A pearl of great price, the free education system of Sri Lanka
Ranasinghe, A.

Citation for published version (APA):
Ranasinghe, A. (1999). A pearl of great price, the free education system of Sri Lanka
Chapter 2

Economics of Education: A Theoretical Background

2.1 Introduction

Because the main objective of this book is to explore the role of family background and varying school quality on individual performance of schooling, this chapter is devoted to a review of the main literature related to the issues. It is not intended as a complete and extensive review; it merely sets the stage by presenting relevant considerations.

The chapter is organised into eight sections including the introduction. Since most of the literature reviewed in this chapter is somehow related to the earnings function, section 2 presents a short review of the earnings function. Section 3 reviews the major theoretical links between earnings and schooling. Various theories on demand for education are reviewed in section 4, while section 5 describes the issues related to the education production function. Section 6 considers demand for schools. Given that the individual acquisition of education is correlated with individual as well as family characteristics, section 7 reviews the pros and cons of alternative education policies for breaking the vicious-circle between family background and education acquisition and reviews the major theoretical contributions to education policies. Section 8 concludes the chapter.

2.2 Earnings Function

According to the human capital theory\(^1\), one’s potential earnings, \(E_t\), is given by

\[
\ln E_t = \ln E_0 + r \sum_{i=1}^{t} C_{ij},
\]

where \(\ln E\) is the natural logarithm of the potential earnings, \(C\) is the cost of education, subscript “\(i\)” represents the individual, and “\(t\)” represents time. The equation also assumes that the rate of return, \(r\) is constant. Since the direct cost of education is not observed, the general practice is to use “a time-equivalent of \(C_t\)”:

\[
s_t = C_t / E_t.
\]

One of the main components of the cost of education is the foregone earnings. Then the above time equivalent is approximately equal to the proportion of one’s time invested in human capital production. Following Ben-Porath (1967) [cf. section 2.4 below] we can assume that \(s_t\) is one for full-time students. Furthermore, it is less than one and decreasing over time during the post-schooling period. Assuming different rates of return for schooling and post-school investment and with a parabolic approximation of the decreasing time-equivalent of post-school investments, we can write that potential earnings is an increasing function of years of full-time education (schooling), labour market experience and its square. Since the observed earnings, \(W_t = (1 - s_t) E_t\), we can

---

\(^1\) See section 2.4 below for a detailed elaboration of the human capital theory.
approximate \ln E_t with \ln W_t with a suitable re-scaling of the earnings function. [See inter-alia, Polachek and Siebert (1993), Chapter 4].

\[ \ln W_t = \alpha_0 + \alpha_1 S_t + \alpha_2 X_t + \alpha_3 X_t^2 + \varepsilon_t \] (2.1)

where \ln W is the natural logarithm of the observed earnings, S is years of schooling and X is the labour-market experience. \( \alpha_0 \) is the initial earnings, \( \alpha_1 \) is the rate of return for schooling, \( \alpha_2 + 2\alpha_3 X_t \) is the rate of return for post-school investment, and \( \varepsilon \) is a random variable representing all other effects on earnings. This is the Mincerian earnings equation, named after Jacob Mincer, the inventor of this theoretical equation [See, Mincer (1970 & 1974)]. Currently, there are many critiques of interpreting equation (2.1) as a causal relationship between earnings and the variables listed in the right-hand side. [See Card (1994) for an excellent review of the alternative estimation methods of the earnings function]. Many versions of equation (2.1) have been estimated for many countries and they generally find that the length of schooling has a key role in determining the level and the dispersion of earnings. The immediate interest of this finding for education economists is the mechanism that determines this relationship. Why does education enhance earnings? There are two broad theoretical interpretations for this association: the human capital theory [See inter-alia, Mincer (1970 & 1974)] and the signalling effect of education [Spence (1973) & Arrow (1973)]. Within the human capital theory it is also interesting to investigate whether it is cognitive development or the personality development during the schooling career that has mostly affected earnings [See Gintis (1971)]. An understanding of the link between education and earnings is important not only from the theoretical perspective, but it is also interesting for policy makers. For example, priorities in public investment in schools have to be re-designed accordingly if personality traits developed during the school career are more important than cognitive ability [Gintis (1971)].

2.3 Link from Education to Earnings

Economists have been divided by their beliefs regarding the association between education and earnings. One group uses the human capital argument to explain the association while the other uses the signalling theoretical arguments. The human capital theory argues that education enhances productivity and workers are paid according to their marginal productivity. More educated people are more productive and are therefore paid more than others are. This theory is based on an implicit assumption that workers’ productivity is observable to employers. In contrast to the human capital theory, the signalling-theorists argue that education does not increase, but instead reveals, the productivity. The labour market is characterised by asymmetric information. A potential employee knows his own ability but firms are unaware of that ability. Individuals possessing a greater number of abilities can acquire more education at a lower cost. In other words, the cost of education is a negative function of ability. Firms then use the reported level of education as a signal of productivity and decide on payments accordingly [See Groot and Hartog (1995)].

Layard and Psacharopoulos (1974): Layard and Psacharopoulos (1974) provide an empirical test regarding the screening (signalling) hypothesis. They argue that if the
screening hypothesis is the only truth, the rate of return to dropouts must be less than the rate of return to course completers, because dropouts do not show the required signal. Second, if screening is the only reason for a wage differential by level of education, tenure must have a negative correlation with the rate of return. This is because over time the employer receives more information about the productivity of the worker and there is therefore no need to use education to guess productivity. Third, they argue that employers can easily test productivity rather than rely on time spent in school. They found no empirical support for any of these hypotheses and therefore concluded in favour of the human capital theory.

Weiss (1983): This paper models the education decision using a mixed screening and human capital model. Weiss (1983) relaxes Spence’s assumption that education does not improve productivity. According to this model, education acts both as a screening device (it reveals the unobserved productivity), and it also increases the productivity. In this mixed model, Weiss (1983) shows that there is an under-investment in education (in contrast to the implication of the screening theory that there is an over-investment in education and the implication of human capital theory that the investment in education is optimal).

The paper begins with three critiques of the screening theory. First, this theory assumes that the more able invest more in education because it is cheaper for them. However, in reality it is well known that the main cost pertaining to full-time education is opportunity cost and the opportunity cost is greater for more able people; hence the screening theory is embroiled in a contradiction. Weiss (1983) shows that if a test certification is introduced, we can still argue that abler people receive more education because they can easily pass the test.

The second critique of the screening theory is that if education is just a sorting device and if only able people choose more education, firms can identify the more able people by only observing their (the individuals’) decision to attain more education. But it does not happen in reality. Weiss (1983) provides several reasons for this phenomenon. First, firms also care whether a person passes the given exam or not. Then, the question becomes why don’t firms give exams rather than wait for schools to administer them? It is efficient to have such tests given by schools rather than by firms. It would also be risky for individuals to be given a test by a firm. Moreover, firms will not be willing to administer such a test because it produces a positive externality for other firms. Therefore, this will not happen unless job applicants pay for the test.

The third critique is that if there are many types of workers, rather than two (good vs. bad), equilibrium does not exist.


2.4 Demand for Education

Gary S. Becker (1967): The human capital theory of demand for education was first developed by Becker (1967) and later reproduced by many economists [See inter-alia, Mincer (1970), Oosterbeek (1992), Hartog (1993) and Card (1994)]. This theory assumes that one chooses the optimal length of schooling to maximise the present value of the net lifetime wealth. This model is estimated in Chapters 3 and 6 of this book. The optimisation problem is as follows:

\[
\text{Max } V_s = - \int_{A_0}^{A_0+S+T} Ke^{-rt} dt + \int_{A_0+S}^{W_s}Ce^{-rt} dt
\]  

where \( V_s \) is the present value of net lifetime wealth. \( K \) is the direct cost of education, \( A_0 \) is the age at which the schooling decision is made, \( S \) is the optimal length of schooling, \( r \) is the discount rate, \( W_s \) is the wage rate for \( S \) years of schooling, \( e \) is the natural logarithm base, \( t \) is time, and \( T \) is the length of working life.

Now, assuming that \( T \) is large and \( K \) is independent of the level of education, we can write the following first-order condition

\[
r(K + W_s) = \frac{\partial W_s}{\partial S} \]

where the left-hand side gives the marginal investment cost of education and the right-hand side gives the marginal rate of return to education. A second order-condition guarantees that the marginal cost increases and the return to education decreases with length of schooling. The equilibrium length of schooling \( S \), and the equilibrium rate of return to education (this is also the discount rate at equilibrium) \( r^* \) are determined by the interaction between the marginal investment cost and the marginal rate of return to education. This theory explores the variation of the length of schooling over a cross-section of individuals in terms of varying cost and rate of return to education. People with low discount rates (and low costs) and those with higher rates of return will stay longer in school than the others do. Becker (1967) further assumes that the cost of education represents opportunity, while the rate of return represents ability. People with better opportunities enjoy education with a relatively lower investment cost and those with greater abilities will enjoy a higher rate of return to education. Becker (1967) thus identifies two sources of individual variations in schooling attainment: opportunities (costs) and abilities (returns). Therefore, the total variation of schooling is due to the variation in ability or the variation in opportunities and the covariance between the two. This theory implies that societies with equal opportunities -- but with individuals having different abilities and societies with unequal opportunities and equally able people -- may experience the same variation of schooling length. However, these two societies can be differentiated by the association between the length of schooling and the rate of return to

---

2 W.S. Woitinsky lecture (1967). This paper has also appeared in the third edition (1993) of Becker (1964) as an addendum to Chapter 3.
education at equilibrium. If opportunities are equal but abilities are not, the rate of return must be positively correlated with the length of schooling; the correlation must be negative otherwise. Becker (1967) and Mincer (1970) have depicted three special cases: equal opportunities, equal abilities and perfect correlation between abilities and opportunities. The first case presumes that everyone in the society has equal opportunities. They are different in ability. In this situation, there exists only one cost curve. The rate of return is different among individuals. In this model, equilibrium occurs along the supply curve. Becker (1967) calls it the elite model. In the second case, there are many cost curves but there is only one rate of return curve. This situation is called the egalitarian model. According to the third case, abilities and opportunities are either negatively or positively correlated. If the correlation is positive, more able people enjoy better opportunities as well. In this case, an outward shift of the return to education curve is associated with a downward shift of the cost curve by the same magnitude. Such a society is attributed with many equilibrium lengths of schooling for a given equilibrium rate of return. If the correlation is negative (abler people have fewer opportunities), an upward shift of the return to education is associated with an upward shift of the cost of education by an equal magnitude. The resulting equilibria are attributed with many rates of return for a given equilibrium length of schooling.

Another interesting issue is the implication of this model for social mobility. In general, the argument is as follows. If the family background positively affects the ability and the opportunity such that pupils with affluent family backgrounds are more capable and they have a greater number of opportunities than others do, an expansion of education will never enhance the social mobility. Pupils with affluent family backgrounds will receive more education because they are more able and they also have greater opportunities. They will also be the rich parents of the next generation since they obtain better-paid jobs (as a result of being more educated).

