.Oosterbeek. 1994. Earning effects of different components of schooling; human capital versus screening. *Review of Economics and Statistics* 76(2): 317-21.

Harmon, C., V.Hogan and I.Walker. 2003. Dispersion in the economic return to schooling. *Labour Economics* 100(2): 205-15.

Harmon, C., H.Oosterbeek and I.Walker. 2003. The returns to education: microeconomics. *Journal of Economic Surveys* 17(2): 115-156.

Heckman, J. 1979. Sample specification bias as a specification error. *Econometrica* 47: 153-161.

Krueger, A. 2002. Inequality, too much of a good thing. Princeton University. Industrial Relations Section wp n.466.

Krueger, A. and M.Lindahl. 2001. Education for growth: why and for whom ? *Journal of Economic Literature* 39: 1101-36.

Lang, K., and D.Kropp. 1986. Human capital versus sorting: the effects of compulsory attendance laws. *Quarterly Journal of Economics* 101(3): 609-24.

Layard, R., and G.Psacharopulos. 1974. The screening hypothesis and the returns to education. *Journal of Political Economy.* 82: 985-98.

Layard, R., S.Nickell and R.Jackman. 1991. *Unemployment.* Oxford University Press, Oxford.

Lazear, E. 1995. *Personnel Economics.* MIT Press.

Lazear, E. 2004. The Peter principle: a theory of decline. *Journal of Political Economy* 112(1): S141-S163.

Lucas, R. 1988. On the mechanics of economic development. *Journal of Monetary Economics* 22: 3-42.

Lynch, L. 1994. (ed). *Training and the private sector: international comparisons.* Chicago. University of Chicago Press.

Miller, P., C.Mulvey C. and N.Martin. 1997. Family characteristics and the returns to schooling: evidence on gender differences from a sample of Australian twins, *Economica* 64: 119-136.

Mincer, J. 1974. *Schooling, Experience, and Earnings*. New York: NBER.
Murnane,R., J.Willett and F.Levy. 1995. The Growing Importance of Cognitive Skills in Wage Determination. *Review of Economics and Statistics* 77(2) : 251–66.
Murnane, R., J.Willett, M.Braatz and Y.Duhaldeborde. 2001. Do Different Dimensions of Male High School Students' Skills Predict Labor Market Success a Decade Later? Evidence from the NLSY. *Economics of Education Review* 20 (4): 311-320.
Murphy, K., and F.Welch. 1992. The structure of wages. *Quarterly Journal of Economics* 10(4): 283-96.
OECD 2001. *Education at a glance*. Paris
Riley, J. 1979. Testing the Educational Screening Hypothesis. *Journal of Political Economy* 87(5): s227-52.
Riley, J. 2001. Silver signals: twenty-five years of screening and signaling. *Journal of Economic Literature* 39: 432-78.
Romer, P. M. 1990a. Endogenous Technological Change. *Journal of Political Economy* 98: S71-S102.
Romer, P. M. 1990b. Human capital and growth: theory and evidence. *Carnegie-Rochester conference series on public policy* 32: 251-285.
Shapiro,C. and J.Stiglitz. 1984. Equilibrium Unemployment as a Worker Discipline Device. *American Economic Review* 74(3): 433–44.
Sianesi, B. and J.Van Reenen. 2003. The returns to education: macroeconomics. *Journal of Economic Surveys* 17(2): 157-200.
Spence, M. 1973. Job market signalling. *Quarterly Journal of Economics* 87: 355-379.
Stiglitz, J. 1975. The theory of "screening", education and the distribution of income. *American Economic Review* 64/3: 283-300.
Teulings, C. and T.vanRens 2002. Education, growth and income inequality. Timbergen Institute discussion paper n.001/3
Thurow, L. 1975. *Generating inequalities.* New York. Basic Book.
Topel, R. 1999. Labour Market and Economic Growth, in Ashenfelter, O. and D.Card (eds.). *Handbook of Labour Economics* – vol.3, Amsterdam: North Holland.
Trostel,P., I.Walker and P.Wooley. 2002. Estimates of the econometric return to schooling for 28 countries. *Labour Economics* 9(1): 1-16.
Weiss, A. 1990. *Efficiency wages: models of unemployment, layoffs, and wage dispersion.* Princeton University Press, Princeton
Weiss, A. 1995. Human capital vs. signalling explanation of wages. *Journal of Economic Perspectives* 9(4): 133-54.
Willis, R. 1986. Wage determinants: a survey and reinterpretation of the human capital earning functions. in Ashenfelter, O. and R.Layard (eds), *Handbook of labour economics*, Amsterdam: North Holland.