Thus far this model assumes that ability affects only the rate of return to education. Griliches (1977) assumes that ability also affects the intercept of the education earnings profile and therefore, more able people also have a higher opportunity cost as well as a higher rate of return. This leaves the ability effect on schooling undetermined.

Another schooling model developed in the spirit of Becker (1967) is Ben-Porath (1967). Ben-Porath (1967) takes into consideration the dynamic nature of the human capital formation to develop his model.

---

3 The implication of the human capital model on social mobility is briefly discussed in Mincer (1970). Later, there were several works in the economic literature addressing the issue of social mobility. See for example, Becker and Tomes (1986) *inter alia*. 
Ben-Porath (1967): Ben-Porath (1967) provides an extension of the human capital-demand for education model. The model presumes the existence of a human capital production function. It is assumed to be a function of two inputs: individuals own inputs and purchased inputs from the market. Individuals own input is the already accumulated human capital. Individuals use a fraction of the already accumulated human capital for the production of human capital. Assuming Cobb-Douglas production technology, Pen-Porath (1967) defines the following human capital production function:

\[ Q_t = \beta_0 s_t K_t^{\beta_1} D_t^{\beta_2} \]  

(2.4)

where \( Q_t \) is the human capital production, \( K_t \) is the already accumulated human capital, \( s_t \) is the fraction of \( K_t \) used for further production of human capital and \( D_t \) is the other inputs purchased from the market. \( \beta_0 \neq 0, \beta_1, \beta_2 > 0, \beta_1 + \beta_2 < 1, \) and \( 0 < s_t < 1 \) by definition.

For a given level of \( K_t, D_t \) and parameter values, human capital production, \( Q_t \) is a function of \( s_t \) alone. However, by definition, \( s_t \) has an upper bound (\( s_t = 1 \)) and thus so has \( Q_t \). Let us denote the upper bound of \( Q_t \) by \( \overline{Q}_t \).

Individuals minimise the cost of producing human capital subject to a given production function constraint. The cost of producing human capital is twofold: opportunity cost and the direct cost for purchased inputs. Using Ben-Porath’s notations we write \( \alpha_o \) for the rental rate of human capital and \( P_d \) for the price for a unit of purchased inputs. Ben-Porath (1967) also assumes that individuals are price-takers in both markets. Therefore, both \( \alpha_o \) and \( P_d \) are constants. Minimising the cost subject to the given production function, we can derive the cost function. The first derivative of the cost function with respect to \( Q \) yields the marginal cost curve. The marginal cost curve is given below.

\[ MC_t = \frac{\alpha_o}{\beta_0 \beta_1} \left( \frac{\beta_1 P_d}{\beta_2 \alpha_o} \right)^{\beta_2 / \beta_1} \left( \frac{Q_t}{\beta_0} \right)^{(1 / (\beta_1 + \beta_2)) - 1} \]  

(2.5)

Marginal cost is increasing with \( \alpha_o, P_d \) and with \( Q \). Marginal cost is zero when \( Q \) is zero. Now, remember that \( Q \) has an upper bound. This can be incorporated with equation (2.5) to show that, for a given time, the short-run marginal cost is increasing with \( Q \) until \( Q \) reaches its upper bound. After that, the marginal cost is perfectly elastic to \( Q \). In other words, the short-run marginal cost curve will be parallel to the vertical axis representing the marginal cost at the upper bound of \( Q, \overline{Q}_t \).

One further remark on equation (2.4) is in order. That is about the definition of \( K_t \), the human capital stock at year \( t \). If depreciation is not allowed, \( K_t \) is the accumulated human capital from the beginning to year \( t \). Thus, \( K_t = \int_0^t Q_v dv \). This means that \( K_t \) is growing over time such that \( K_t > K_{t+1} \) for any \( t \). Once again this implies that the upper bound of \( Q_t \),
(Q_t) is also increasing over time. The implication of this on the marginal cost curve is that the vertical segment of the short-run marginal cost curve will shift outward over time, indicating a higher cost to achieve the upper bound of the production function.

This is all about the cost side. Regarding the benefits side, Ben-Porath (1967) defines the "demand price" for human capital, P_t, as the present value of the rental rate of human capital, \( \alpha o \). Once again, individuals are price-takers and therefore P_t is also the marginal benefit curve. P_t is given in equation (2.6).

\[
P_t = \alpha_o \int_t^T e^{-rv} dv = \frac{\alpha_o}{r} [1 - e^{-r(T-t)}] \tag{2.6}
\]

where r is the discount rate and all other parameters are already defined. Equation (2.6) shows that the marginal benefit to investment in human capital is independent of Q and its negative first derivative to t indicates that the marginal benefits to investment in human capital decrease over time.

The intersection between equations (2.5) and (2.6) determines the equilibrium amount of Q_t. Remember the following properties of the marginal cost curve. The marginal cost curve starts from the origin. It increases with Q until the upper bound of Q is reached. At the upper bound, marginal cost is vertical in the short-run. This together with the horizontal marginal benefits curve suggests that one’s lifecycle can be divided into three phases according to the nature of the equilibrium of the human capital investment. In the younger years, the marginal cost curve reaches the vertical segment very soon. The marginal benefit is also very high. Therefore, it is likely that equilibrium occurs at the vertical segment of the marginal cost curve, indicating that s_t is one (full-time schooling). Over time the vertical segment of the cost curve shifts outward and the marginal benefit curve shifts downward; the equilibrium point gradually reaches the upward sloping segment of the cost curve. This characterises the second phase of the lifecycle. In the second phase s_t is less than one but greater than zero, indicating that only a fraction of accumulated human capital is re-invested in human capital production. During this period, people invest in part-time education and on-the-job-training, etc. At the end of the lifecycle or at the retirement age (when t=T), the marginal benefit will be zero and equilibrium occurs at the horizontal line. This characterises the third phase of the lifecycle.

According to Ben-Porath (1967), whether someone is in full-time education or not depends on whether the equilibrium occurs along the vertical segment of the marginal cost curve or not. Therefore, all the factors determining the position of the vertical segment of the marginal cost curve and the location of the marginal benefit curve will be in the reduced form of the optimal length of schooling equation.

---

5 Ben-Porath (1967) also introduces a depreciation rate. See equation (10) in Ben-Porath (1967). In this chapter we ignore the depreciation rate for the sake of simplicity.
In general, the human capital model of Becker that we have presented in the previous section, and the Ben-Porath (1967) model generate similar testable hypotheses regarding demand for education. However, some particular characteristics of Ben-Porath (1967), which are different from the Becker model deserve further remarks.

One important difference between the Ben-Porath (1967) and Becker (1967) models is the role of age. Becker (1967) does not have a direct role for the age as a determinant of the equilibrium length of schooling. However, Ben-Porath (1967) has a direct role for age. As we have already noted, age is one of the prime factors determining the location of the vertical segment of the short-run marginal cost curve and the location of the marginal benefits curve.

The role of ability and family wealth is also different in the two models. According to Becker (1967), ability affects the rate of return. More able people stay longer in full-time education because they have a higher rate of return. However, according to Ben-Porath (1967), the ability effect channels through cost [c.f. equation (2.3) & (2.4) $\beta_i$ is the ability parameter]. According to Becker’s model, unless we assume that ability is a function of family wealth, the effect of family wealth channels only through the cost function. Pupils from more affluent families have a lower discount rate. In fact, in the Ben-Porath model family wealth does not have any direct role. The only channel in which it may affect the schooling decision is through the effect on the interest rate, $r$ that appears in the marginal benefit curve [c.f. equation (2.6)].

Finally, according to Becker (1967), people invest in full-time education when they are young only because otherwise the period over which they can receive returns on their investment is shorter. However, according to Ben-Porath (1967), higher opportunity cost and lower marginal benefits for elderly people will also preclude them from investing in full-time education.

**Uncertainty and Investment in Education:** Levhari and Weiss (1974) is one of the pioneering works where uncertainty is introduced into the human capital model. This model assumes that individuals maximise the expected utility subject to the wealth constraint. The expected utility is a function of current and future consumption, $E[U(C_0, C_1)]$ and the wealth constraint is given as

$$C_1 = [A + (1-\lambda)Y_0 - C_0](1+r) + Y_1$$

(2.7)

where $C$ is consumption, subscripts 0 and 1 indicate present and future, respectively. $Y$ is earnings, $\lambda$ is the proportion of present time allocated for investment in human capital, and $r$ is the market rate of interest. $A$ is the initial (inherited) wealth.

This model assumes that education is a pure-investment good. Utility is generated only from present and future consumption. Furthermore, $Y_1$, future earnings are assumed to be a function of human capital accumulation, $\lambda$ and future unknown state, $\mu$. $\mu$ represents two types of uncertainties, *uncertain inputs* such as individual ability and school quality.
and the uncertain outputs, such as demand and supply conditions in the future labour market for a stipulated amount of human capital accumulation. This yields the future earnings function, \( Y_t = f(\lambda, \mu) \) with \( f_\lambda(\lambda, \mu) > 0, f_\mu(\lambda, \mu) > 0, f_{\lambda\lambda}(\lambda, \mu) < 0, f_{\mu\mu}(\lambda, \mu) < 0 \) and \( f_{\lambda\mu}(\lambda, \mu) > 0 \).

Levhari and Weiss (1974) have used a "more special form" for the future earnings function.

\[
Y_t = Y_0 + \lambda Y_0 Y(\lambda, \mu) \tag{2.8}
\]

where \( \gamma(\lambda, \mu) \) is the average rate of return for investment in human capital.

Given that the rate of return for investment in physical capital is certain and the rate of return for investment in human capital is uncertain (or, the latter is riskier than the former), Levhari and Weiss (1974) show that the former is less (greater) than the latter, in equilibrium, if risk of investment in schooling increases (decreases) with the length of schooling.

The Levhari and Weiss (1974) model has three important implications. One is on the association between family wealth and investment in human capital. In the previous models, with regard to certainty, we observed that family wealth does not have any role in the investment decision unless either an imperfect capital market is assumed or individual ability is assumed to be a positive function of family wealth. The Levhari and Weiss (1974) model predicts that even in the absence of capital market imperfections or wealth effects on ability, family wealth can have a positive effect on the investment decision. Given their assumptions of risk increasing with the level of education and of decreasing absolute risk-aversion, investment in risky ventures will become a normal good, and therefore more wealthy people tend to invest more in risky human capital.

The second implication of this model is that the effect of a changing market rate of interest on investment in human capital is unpredictable. An increase of the interest rate has two effects: a substitution effect and an income effect. The substitution effect is negative: the higher the market rate of interest the less will be invested in human capital. However, the sign of the income effect is conditional upon whether the respondent is a net-borrower or a net-saver in the current period. If the individual is a net-borrower, an increasing rate of interest will reduce the value of wealth and thereby further reduce the investment. This indicates that the interest rate effect on the investment in human capital of net-borrowers is negative. However, the income effect is positive for net-savers. This makes the interest rate effect unpredictable for net-savers.

The third implication of Levhari and Weiss (1974) is that the effect of changing risk on the investment decision is unpredictable.

The Levhari and Weiss model (1974) is limited in that they analyze the effects of only one source of uncertainties. Namely, the uncertainties associated with future earnings. The
analysis in Levhari and Weiss (1974) assumes that there is no uncertainty in the schooling career and the probability of being unemployed after the schooling career is zero.

However, in deriving his dropout model, Manski (1989) presumes that the schooling outcomes are unpredictable. He assumes that the decision of completing a stipulated programme depends on an endogenous factor (willingness to complete) and an exogenous factor (ability to complete). Manski (1989) assumes that the ability to complete a given programme is not known to the decision-maker. He realises his competence only after attending the programme. Manski (1989) then shows that one will drop out before the end of the programme once he realises that he cannot complete it. In this sense, dropout as well as completion is an outcome of the optimisation behaviour. A complete review of the Manski paper is given in section 2.7 of this chapter.

Groot and Oosterbeek (1992) develop an extended version of the standard human capital model to incorporate the effects of employment probability, unemployment benefits, expected earnings, and the dispersion of expected earnings on demand for schooling.

Altonji (1993) addresses the effect of uncertainty associated with educational outcomes on the demand for and the rate of return to education.

**Groot and Oosterbeek (1992): Effects of Employment Probability on Demand for Schooling:** This paper extends the human capital model by assuming that people consider the future employment prospects in their decision on the length of schooling. The model also assumes that future earnings are not known with certainty. There are two sources of uncertainty in this model: uncertain employment prospects and uncertain earnings prospect for a given job. The unemployment probability is assumed to be negatively correlated with length of schooling. It is further assumed that at the beginning of each period, there is a single wage offer. Individuals accept the offer if and only if the offered wage rate \( W(S) \) is greater than the unemployment benefits \( b \) and stay unemployed otherwise. The unemployment benefits are independent of the length of schooling. The variance of the wage offer distribution is positively correlated with the length of schooling, indicating that there are more opportunities for more educated people.

With some mathematical derivations, Groot and Oosterbeek (1992) show that the marginal cost of education consists of three elements: direct cost of education, foregone earnings and unemployment benefits. The marginal returns curve has four effects: the effect of the unemployment probability, the wage effect, the distribution effect, and the effect of length of schooling.

According to the authors, the effect of changing unemployment probability on the length of schooling is unpredictable. The effect of a change in unemployment probability on the cost curve is conditional upon whether the unemployment benefits are greater (less) than the expected earnings. If the unemployment benefits are greater (less) than the expected earnings, the effect is negative (positive). In the first case, the marginal cost curve shifts...
downward with an increase in the unemployment probability and the marginal cost curve shifts upward if the unemployment benefits are less than the expected earnings. The effect of changing unemployment probability on the marginal returns is negative, indicating that the marginal return curve shifts downward with increasing unemployment probability. This easily shows that the effects of changing unemployment probability on the length of schooling is unpredictable if the unemployment benefits are greater than the expected earnings and the effect is clearly negative if the unemployment benefits are less than the expected earnings.

Changing unemployment benefits have a positive effect on the marginal costs while its effect on the marginal returns is unpredictable. Therefore, the overall effect of changing unemployment benefits on the length of schooling is also unpredictable. The effects of changing earnings are also unpredictable because it changes both the marginal cost and the marginal return towards the same direction.

Finally, Groot and Oosterbeek (1992) show that the effects of changing variation of wage offer distribution on the length of schooling is negative. It changes the marginal cost and the marginal return towards opposite directions.

**Extension of Becker Model to encounter the utility effect of Education:** Lazear (1977) provides a theoretical and econometric framework to examine whether education is a consumption good or an investment good. The issue of whether education is an investment or a consumption good goes far beyond the beginning of the subject of education economics itself. Lazear (1977) argues that if education is an investment good, one would choose a length of education to maximise the lifetime wealth (which he calls the opportunity set). According to Lazear's definition, the opportunity set consists of two elements: goods and services consumable (including leisure), \( X_e \) and education, \( S_c \). In other words, \( X_e \) is the non-human wealth and \( S_c \) is the human wealth possessed by individuals. The opportunity set at birth is \( (X_e, 0) \), where \( X_e \) is the possession of non-human wealth at birth. Lazear (1977) assumes that an individual invests part of his non-human wealth to produce education. The technology of the education production function, \( S = f(X^p) \) is assumed to be monotonic and continuously differentiable. \( X^p \) is the part of non-human capital invested to produce education. Accumulating human capital in terms of length of education changes the opportunity set to \( (X_e, S_c) \), where \( X_e, S_c \) are already defined. Further, Lazear assumes that the individual can sell his accumulated education, \( S \) for a given unit price of education, \( P_S \). When \( X^p \) units of initial wealth are invested in education, the initial wealth reduces to \( X_e - X^p \). However, he can sell his accumulated education at a given unit price \( P_S \). The addition to the non-human wealth is designated by \( f(X^p)P_S \). Then, \( X_e \) can be defined as \( X_e = X_e + [f(X^p)]P_S - X^p \).

If the individual maximises wealth (if education is an investment), he will choose the length of education \( S \) such that the education production function, \( f(X^p) \) will be maximised. Because \( X_e = X_e + [f(X^p)]P_S - X^p \) this will also maximise the opportunity set.
However, if education is a consumption good, it must be an argument in the utility function. Let \( U = U(X^c, S^c) \) be the well-behaved utility function. The rational individual then maximises this utility function subject to the opportunity constraint. Lazear (1977) shows that if the individual is maximising utility rather than wealth, the optimum will never occur at the wealth maximising level of education unless there is a kink in the education production function at the wealth maximising level. However, this possibility is ruled out by the assumption of a continuously differentiable education production function. Therefore, the above theoretical framework yields two different optimal lengths of education, depending on whether the individual is a wealth maximiser or a utility maximiser.

By using specific mathematical formulations for the utility and education production functions, Lazear derives two demand functions for wealth maximisers and for utility maximisers. Lazear (1977) specifies \( U = X^c e^S \) to be the utility function and \( S = AX^p \) to be the education production function. This yields

\[
S^* = (AYPS)^{(1-\gamma)}
\]

for the wealth maximising education demand function and

\[
S = (1/P_S)(X^p - X^e) + (1/\theta)(1/P_S)(X^p/\gamma S) - (1/\theta)
\]

for the utility maximising education demand function.

According to the utility function, it is clear that if education is a consumption good, \( \theta \) must be positive. Lazear (1977) further argues that the pre-condition for the education to be an investment good is that the price of education, \( P_S \) must be positive.

Lazear (1977) estimates the education production function and the utility maximising education demand function (equation 2.10). In order to estimate these functions, one has to have observations on \( S \), the length of schooling, \( P_S \), the unit price of education, \( X^c \), possession of non-human capital at birth and \( X^p \), the part invested in education production. Among these variables, only \( S \) can be directly obtained from a sample survey. All others are conceptual variables. Lazear (1977) assumes that \( 1/P_S \), the inverse of the unit price of education, is a linear function of some selected exogenous individual characteristics. \( X^c \) is the present value of lifetime income if the individual does not attend school (i.e. \( S = 0 \)). He derives this from an earnings function and by assuming that the market rate of interest is 10 percent and the length of a working life is 65 years. \( X^p \) is also derived from the same earnings function. He further includes several exogenous variables in the education production function. These yield the following two equations.

\[
S = Z\beta(X^p - X^e) + Z\beta(X^p/\theta S) - 1/\theta
\]

is the utility maximising education demand function and

\[
\ln S = \gamma \ln A + \gamma \ln X^p + \ln X^e
\]

6 The full cost of education is the sum of opportunity cost and direct cost. Lazear (1977) uses an imputation method to calculate the total cost of education. For details see Lazear (1977).
is the education production function. $Z$ is a row vector of exogenous variables, $\beta$ is a column vector of coefficients. $X_p$ is the full cost of education. $\ln X$ is a vector of exogenous variables determining $1/P_s$. Calculation of this variable is given in Lazear (1977) and we will not elaborate upon it here.

Once the data is available, we can use the OLS to estimate equation (2.12). However, the estimation of equation (2.11) encounters several problems. First, this relationship is not intrinsically linear. Without knowing the value of $\gamma$, this equation is not estimable. However, equation (2.12) gives an unbiased estimate of $\gamma$. Lazear (1977) uses this to estimate equation (2.11). The second problem associated with the estimation of equation (2.11) is that $S$, the length of schooling, appears in both sides of the equation. This gives a non-zero correlation between the error term in equation (2.11) and the explanatory variables of the model. Lazear uses the two-stage estimation method (2SLS) with some instrumental variables to overcome this problem. He regressed all the explanatory variables on a set of instrumental variables and used the predictions of these regressions as the explanatory variables in equation (2.11). The third problem encountered with the estimation of equation (2.11) is that this equation over-identifies all the regression parameters. To overcome this problem, Lazear assumes that the intercept of the 2SLS estimate of equation (2.11) yields an unbiased estimate of $\theta$. The application of the same 2SLS procedure to the following model gives an unbiased estimate for $\beta$

$$S + 1/\hat{\theta} = \hat{Z}[X_pX_e] + (X_p/\hat{\gamma}S)$$

(2.13)

where $\hat{\theta}$ is an estimate of $\theta$ using the 2SLS procedure described above. Lazear found that $\theta$ is negative, thus indicating that education is bad from a consumption point of view. It gives dis-utility to the consumer. The estimated $\beta$'s were used to derive the unit price of education, $P_s$, which was found to be positive. This alone, according to Lazear, proves that education is an investment good. By using equation (2.9) and estimates of $\gamma$ and $P_s$, Lazear (1977) calculated the wealth maximising level of schooling, $S^*$ and found that the realised length of schooling, $S$ is less than the wealth maximising length, $S^*$ which indicates that people stop education before the wealth maximising length, because "consumption" of education is a "distasteful" activity.

Signalling theories on Demand for Education: The signalling theories discussed above also include a perspective on demand for schooling. Spence (1973) described demand for education in terms of a signal for unobserved ability. Spence's (1973) theory is restated in Gibbons (1992). Gibbons (1992) presented the same model as an extensive form game. Spence (1973) assumes that the productivity of a worker is unknown to the employers and

---

7 Kodde (1985) applied a different approach to test the same hypothesis. According to human capital theory, direct cost and the opportunity cost of education must have the same effect on education choice. However, Kodde's estimates show a difference. He asserts that the opportunity cost effect is greater than the direct cost effect. He attributed this difference to a wealth effect associated with changing opportunity costs. He concludes that education is a normal good.

8 An extensive form representation of a game means that the game is sequential. Player one acts first and player two reacts in his optimal way.
the cost of education is negatively related with ability such that more able people learn relatively cheaply as compared to less able counterparts. Furthermore, the cost of education for a given level of ability is an increasing function of the level of schooling. In the first move of the game, workers choose the optimal length of education depending on their own abilities. In the second stage profit maximising firms offer wages simultaneously and the worker accepts the one offering the highest wage rate. Payoff to the worker is $w - c(\eta, S)$, where $w$ is the wage offer, $c(\eta, S)$ is the cost of education, $\eta$ is the ability index, which assumes $H$ for high ability and $L$ for low ability with known probability distribution, and $S$ is the level of education. The payoff to the firm offering an acceptable wage is $y(\eta, S) - w$, where $y(\eta, S)$ is the production function. Payoff to the other firms offering a lower wage rate is zero.

In the presence of information asymmetry, this model may reach three different types of equilibrium, *pooling*, *separating* and *hybrid* equilibrium, depending on the payoff to the worker with different strategies. In the pooling equilibrium, both high and less able workers signal the same level of education such that education will not give a proper signal regarding the ability. Pooling equilibrium is the outcome if $w(H) - c[L, S(H)] > w(L) - c[L, S(L)]$. That is, the payoff to the less able workers for signalling the level of education of more able workers is greater than the payoff to the same worker for reporting the education level of less able workers. If the opposite is true, a separating equilibrium will occur, where the less able workers report lower educational attainment as compared to the education level of the more able workers. In the hybrid equilibrium, one group fixes its own signal and the other will randomly allocate themselves between the two education levels.

However, Spence (1973) shows two conditions guaranteeing the existence of a separating equilibrium. The two conditions are: i.) cost of schooling is a negative function of ability and ii.) there are sufficient numbers of signals within the appropriate cost range [see Spence, *op.cit*, p. 367].

On the variation of educational attainment, the signalling theory hypothesises that ability variations are the principal determinants of low and high educational attainment. Under the given set of assumptions those with high ability report a longer duration of education.

**Family Background and Investment in Education:** Thus far the reviewed literature on demand for schooling suggests that the optimal length of schooling varies over individuals with different abilities and different opportunities. Another issue that is very much related with the individual variation of the length of schooling is the family background effect. In general, the theories we have reviewed so far imply that the variables measuring the family background are one way or another related to the abilities and (or) the opportunities, and thereby the optimal length of schooling varies with the variables which measure family

---

9 In fact, this is a rather questionable presumption. This theory mainly considers the psychic cost of education. Opportunity cost, the principal factor in human capital theory, is apparently ignored.
background. For example, parents' education, their income, wealth and occupational status influence the ability and the educational opportunities of their offspring. Children of well-to-do parents are more capable than their counterparts. This immediately suggests that the children from well-to-do families gain more benefits (high marginal rate of return) from a given length of schooling and the cost of schooling is lower. Consequently, the optimal length of schooling of the children of the relatively less affluent families is lower than the optimal length of schooling of their wealthy counterparts. Ben-Porath (1967) reaches the same conclusion. How we introduce family background into the respective models determines the implications.

Under the signalling framework family background is an index (an observable, unalterable attribute Spence, *op.cit.*, p. 357). Spence (1973) explains the effects of indices on the equilibrium assuming that ability is randomly distributed over different values of an index. Under this circumstance the index affects the equilibrium if different values of an index represent different opportunity sets. For example, if the gender is the index (this is also a background variable), and if male and female have different opportunity sets, each will find a different equilibrium. In this case, *ceteris-paribus* the group with a relatively larger opportunity set will report a higher length of schooling.

However, the literature related with the family background effect in general assumes that family background has a positive effect on ability. Pupils with a better family background are more able (healthy, wealthy and wise), and given that the two conditions for separating equilibrium exist, this implies that family background has a positive effect on schooling. Introduction of a wealth effect will reinforce this, unless we assume that wealthy people are less able.10

Becker and Tomes (1986) and Becker (1989) have developed an economic theory analysing the family effect on investment in human capital formation11. Various empirical investigations of the issue are available. (For example, Bowles (1972), Conlisk (1974, 1977 and 1984), Menchik (1979), Zimmerman (1992) and Solon (1992)12. Becker and Tomes (1986) is based on an anthropological study of the relation of parents' height and the height of their children. [See Galton (1886)]. Galton's study indicates that the height of an individual can be expressed as the weighted-sum of the height of parents and the average height of the population. This suggests that when parents are taller than average, their children will also be taller than average. However, these children are shorter than their parents. If the parents are shorter than the average height, their children will also be

---

10 If wealthy people are less able they can compensate for the higher cost of schooling by using family funds.
11 This first appeared as a Discussion Paper in 1984 and is also available in the third edition (1993) of Becker (1967).
12 Most of these empirical investigations claim that the regression of children's income on the income of parents with OLS is downward biased due to the omitted variables and measurement errors. They have applied many different methods to correct the results for those biases and conclude that the parents' income is highly related with the income of their children, thus indicating that the intergenerational mobility of the society towards the average is slower than what is predicted when ignoring these biases.
shorter than average but taller than their parents. This finding leads Galton to conclude that the height of people converges towards the average height of the community over several generations (*regress towards the mediocrity*).

The Becker and Tomes (1986) model is precisely elaborated in Goldberger (1989) with some critiques. [Also see Polachek and Siebert (1993), Chap. 3]. According to Goldberger’s exposition, parents’ income \( (Y_{t+1}) \) is the sum of their own consumption \( (C_{t+1}) \) and the investment in their offspring \( (H) \). Thus, \( Y_{t+1} = C_{t+1} + H \). Income of the offspring \( (Y_t) \) is \( Y_t = mH + L_t \) where \( m = 1+r \). \( r \) is the market rate of interest of generation \( t \), \( L \) is the luck of the child. We can incorporate child’s ability, the endowments of parental genes, and labour market luck into this component. Becker and Tomes (1986), assuming a Cobb-Douglas utility function of altruistic parents (parents’ utility is a function of children’s income and parents’ consumption), derive the conditions for optimal allocation of parents’ lifetime earnings over their own consumption and investment in their children.

The utility function of the altruistic parents is written as follows:

\[
U = \alpha \ln Y_t + (1 - \alpha) \ln C_{t+1} \tag{2.14}
\]

where all the variables are already defined. By maximising equation (2.14) subject to parents’ and children’s income constraints, we can derive the rules for the optimal allocation of parents’ lifetime earnings over their own consumption and investment in human capital formation of their children. These conditions are given below.

\[
H = \alpha Y_{t+1} - (1 - \alpha) L / m \tag{2.15}
\]

\[
C = (1 - \alpha) Y_{t+1} + (1 - \alpha) L / m \tag{2.16}
\]

Equation (2.15) and the definition of children’s income given in the text show that a unitary increase in parents’ income increase parents’ investment by \( \alpha \) units (which is essentially a positive fraction) and in turn, children’s income rises by \( \alpha m \) units. A unitary change in the luck of children will lower the investment by \( (1 - \alpha) / m \) units, and in turn it changes children’s income by \( \alpha \) units. This yields the following equilibrium for children’s income:

\[
Y_t = b Y_{t+1} + \alpha L \tag{2.17}
\]

where \( b = \alpha m \). In equation (2.17) above, \( L \) is decomposed into two parts as \( e \), the endowments and \( u \), the market luck. The endowment represents the human capital that a child receives from their parents without any effort, for example in genes. Substituting this into equation (2.17), we can re-write the same equation as follows.

\[
Y_t = b Y_{t+1} + \alpha e_t + \alpha u_t \tag{2.17*}
\]

By following Galton (1886), Becker and Tomes (1986) define that the parents’ genes and other inherited attributes transfer to their descendants in the same way as Galton’s height transformation. Thus,

\[
e_t = (1 - c) \bar{e} + c e_{t+1} + v_t \tag{2.18}
\]

where \( \bar{e} \) is the average endowments of the community under consideration, \( c \) is the *inheritability parameter* [Becker and Tomes *op.cit*, p. S5] and \( v \) is an error term.
These conditions suggest that increasing parents’ income raises both parents’ own consumption as well as their investment in children. Increases in the factors assembled into luck will have a positive effect on children’s earnings. However, both the parents’ income and luck will increase children’s earnings to a lesser extent than the increase in parents’ income and (or) children’s luck. Becker and Tomes (1986) call it “offsetting effects”. For example, ability of children will have a negative effect on the parents’ investment decision. Parents, who know the abilities of their children, tend to reduce investment in more capable children and increase the investment in less capable children. They further draw policy implications from this model that public education policies and other subsidy policies will have a very small effect on demand for schooling because parents withdraw a certain percentage of their private investment in human capital as the government increases subsidies. Goldberger (1989) criticises this policy statement. His argument is that the above “offsetting” effect can occur if and only if the rate of return to parental investment (H) does not increase with luck (L).

The Becker and Tomes model further assumes that the transfers of assets from parents to children can happen in two alternative forms: as an investment in human capital formation and in terms of bequeathal. Economically affluent parents transfer more than others. They invest more in human capital formation in their children and they also bequeath more than others. The composition of transfers changes with parents’ income. Poor parents transfer more in terms of investment in human capital formation. Economically affluent parents transfer more in terms of bequeathal. This is driven by the assumptions within the theory. The theory assumes that investment in human capital is subject to the law of diminishing returns. The marginal rate of return to human capital therefore decreases with the amount invested, but an individual cannot affect the market rate of return for non-human assets. Economically affluent parents invest in human capital formation more than the others. Therefore, it is more likely that for economically affluent families, the marginal rate of return for human capital is equal to that for non-human capital. For them, any further investment in human capital is worthless. Consequently, an increase in their income will increase bequeathal only, not the investment in human capital.

### 2.5 Education Production Functions

A systematic investigation of the relationship between school inputs and student performance is condensed in the education production function. This idea became very popular among researchers after the controversial Coleman report of 1966. The Coleman report identified that the school inputs have a very marginal influence on student achievement. According to this report, the most important factor is the family background and the peer group effect. There was a plethora of studies on the association between school inputs and student performance after the Coleman report. Hanushek (1986) has reviewed the major issues in this area. Although a huge amount of work has been conducted since the time of this review, it seems that the issues raised by Hanushek are still valid.
Hanushek (1986) pointed out three drawbacks in the existing literature on the education production process. First, is the absence of a proper theoretical framework to justify the presumed production process. Although this is true in general, Polachek et al. (1978) and Levin (1980) have developed some reasonable theoretical structures. We will discuss these two models in detail. The second and third drawbacks, according to Hanushek (1986), are the difficulties associated with identifying inputs, and output of the education production process.

A mapping rule of one or several inputs into one measure of output is called a production function. Such a mapping must satisfy several mathematical properties. It must be a single valued, continuous, and well-defined function over the input set, yielding non-negative output. It has continuous first and second-order partial derivatives. The first derivative is positive and the second derivative is negative by assumption. Out of the 147 estimates of education production functions that Hanushek (1986) has reviewed, more than half of the estimates have revealed an insignificant association between school inputs and output. More surprisingly, there are many production relationships with negative signs that are even statistically significant. It is clear that the standard production theory will provide very little guidance in the interpretation of these findings. In fact, according to Hanushek (1986), for many of these estimated relations the choice of specific mathematical functional forms to represent the production process is not justified. One such attempt is Levin (1980).

Levin (1980) has used John Carroll’s (1963) theory of learning as the baseline theoretical foundation. According to Levin (1980), Carroll’s theory identifies two major factors determining one’s learning; internal factors and external factors. Internal factors are the student’s characteristics, which include aptitude, ability to acquire new knowledge and the time he is willing to spend on learning. External factors are the factors beyond the student’s control. Quality of teaching and the time devoted to teaching a given lesson (opportunities) are the external factors emphasised by the learning theory of Carroll. Quality of instruction is assumed to depend on teachers’ capacity and effort level.

Levin (1980) emphasises the importance of teachers’ capacity, effort and time factors as separate determinants of the education production process. Levin argues that teachers with the same academic capacity may work differently in different school systems. Teachers working at schools with more incentives will, in general, work harder than their counterparts with equal qualifications.

Levin uses this theoretical structure to disqualify most of the estimated education production functions. He argues that, in general, measures of teacher input variables are limited to educational qualifications and teacher experience. These represent only the

---

13 See Harbison and Hanushek (1992) for more regression results for selected developing countries as well.
capacity dimension of teachers. Levin’s theory implies the following education production function

\[ Q = A(bX)^a \]  

(2.19)

where \( Q \) is the output, \( X \) is the teachers’ capacity measure and \( b \) is a fraction measuring the effort level. If one works at his full capacity, \( b \) is one and it is a positive fraction otherwise. Equation (2.19) clearly implies that teachers with the same capacity may report different output levels depending on the size of \( b \).

Levin (1980) has fully described the external factors from Carroll’s theory of learning. However, he pays less attention to the internal factors. Summers and Wolfe (1977) and Polachek et al. (1978) are two of the few models that articulate the complete Carroll model into an economic theoretical framework.

Summers and Wolfe (1977) estimate the following function:

\[ \Delta A_t = A(GSES,TQ,SQ,PG) \]  

(2.20)

where \( A \) is a measure of achievements, change in \( A \) is the value-added. GSES is the Genetic and Socio-economic characteristics, TQ is teachers’ quality, SQ is quality of school inputs other than teachers, and PG is peer-group characteristics.

Summers and Wolfe (1977) depart from the traditional estimates of education production functions on two grounds. First, they use a measure of value added as the output variable in the education production function. According to their specification, the output of a school is not comprised of the academic achievements themselves, but rather the change of the achievements under a given arrangement of school inputs. Second, they argue that the effect of a given input also depends on the characteristics of pupils. Mere introduction of individual attributes into the right-hand side will not be sufficient. Different pupils receive different benefits from a given amount of inputs even after controlling for their socio-economic characteristics. Therefore, Summers and Wolfe (1977) introduce many interaction terms into their specification. Interaction terms between school quality variables and individual characteristics reveal that different individuals have different effects from the same input combination. For example, their estimates show that more able students receive more benefits from experienced teachers. The effect of teacher experience on weak children is negative.

Polachek et al. (1978) derive and estimate an education production function for a selected sample of undergraduate students from the University of North Carolina at Chapel Hill. This model assumes a CES production technology and defines production as a function of school and pupils’ inputs. The output of this production function is the proportion of correct responses for 50 questions on principles of economics. Inputs are threefold: number of hours a student participates in lectures and seminars, number of hours a student spends on reading and preparation, and the individual’s scores on the quantitative portion of the SAT exam. Gender and family background variables are also used for controlling purposes.
According to their findings, initial ability as measured by the SAT scores is the most effective factor in determining one’s success in education. Individual study is more effective than class attendance. Individual study along with class attendance will compensate for lower initial ability.

Mora (1997) estimates average daily attendance and school dropout decisions within a utility maximisation framework. The model is applied for the US-ethnic minority groups. Let $U_{i,j}^A$ be the utility from attending school $j$ by individual $i$ and, $U_{i,j}^N$ be the utility from not attending. A child will attend school $j$ if $U_{i,j}^A > U_{i,j}^N$. Mora assumes that $U_{i,j}^A - U_{i,j}^N = Q_j \alpha_1 + Z_{i,j} \alpha_2 + z_j \alpha_3 - C_{i,j} \alpha_4 + \epsilon_i + \epsilon_j$, where $Q$ is a vector of school characteristics, $Z$ is a vector of the characteristics of individual $i$ attending school $j$, $z$ is a vector of average pupil characteristics of the school (peer-group), $C$ is opportunity costs, and there are also individual and school-specific error terms. When we take the school average of the above expression this gives:

$$n^{-1} \sum_{i=1}^{n} (U_{i,j}^A - U_{i,j}^N) = n^{-1} \sum_{i=1}^{n} (Q_j \alpha_1 + Z_{i,j} \alpha_2 + z_j \alpha_3 - C_{i,j} \alpha_4 + \epsilon_i + \epsilon_j), \quad \text{or} \quad \text{ADA}_j = Q_j \alpha_1 + z_j (\alpha_2 + \alpha_3) - c_j \alpha_4 + \epsilon_j,$$

where ADA is the average daily attendance and $c$ is the average opportunity costs (i.e. the price) of attending $j$.

The author also models the dropout decision with a similar set of explanatory variables. However, since the individual level data is not available for opportunity costs, it is proxied by using the residuals from the average equation. Empirical findings support the argument that school inputs matter. An educational voucher system is proposed as a remedy for the high dropout rate and the irregular attendance of minority groups.

However, none of these models has considered the institutional set up and incentives which go into the effort variable of Levin (1980).

As noted above, the second criticism of Hanushek (1986) regards the choice of input variables. Hanushek (1986), Harbison and Hanushek (1992) and Hanushek and Jamison (1992) have listed all the major works in the field of education production functions. According to their reviews, the most popular input variables are $a.$ Teacher/Pupil Ratio, $b.$ Teacher Education, $c.$ Teacher Experience, $d.$ Teacher salary, $e.$ Expenditure per pupil and $f.$ Selected administrative and physical facility variables. All these variables are presumed to have a positive effect on the school output. However, according to their reviews in econometric research, these variables are statistically insignificant and there are many cases where the regression coefficients are negative. This is rather similar for both developed and developing countries. In fact the issue that school inputs are not significant or less significant is one of the earliest problems raised by the pioneers of the education production function estimates [Coleman et al. 1966]. Loeb and Bound (1996), through an extensive review of the literature, have observed that, in general, school inputs are
significant when the labour market variables are used as education output, whereas they are insignificant when the academic outcomes are used as the education output. Loeb and Bound (1996) have several explanations for that. The most viable explanation is related to the differences in the type of data used in two cases. When the academic performance is used as the output variable, the analysis is done at the micro level and both the output and the input variables are generated contemporaneously. When the labour market output is used, school input variables are aggregated [In, for example, average quality of education in the district of birth. See Card and Krueger (1992a) inter-alia].

While the second problem raised by Hanushek (1986) has not yet been resolved, the third problem is also enormously important. This concerns the choice of school output. What do schools produce? Hanushek (1986) correctly pointed out that “Education is a service that transforms fixed quantities of inputs (that is individuals) into individuals with different qualities”. This alone makes the estimates of education production functions different from the standard production theory. The standard production theory postulates the process through which fixed amounts of inputs convert into amounts of output, whereas the education system changes the quality of fixed amount of student inputs using various school inputs, such as teachers and other material inputs. This complicates the education production process.

The most popular measure of academic performance is the standardised test scores [Loeb and Bound (1996) inter-alia], while student attitudes, school attendance [Mora (1997) inter-alia] and college continuation rates have been used in different studies as well. Use of labour market performance such as wages also has a long history. Behrman and Birdsall (1983), Card and Krueger (1992a and 1992b) are only examples from a long list. Heckman et al. (1996) have an excellent review of the education production functions with wages as the output variable.

As previously noted, review of the literature relative to those two types of educational outputs is already available [See, Hanushek (1986) and Heckman et al. (1996)]. However, the third variety of education production functions, joint-product production processes, are not fully elaborated elsewhere. We therefore skip the first two and expand on the third one.

Rao (1969) has provided a clear exposition on joint-product production processes. There are two different situations with more than one output produced by the same firm at the same time: multiple-product and joint-product. In the case of multiple-output, the same type of inputs are used to produce several outputs. The production processes are different from each other. Use of some input in one production process implies that it cannot be used to produce the other output. This is called complete input exhaustion. An example of a multiple-product production process is wheat and livestock. In the case of a joint-product, the same amount of input is used to produce more than one output produced by the same production process. In this case use of a certain amount of input to produce one output does not mean that this amount is not available to produce the other output. This is called zero input exhaustion. A typical example of such a system is the production of wool.
and mutton [Vinod (1968)], or sugar and molasses [Rao (1969)]. It is more appropriate to write this production function in its implicit form \( F(y_1, y_2, x_1, x_i) = 0 \), or equivalently\(^{14}\), \( f(y_1, y_2) = g(x_1, x_i) \).

Chizmar and Zak (1983 and 1984) and Chizmar and McCarney (1984) furnish some of the earliest applications of the joint-product production systems to the education production function. Chizmar and Zak (1983) estimate a joint-production function for the undergraduate economics teaching at the Illinois State University. This model assumes that undergraduate teaching produces two outputs. One is the knowledge of the principles of economics \((Y_1)\) (cognitive ability) and the other is the attitude toward economics \((Y_2)\). \(Y_1\) is measured using exam scores of principles of economics at the end of the course and a 14 item Likert scale instrument was used to generate attitude scores. This attitude test was performed twice, once at the beginning of the course and then at the end. The end of the course test score was one output in the joint-production process. Selected variables on students' background, their prior achievement, ability, and effort were the inputs of the presumed joint-production process. Chizmar and Zak (1983) assume the Cobb-Douglas specification of the production technology.

\[
\alpha_1 Y_1 + \alpha_2 Y_2 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon \tag{2.21}
\]

where \(Y_1\) and \(Y_2\) are already defined. \(X_1\) to \(X_k\) are the inputs, all are measured in natural logarithms and \(\varepsilon\) is a random error term. Hotelling's (1936) canonical correlation technique was used to estimate this model.

In another application, Gyimah-Brempong and Gyapong (1991) estimate a joint-product education production function with Mathematics and English scores as the two joint-products. They introduce socio-economic characteristics of the community as a set of inputs. They argue that socio-economic characteristics of the community can have two effects on the school output: an indirect effect via its effect on the school characteristics and a direct effect. Schools serving educated and affluent communities are generally enriched with more facilities and they are more efficient. Children from such communities are "healthy, wealthy and wise". In order to test whether there is a direct effect of socio-economic background of the community variables, Gyimah-Brempong and Gyapong (1991) estimate two versions of their model: one with socio-economic characteristics and the other without. They found that socio-economic variables have a direct effect on the school output. They state that the most plausible socio-economic variable in this context is the measure of educational attainments of the community. They further conclude that effects of school inputs are not easy to capture because of the high collinearity between school inputs and socio-economic variables.

\(^{14}\) This is true if and only if the partial derivative of the implicit function with respect to \(y\) is non-zero for the given values of \(x\) and \(y\) that satisfy the implicit function. See Chiang (1984), *implicit function theorem*, pp. 204-206.
2.6 Demand for Schools

The human capital theory [Becker (1967)] implicitly assumes that schools are homogenous. This, in turn, suggests that the choice of school is irrelevant. However, as we already learned in the previous section, the entire literature on the education production function presumes that schools are different from each other in productivity. This presumption has an immediate implication on the individual behaviour on school choice. If schools are different from each other in productivity, there is no reason for us to believe that a rational individual chooses a school at random. He chooses the best school subject to his budget and other constraints.

This motivates the discussion in this section. The literature on the demand for schools is rather undeveloped. According to the available literature, the earliest works on the issue were published in the 1980s [See Fuller et al. (1980)]. And from the 1980s to today, only a few works have been published. Economic theories and statistical techniques are very specific. Our review of the literature on the issue shows that two types of statistical techniques were applied for estimation purposes. Venti and Wise (1982) have estimated a system of equations for application, choice of college quality, and admission decisions simultaneously. Gertler and Glewwe (1990) have applied the Nested-Multinomial (NMNL) of McFadden (1981). The conceptual framework for Gertler and Glewwe (1990) is provided by Small and Rosen (1981).

Venti and Wise (1982) argue that estimation of demand for school and admission decisions without considering the application decision is biased because of the selectivity problem. It is more accurate, they say, to first model the application decision. If one decides to apply, then the problem is the choice of school to which one applies. After application to a college, the college authority will decide whether to accept the application and offer a placement or to reject. Venti and Wise (1982) assume that the utility associated with not applying for a college \( U_0 \) is a linear function of socio-economic characteristics of the individual. Thus,

\[
U_0 = Z_1 b_0 + e_0 \tag{2.22}
\]

where \( Z_1 \) is a vector of socio-economic characteristics, \( b_0 \) is a vector of estimable coefficients and \( e_0 \) is a random error term. Venti and Wise also assume that the quality of the most preferred school is a function of socio-economic variables.

\[
Q = Z_2 \beta_2 + e_2 \tag{2.23}
\]

where \( Q \) is the quality of the most preferred school, \( \beta_2 \) is a vector of estimable coefficients and \( e_2 \) is a random error term. Venti and Wise (1982) assume that the average SAT scores of a given school represent the quality of that school.

Then, the utility from continuing education \( U_1 \) is

\[
U_1 = Z_1 b_1 + Q c_1 + e_1 \tag{2.24}
\]
Equation (2.22) and (2.24) together show that one will apply for the most desirable college if and only if $U_i \geq U_0$.

Then the probability of application is equal to the probability of $U_i \geq U_0$, $Pr(U_i \geq U_0)$. By substituting equation (2.23) into (2.24), we can write the probability of application

$$Pr(U_i \geq U_0) = Pr[Z_i(b_1 + c_1b_2 - b_0) + (c_1e_2 + c_1 - e_0) \geq 0]$$

(2.25)

Once the application is submitted, the authority of the school will consider whether the applicant is qualified. An applicant will be admitted only if his academic potential ($A_P$) is not less than the required minimum standard ($L$) by the college. Venti and Wise (1982) assume that the minimum standard of a college is a positive function of the quality of the college. Let $L_j = \alpha_1Q_j + u_j$, where subscript $j$ refers to the $j^{th}$ college, and $u_j$ is a random term. Applicant potential is assumed to be a function of his high-school rank ($R$), SAT scores ($T$), socio-economic characteristics of the applicant ($Z_2$)$^{15}$, and a random error term. Let

$$A_{Pij} = \alpha_1T_i + \alpha_2R_i + Z_{2i}\delta + e_{ij}$$

(2.26)

be a linear approximation determining the college's assessment of applicants' academic potential. It is now clear that an applicant will be admitted to a college only if the academic potential is not less than the required minimum standard by the college. Equation (2.27) gives the admission probability.

$$Pr(\text{Admit}_{ij}) = Pr(A_{Pij} \geq L_j) = Pr[(\alpha_1T_i + \alpha_2R_i + Z_{2i}\delta) - \alpha_1Q_j + (e_{ij} - u_j) \geq 0]$$

(2.27)

Venti and Wise (1982) estimate equations (2.23), (2.25) and (2.27). Estimation of this model requires several types of data. Venti and Wise (1982) use selected variables to represent academic and non-academic achievements of individuals, quality of the education at high school, socio-economic background of individuals, and local labour market information. The academic achievements of respondents are measured by the Scholastic Aptitude Test (SAT) scores of the responding individuals. For those who had not taken the SAT, either the American College Test (ACT), or the Educational Testing Service (ETS) scores were used to produce their "SAT equivalent" scores. Venti and Wise (1982) also use the high school class rank of individuals as a measure of academic achievement. Non-academic achievements are measured by the leadership quality of respondents (whether the respondent is a leader of the student community at high school or not) and the athletic achievements of the respondents. Quality of the high school attended is approximated by the percentage of students from a given high school enrolled in two- or four-year college programs. Quality of the college, the dependent variable of equation (2.18), is the average SAT score of the students entering the college. Family income, parents' education, race, and residence of the respondent serve as socio-economic variables in this model. Local wage rates and unemployment rates are the local labour market variables.

---

$^{15}$ Note that we have used subscripts 1 and 2 with Z. This is just to differentiate between two sets of socio-economic characteristics used in different equations.
According to their findings, the academic achievements of students (SAT and its equivalent) play the most important role in application, choice of college quality and admission decisions. Nevertheless, other variables such as high school class rank are also important. One of the most remarkable findings of Venti and Wise (1982) is that individual SAT (or equivalent) scores bear more on application and college quality than on the admission decision.

An advantage of the Venti and Wise (1982) model is that it corrects the college quality equation (2.23) and admission model (2.27) for the selectivity bias. However, all the criticism regarding the selection of a single measure for school quality is applicable to their model as well. There is no solid theoretical foundation for the school quality model.

An alternative specification and estimation method is used by Gertler and Glewwe (1990). According to their specification, there are G+1 options for an individual for whom there are G different schools in the vicinity of his residence. First of all, he is to decide whether or not to continue education. If the no-schooling option is chosen, his utility function is given below.

\[ U_o = \alpha_1 Y + \alpha_2 Y^2 + \varepsilon_o \quad (2.28) \]

where \( U \) is the utility, \( Y \) is family income and \( \varepsilon \) is a random error term. Subscript "o" indicates the no-schooling option. Square of income is introduced into this utility function on two arguments. Since the decision rule compares utilities associated with two alternatives and income is a determinant of the utility associated with any alternative, the income effect will be cancelled out from the demand function if it appears in the utility function in linear form. This does not allow the inclusion of income as an explanatory variable in the demand function. An alternative method is to introduce varying coefficients of income for the utility functions associated with different options. However, Gertler et al (1990) argue\(^{16}\) that this violates the “maximisation of stable utility function”.

If he decides to continue education, he has to choose a school. If the schooling option is chosen, his utility depends on the school selected for education. Utility derived from the schooling option is assumed to be a function of the increment of the human capital stock from one year of education in the selected school and consumption of goods and services other than education. Gertler and Glewwe (1990) assume that increment of human capital stock is a function of school, individual and family characteristics. Utility associated with school “i” is given below.

\[ U_i = \gamma_i Z_i + \alpha_1 (Y - K_i - wh_i) + \alpha_2 (Y - K_i - wh_i)^2 \quad (2.29) \]

where \( \gamma_i Z_i \) is the function determining increment of human capital, \( Z \) is a vector of school, family and individual characteristics, \( Y \) is family income, \( K \) is direct cost of the schooling option (tuition, books, etc.), \( w \) is the opportunity cost of education and \( h \) is the number of hours that a child cannot work due to school attendance. The term in the parentheses is the

\(^{16}\) This point is clearly elaborated in Gertler et al. (1987) and Gertler et al. (1994).
net income available for consumption of goods and services other than education\textsuperscript{17}. Gertler and Glewwe (1990) assume two types of schools: local schools and faraway schools. The maintained hypothesis of this classification is that parents choose faraway schools because the quality of education in these schools is better than the quality of the local schools\textsuperscript{18}. This leads to three alternative choices for the individuals in their sample: no schooling option, local school or faraway school. Gertler and Glewwe (1990) estimate this model using the Nested Multinomial Logit of McFadden (1981). The probability of no-schooling and choosing $i^{th}$ category of schooling is given as

\begin{equation}
\pi_0 = \frac{\exp(U_0)}{\exp(U_0)+[\exp(U_l/\sigma)+\exp(U_f/\sigma)]^\sigma} \tag{2.30.1}
\end{equation}

\begin{equation}
\pi_i = \frac{\exp(U_i/\sigma)}{[\exp(U_l/\sigma)+\exp(U_f/\sigma)]^\sigma} \tag{2.30.2}
\end{equation}

where $\sigma$ is one minus the correlation of error terms in the conditional utility functions of the local and faraway schooling options, subscripts "l" and "f" denote local and faraway schooling options, respectively. They next use Small and Rosen (1981) to calculate the willingness to pay for reducing travel time.

Gertler and Glewwe (1990) argue that the price elasticity does not reveal the overall welfare effect of charging user fees. Price elasticity estimates do not consider the welfare effect of utilising the funds raised in terms of user fees. Therefore, Gertler and Glewwe (1990) estimate the “welfare neutral price” of education. Their estimation assumes that user fees are invested to build schools with more facilities in the vicinity of the respondents’ residence. More specifically, they estimate the willingness to pay for reducing travel time from one and two hours to zero. The estimation procedure is fully articulated in Small and Rosen (1981). Willingness to pay for reducing travel time to zero is calculated for several income quartiles. Estimates of willingness to pay are then compared with the operating cost of schools. Their estimates show that the willingness to pay for reducing travel time by one hour is smaller than the operating cost. However, the willingness to pay is much higher than the operating cost when the travel time is reduced by two hours.

\textsuperscript{17} Inclusion of children’s opportunity cost into this seems rather unrealistic because they have already decided to attend school before choosing a school. Therefore, for school choice, opportunity costs is a sunken cost.

\textsuperscript{18} This assumption is justified within the context where Gertler and Glewwe (1990), estimate the model. This is an application to rural Peru. According to Gertler and Glewwe (1990) the quality of education in rural Peru is not good on average as compared to the average quality of education in the urban sector. Therefore, parents living in the rural sector tend to send their children to faraway urban schools to give them a better education.
2.7 Policy Issues

Most of the literature surveyed in previous sections shows that the schooling decision is positively correlated with family background. This positive correlation arises from two sources. First, the family background has a direct effect on the optimal length of schooling through its effect on "opportunities" and "abilities". The second link is generated through the effect of family background on school choice and in turn, the effect of school quality on the length of schooling. This creates a vicious circle which transmits the inequality from generation to generation. Children of poor parents will leave schools early because their parents are poor. Those children experience a disadvantaged situation in the labour market because they are less educated. Their children will, in turn, be less educated and poor for the same reasons. In other words, they are poor because their parents are poor. Poverty will transmit from one generation to the next through this vicious circle.

Policy makers and social scientists are interested in how to break this vicious circle. According to Becker and Tomes (1986), *ceteris-paribus* this will be broken within several generations because the "earnings regress towards the mean" over generations. The intensity of this convergence depends on the slope of the regression of children's earnings on parents' earnings. The slope has to be a positive fraction and the regression line should intersect the 45-degree line at the positive quadrant. The smaller the slope (closer to the zero), the sooner will be the convergence. And given that rich parents invest a smaller percentage of their income on children's human capital formation [c.f. the section on family background and investment in education in this chapter], education will expedite the process of convergence [see also Polacheck and Siebert (1993), p. 51]. Therefore, Becker and Tomes (1986) suggest that state intervention is not necessary. However, many empirical applications of the Becker-Tomes model have resulted in a relatively large slope of the regression line, indicating that convergence will not occur soon.

There are also theoretical arguments stating that a policy intervention will be ineffective because none of these policies will reduce the early dropouts from full-time education. For example, Manski (1989) develops a theoretical model to evaluate the effects of some selected education policies on the dropout phenomenon. This theoretical model is developed by assuming that the enrolment in a given education programme and the completion of that programme are two different decisions. The enrolment decision is entirely an endogenous decision to the decision-maker, whereas completion of the programme depends on an endogenous factor (*willingness to complete*) and an exogenous factor (*ability to complete*). Manski (1989) assumes that the ability to complete a given programme is not known to the decision-maker. He realises his competence only after attending the programme. Manski (1989) shows that one will drop out before the end of the programme once he realises that he is not competent enough to complete it. In this sense, dropout as well as the completion is an outcome of the optimisation behaviour. He uses this theoretical argument to show that scholarship programmes or any other incentive
packages will not reduce the dropouts substantially. Instead, these policies may increase dropout rates.

At the end of compulsory education one faces three alternative options: go to the labour market and work for a wage, continue education and complete the next level, and continue education and drop out before the end of that level. The utilities associated with these options are designated by $U_w$, $U_c$, and $U_D$, respectively. For the purpose of modelling, Manski (1989) assumes\(^{19}\) that $U_D < U_w < U_c$. He defines an ex-ante probability of completing the stipulated programme, $P_r$. This probability is varying over individuals. One will enrol in post-compulsory education if and only if $P_r U_c + (1 - P_r) U_D > U_w$, the expected utility from enrolling is greater than the utility derived from working. One will be indifferent if the former is equal to the latter. Using this indifference condition, we can solve the above enrolment criterion for the threshold probability of completing the stipulated programme. The enrolment criterion can then be re-stated using the threshold probability. One will choose the enrolment option if his ex-ante probability is greater than or equal to the threshold probability. Manski (1989) assumes that the ex-ante probability has a given distribution.

Policy makers can influence the enrolment decision either by increasing $P_r$, the ex-ante probability of completing a given programme or by increasing $U_c$, utility derived from continuing education. Various types of educational subsidies increase $U_c$, and $P_r$ can be increased by lowering the standards of education. These encourage the less competent and less enthusiastic people to enter the programme. A greater proportion among them will drop out before the end of the programme and therefore increase the dropout rates.

Despite the Becker-Tomes claim that policy interventions are not required and Manski’s claim that policy interventions are ineffective, there are many other reasons for policy makers to intervene within the free market equilibrium of the education system. Equity, efficiency and public concern are considered as three very important reasons [See Polachek and Siebert (1993), pp. 59-67]. Without appropriate state intervention, poor people will not be able to afford their schooling and it is therefore socially unfair to leave the education solely to the private sector. It is necessary to have education subsidies in order to attain an egalitarian society. This is the equity objective of education policy.

Education is an activity with many positive externalities. Therefore, the equilibrium length of schooling attained by wealth maximising individuals is less than the social optimal and hence the resource allocation will not be socially efficient. This is an efficiency justification for state intervention in terms of subsidy policies and compulsory education, etc. Public choice theory refers to the impact of interest groups in politics. There are many interest groups with education as one of the prime targets. For example, teachers'\(^{19}\) Note that all other possible orderings generate trivial behavioural patterns. For example, if $U_w$ has the highest value, everyone will work, and if $U_D$ is the highest, everyone will enrol and drop out before the end of the programme.
associations in the USA and many European countries are strongly influencing the
formation of education policies of their respective countries. In Sri Lanka, almost all
voters in general, and lobbying groups led by university students in particular, are
extremely interested in the matter.

There are three motives that generate massive policy interventions in the field of
education. One can divide the complete range of currently practised education policies
into: 1.) the compulsory education policy, 2.) subsidy policies, and 3.) policies aimed at
the improvement of the quality of education. This section presents remarks pertaining to
the evaluation of these policy options within a given theoretical framework.

Compulsory Education: Education is compulsory until a certain level everywhere in the
world. The legal enforcement of compulsory education compels parents to send their
children to full-time education until the socially determined level is reached. This policy is
not without drawbacks. For example, from the policy makers' perspective, proper
implementation of compulsory education policy is rather difficult. For the effective
implementation of this policy there should be a mechanism to monitor the system. There
should be a mechanism to check whether parents send their children to school and to
punish the parents who disobey the law. This system requires a bureaucracy and therefore
inefficiencies associated with a bureaucracy will affect the proper implementation of the
completion of compulsory education. From the individual’s perspective, this policy will be
felt to be unfair by those for whom the unconstrained equilibrium length of schooling is
less than the compulsory length, because under the latter, their perceived marginal rate of
return is less than the marginal cost. However, in hindsight the policy may benefit them. A
key issue is of course the optimal length of compulsory education.

Education Subsidies: There are many different forms of education subsidies, varying
from completely free education (case of Sri Lanka) to providing various types of education
subsidies. In fact, the argument is that some people leave school before the socially
desirable level because the cost of education is higher for them. Provision of subsidies
presumably lowers the cost and thereby encourages pupils to stay in full-time education for
a longer period than their free market equilibrium level. Currently, there are many different
forms of subsidy policies. Free education by the public schools, soft loans, scholarships,
grants, education auctions, and education vouchers are but a few examples [For a detailed
discussion of various types of education subsidy policies, see Johnes (1993), Chap. 7].
While each subsidy system has its own merits and demerits, Peltzman (1973) [See
Polachek and Siebert (1992) for a simpler graphical exposition] contributes a general
criticism regarding the effectiveness of education subsidies. Peltzman (1973) compares
three types of subsidy schemes: offering a sum of money that can be used only for
education, rebating the cost of education subject to a certain maximum, and offering
subsidised education in a completely different and secluded set-up (he calls this in-kind
subsidy). The first two options always shift at least a part of the budget constraint to the
right, allowing the individual to choose any desirable amount of education and other
commodities. Therefore, these schemes would not have any negative impact on the choice of education by individuals. However, the third scheme, (which he argues is quite often the case in the US higher education), would motivate negative reactions at least among a segment of the population and thereby the total effect on the society would be unpredictable.

The so-called Peltzman effect argues that the in-kind subsidy affects the schooling decisions of different social classes differently. It will increase the amount of schooling of the poor while it has no effect on the schooling decision of the rich people. However, the effect of this policy on the amount of schooling attained by the middle-class would be negative. Therefore, whether the in-kind subsidy scheme will increase or decrease the average educational attainments of the society depends on the relative size of the middle-class. If the middle class is larger than the other two, it will be counter-productive.

This theory is based on the assumption that the education system consists of two secluded set-ups: subsidised schools managed by the state and the unsubsidised schools. No one is allowed to move between them. The subsidised option offers free education up to a certain limit and presumably this limit is higher than the free market equilibrium of the poor and it is smaller than the free market equilibrium of the middle-class and the rich people. An individual can choose either the subsidised option or the unsubsidised one. Once the choice is made, he has to stick to it. For example, choosing unsubsidised education means a complete rejection of the subsidy. Or in other words, choosing subsidised education means that one has to remain in the subsidised education system until the end and he cannot move back to the unsubsidised option once the subsidised education is completed.

If the subsidised option is chosen one can spend all income to buy other commodities. This means that one can move to a corner of the budget constraint: buy no additional education, spend all money for other commodities and utilise only the subsidised amount of education. If the non-subsidised option is chosen, the individual moves to the free market equilibrium.

In the case of poor people, the subsidised amount of education is greater than the free market equilibrium amount by construction of the model. Therefore, the best alternative for them is to move to the subsidised education. This increases the consumption of education and other commodities. At the other extreme, rich people consume a larger amount of education as well as more of the other commodities. Therefore, the marginal utility from education and other commodities is very small. If they decide to move into the subsidised education, they have to reduce the consumption of education by a substantial amount. This is associated with some loss of utility. They can also increase the consumption of other commodities by moving to a corner of the budget constraint. This presumably increases the total utility. However, the utility gain will not be sufficient to

---

20 Peltzman (1973) refers to the monetary value of education. However, it can easily be extended to the units of human capital or to length of schooling.
compensate for the loss of utility due to the reduction of the consumption of education. The rich people will therefore stick to the free market equilibrium. The free market equilibrium amount of education of the middle-class is larger than the subsidised amount of education. However, Peltzman (1973) argues that for an individual from the middle-class, utility loss associated with reducing education is smaller than the utility gain due to the increase of the consumption of other commodities. Therefore, the middle-class will reduce the consumption of education with the in-kind subsidy.

General subsidy programmes such as free education are criticised for their cost ineffectiveness because they can also include subsidies for those who can pay. Specifically, the policy of free education can even be counter-productive. This policy will minimise the direct cost of education subject to a substantial amount of supplementary cost. Therefore, this policy will not encourage those who quit school earlier due to reasons other than the direct cost and those who cannot bear even the supplementary cost. For example, a free education policy will not help less able and less motivated students. Given the assumption that family background has a positive effect on these variables, all these imply that a greater part of the benefits of free education will be accrued by students with advantaged family backgrounds. From this perspective, free education subsidises the education of well-off children.

An alternative policy to overcome this drawback is to arrange individual specific subsidies. Scholarships, grants and soft loans are the most popular individual specific subsidy programmes. Schools or the policy makers have to identify the people really in need of the subsidy and allocate grants for them. In general, scholarship programmes are on a need basis and a merit basis. Granting scholarships on a merit basis is a rather easy task. However, the need base raises some practical problems. First of all, it is not easy to identify eligible people without the effect of inefficiencies and favouritism. Therefore, the programme may fail to derive the desired outcome. Effectiveness of the merit-base subsidy programmes is rather questionable. In an egalitarian society [c.f. section 2.4], where educational variations exist only because some people have more opportunities (low cost) than others, merit-base subsidy programmes will shift the cost curves downward in a parallel way. This will increase the optimal length of schooling of all groups and therefore the average length of schooling of the society will rise. It may also reduce the disparity of educational attainment because the equilibrium of the group with more opportunities cannot exceed the highest level determined by the education system. According to the elite model, the prediction will be entirely different: merit-base subsidies will increase the disparity.

Effectiveness of the need-base subsidy programmes depends on the relationship between the economic environment of the individual, and the costs and the rate of return to education. If the conventional belief (that rich are more able and have more opportunities) holds, need-base subsidy policies will always reduce the disparity in education. However,
earnings disparity will not be affected in an elite society (able people: elite will always have a higher rate of return to education in equilibrium).

Another criticism of these programmes is that they do not give any incentive to schools to work hard. Therefore, others recommend more sophisticated general and specific subsidy programmes such as education auctioning and education vouchers.

Education auctioning involves the use of a closed-bid auction to determine the price of education and thereby to encourage schools to work more efficiently. A central authority responsible for education can set the quality targets of education and then call closed-bids from all the schools that guarantee the minimum quality standard for a per-pupil price for each subject area and number of students that they can recruit for the coming academic year. Finally, the central authority can accept the bids with lowest per-pupil price with the guarantee of minimum education standard. Students’ fees are paid by the government or by other donors. This system reduces the public expenditure of schooling by increasing the competition between schools whereby schools are motivated to provide education at their lowest cost. This system presumably increases the efficiency that in turn reduces the problem of x-efficiency in the education production process.

In education, the voucher system proposes to grant a voucher to all eligible pupils that they can spend at any school. This gives more autonomy to the pupils in choosing a school. Therefore, schools will have to be more efficient and competitive in attracting more pupils. There are many variants of the voucher system offering different degrees of autonomy to pupils [see Levin (1992), Johnes (1993), Chap. 7 and also Rouse (1998)]. This system can also be combined with the auction system. First nominate schools and determine the price on the basis of a closed-bid auction, and then determine the value of the vouchers.

Improving the Quality of Schooling: A plethora of studies on the education production function and school efficiency have emerged with mixed results [see section 2.5 of this chapter]. In general, the literature, both at individual as well as aggregate levels, suggests that there is no consistent relationship between school inputs and outputs [See Hanushek et al. (1992)]. Therefore, it is rather difficult to suggest the most cost-effective ways to improve the quality of schools. However, recent works on this issue have tended to recommend more structural changes and changes in administration systems and incentive schemes rather than to pump more and better inputs into schools. [See for example, Hanushek et al. (1992) and Levin (1997)].

This argument is precisely elaborated in Levin (1997), where he explains the education production process as an inefficient system. There are a number of reasons for this claim. There are several factors that promote efficiency: a clear objective function with measurable outcomes, incentive schemes, access to useful information, adaptability to changing circumstances, and use of the most productive technology consistent with the cost constraints. Levin (1997) argues that the school production process satisfies none of
these conditions. Therefore, "A tacit assumption will be that greater efficiency gains in educational production can be derived from dramatic organisational changes than from tinkering with reallocations of existing school inputs..."[Levin (1997), p. 303].

Hanushek et al. (1992)] is very similar to Levin (1997) in his assertion of policy implications. Hanushek et al. (1992), having observed the same inefficient school production process, recommend incentive schemes for teachers and schools on a merit basis. Schools and teachers that do better must be paid more than the others. However, difficulties associated with measuring performance hinder the efficiency of this method.

Presumably, improving school quality will affect the optimal schooling length by its effect on the ability. Good schools always train pupils better. We can therefore expect that quality of schooling directly increases the ability [see Card and Krueger (1992a, 1992b)] and thereby increases the optimal length of schooling. However, at the same time we expect that quality of education and supplementary cost of education are always positively correlated, so improving the quality of schools will associate with a backward shift of the cost curve. This makes the effect of quality of school on education enrolment unpredictable at a theoretical level21.

2.8 Summary, Conclusions and Implications on the Present Study

2.8.1 Summary and Conclusions

The main purpose of this chapter has been to review the literature related with the issues addressed by this book. Namely, factors determining the educational performance and school choice at the individual level and the understanding of the technological aspects of the education production process. In addition to the literature, this chapter began with a very brief review of the functional relationship between the level of education and earnings. The literature reviewed in this chapter has explored three major links: education as a signal [Spence (1973)], cognitive ability and personality traits [Gintis (1971)].

Theoretical and empirical literature related with the issues of demand for education explores several possibilities for family background effect on educational performances. Under uncertainty and with given assumptions of risk increasing with level of education and decreasing absolute risk-aversion, wealthy people will invest more in education. If the education system and the labour market returns to education are certain, family background can still affect the schooling decision via imperfect capital market and ability. Due to the capital market imperfections, people have to rely on self-financing or they have to mortgage their physical assets to finance education. This makes education affordable only for children from affluent families. Furthermore, some children cannot attend school

21 Of course there are many other sources that make this effect unpredictable. For example, if quality will increase ability it will also increase the opportunity cost.
simply because they have to work to compensate for the low income of their parents. On the other hand, ability is not uniformly distributed: some are more capable than others. Theoretical conventions further prompt us to believe that children with affluent family backgrounds are more capable. What this story tells us is that unless an intervention occurs, family background effect on educational and labour market performances will remain, and therefore a vicious-circle of poverty will persist over generations. Theories on the school production process and the demand for schools show us that this transmission will be reinforced by the existing pattern of facility distribution and the mechanism that allocates pupils with different family backgrounds over schools.

All these focus on the issues related with education policy. There is vast literature on the education policy issues which explores different forms of interventions: law enforcement, subsidising education, soft loan schemes, scholarships, education vouchers, auctioning education, and improving and equalising the quality of education.

The main theories on the schooling decision that we have reviewed in this chapter are Becker (1967) and Ben-Porath (1967). Overall themes of the present study (i.e. free education, other subsidy programmes, efficiency (quality) of schools and the demand for quality of schools) can be combined with these theories to give a rough sketch of the theoretical foundation of this study.

As we have already explored in section 2.7, all the education subsidy policies divide into general subsidies and individual specific subsidies. The general subsidies benefit all the age-eligible pupils equally, whereas individual specific subsidies benefit only those who are selected for the subsidy. The free education policy is an example of the general subsidy programmes. Regardless of whether they are general or specific, all the subsidy programmes presumably reduce the cost of education. Within Becker’s framework, education subsidy programmes shift the marginal cost functions to the right, enabling beneficiaries to obtain a given length of schooling at a lower cost. However, as we have shown in section 2.7, general subsidy programmes will not necessarily reduce the disparity of educational attainment between rich and poor people. General subsidy programmes, by reducing the cost of education equally, increase the optimal length of schooling of all the beneficiaries. However, individual specific subsidies such as scholarships on a need basis, will reduce the cost only of the beneficiaries. Therefore, such a policy increases the length of schooling of the economically less affluent pupils towards the optimal length of schooling of economically affluent pupils. However, its effects on labour market returns depend on the underlying model. For example, under the elite model, where the individual variation of schooling is due to the variations of personal capabilities, the marginal rate of return to education of the elite group is higher than the marginal rate of return of the backward group even after the subsidy. (I.e. the rate of return curve of the elite will always locate above the rate of return curve of the backward group). Under the egalitarian model, the rate of return to education will be equalised for all after the subsidy.
Ben-Porath (1967) also presents some interesting implications concerning the education subsidies. Under the Ben-Porath (1967) model, education subsidies will lower the price of purchased inputs \( (P_d) \) in equation (2.5). The first derivative of equation (2.5) with respect to \( P_d \) is positive, indicating that the lower \( P_d \) rotates the marginal cost downward through the origin. However, the vertical segment of the short-run marginal cost curve is unaffected. These clearly indicate that subsidy programmes increase the length of full-time education. Implications on the effect of general and specific subsidy programmes are on par with those of Becker (1967).

The implications of Becker (1967) and Ben-Porath (1967) differ on improving quality of schooling and keeping cost constant. Under the Becker framework, quality of schooling can be introduced through the rate of return. Pupils being trained in resource-rich schools are more productive and therefore earn a higher rate of return to education [Card and Krueger (1992a, 1992b)]. This shows that \textit{ceteris-paribus}, the quality of education increases the optimal length of schooling.

Ben-Porath’s (1967) model predicts that the quality effect is ambiguous. Quality of education can be incorporated into this model either through variable \( D \) (purchased inputs) or \( \beta_2 \) in equations (2.4) and (2.5). Improved school quality increases the efficiency of the human capital production process [equation (2.4)]. This lowers the long-run marginal cost curve. The vertical segment of the short-run marginal cost curve shifts outward as well, because the efficiency of the human capital production function is now improved. A downward rotation of the long-run marginal cost curve increases the length of full-time education. However, the outward shift of the short-run marginal cost curve shortens the length of full-time education. Therefore, the effect of the school quality on length of schooling is unpredictable in the Ben-Porath model.

If the quality improvement is associated with an increase in the cost of education, (levy a user-fee to improve the quality of education, for example), predictions will be different. According to Becker’s framework, this shifts rate of return and cost curves upward whereby the ultimate effect is unpredictable. According to the Ben-Porath model, such a policy will at least partially offset the rotation of the long-run marginal cost curve. Under the Ben-Porath framework, the ultimate effect is uncertain. However, it is more likely to be negative because only the positive effect of quality improvement is offset.

In the remaining chapters of this book the theoretical framework is concentrated around the simplest version of Becker (1967). Ben-Porath’s framework and the extensions of the Becker model with uncertainty and consumption motives are ignored. Ben-Porath (1967) offers a sound theoretical framework for analysing the investment in human capital in general. However, modelling the length of schooling within the Ben-Porath framework is not straightforward. This requires further assumptions and complicates the model. The uncertainty and consumption motives are beyond the focus of the present study. Therefore, the simplest version of Becker (1967) will suffice for the purpose of this book.
2.8.2 Implications on the Present Study

The theme of the present study is the free education policy in Sri Lanka. This policy provides education without tuition fees for everyone until university education. As far as this theme is concerned, this literature review presents several hypotheses which can be tested empirically. We list these hypotheses below. The remaining chapters of this book test them and draw conclusions using empirical data collected from Sri Lanka.

- There are many links between family background and educational attainments: uncertainty, imperfect capital markets and the relationship between family background and ability. These suggest that the tuition fee is only one among many factors preventing the poor from staying longer in full-time education. Therefore, free education alone will not be a sufficient strategy for breaking the family effect on educational attainment.

- In the presence of facility and quality disparities in schools, family background may enter into the choice of schooling in such a way that pupils with affluent family backgrounds will choose better schools.

- The presence of family effects on school choice will reinforce the family background effect on education because quality of school has a direct effect on educational attainment.