
Measuring the option value of education

RULOF P. BURGER AND FRANCIS J. TEAL

Stellenbosch Economic Working Papers: 15/13

KEYWORDS: EARNINGS REGRESSIONS, RETURNS TO SCHOOLING, DYNAMIC PROGRAMMING

JEL: C31, J24, J31

RULOF BURGER
DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH
PRIVATE BAG X1, 7602
MATIELAND, SOUTH AFRICA
E-MAIL: RULOF@SUN.AC.ZA

FRANCIS J TEAL
CENTRE FOR STUDIES OF AFRICAN ECONOMICS
UNIVERSITY OF OXFORD
E-MAIL: FRANCIS.TEAL@ECONOMICS.OX.AC.UK



UNIVERSITEIT
STELLENBOSCH
UNIVERSITY



A WORKING PAPER OF THE DEPARTMENT OF ECONOMICS AND THE
BUREAU FOR ECONOMIC RESEARCH AT THE UNIVERSITY OF STELLENBOSCH

Measuring the option value of education¹

RULOF P. BURGER^{2,3} AND FRANCIS J. TEAL³

ABSTRACT

Many recent descriptive studies find convex schooling-earnings profiles in developing countries. In these countries forward-looking students should attach option values to completing lower levels of schooling. Another option value may arise due to the uncertain economic environment in which the sequence of enrolment decisions is made. Most theoretical models that are used to motivate and interpret OLS or IV estimates of the returns to schooling assume away convexity in the schooling-earnings profile, uncertainty and the inherently dynamic nature of schooling investment decisions. This paper develops a decomposition technique that calculates the relative importance of different benefits of completing additional schooling years, including the option values associated with convex schooling returns and uncertainty. These components are then estimated on a sample of workers who has revealed a highly convex schooling-earnings profile, and who face considerable uncertainty regarding future wage offers: young black South African men. We find that rationalising the observed school enrolment decisions requires large option values of early schooling levels (mainly associated with convexity rather than uncertainty), as well as a schooling cost function that increases steeply between schooling phases.

Keywords: Earnings Regressions, Returns to Schooling, Dynamic programming
JEL codes: C31, J24, J31

¹ The authors acknowledge the financial support from the University of Stellenbosch Faculty of Economic and Management Sciences' Elite Research Fund.

² Corresponding author: Senior lecturer, Economics Department, University of Stellenbosch.

³ Research associate, Centre for Studies of African Economics, University of Oxford.

1. INTRODUCTION

The internal rate of return to schooling is central to understanding the earnings distribution, as well as the schooling investment decisions made by individuals. Becker and Chiswick (1966) demonstrated how this return can be estimated econometrically, and following the extension and popularisation provided by Mincer (1974), the schooling return soon became one of the most researched parameters in all of economics. However, the wage regression schooling coefficient that is frequently reported as an estimate of the schooling return is more accurately interpreted as the price of schooling from a hedonic market wage equation, a concept which is only loosely related to the parameter of interest. After almost half a century of empirical research Heckman, Lochner and Todd (2006, p. 311) argue that the conventional econometric methods used to estimate this parameter are fundamentally flawed and that the returns parameter remains “widely sought after and rarely obtained”.

Most theoretical models of human capital investment that are used to guide interpretation of the returns to schooling estimates (for example, Becker (1967), Card (1999)) have maintained the assumptions that the schooling investment decision is i) a once-off decision regarding how many years of schooling to complete, rather than a sequence of enrolment decisions; ii) made with perfect foresight regarding future schooling costs and wage offers; and iii) that the proportional effect of schooling on wages is either constant or decreasing in schooling years. Such models ignore many potentially important aspects of the schooling investment decision, such as the option value of education that may arise due to either convexity in the schooling-earnings profile or uncertainty regarding future wage offers or schooling costs (Heckman, Lochner & Todd (2008)). In fact, without such option values it is difficult to explain why, when faced with a convex schooling-earnings profile, so many individuals choose to complete the initial low-yielding school years and then stop investing at a point where the returns are still rising.

The paper begins by reviewing the treatment of convexity and uncertainty in the returns to education literature in section 2. In order to investigate these issues, section 3 develops a dynamic model of individual schooling investment based on that of Belzil & Hansen (2002). A decomposition technique that calculates the relative importance of different benefits of obtaining additional schooling, including the option values associated with convex schooling returns and uncertainty, is then derived. Sections 4 briefly discusses the panel data used to estimate the model, before estimating each of the components of the schooling benefit on a sample of workers who has revealed a highly convex schooling-earnings profile, as well as considerable uncertainty regarding wage offers: young black South African men.

2. ESTIMATING THE RETURN TO SCHOOLING INVESTMENT

Much of the returns to schooling literature assume that the effect of schooling on log earnings is either constant or a decreasing function of schooling year. This assumption was initially motivated by evidence from early Mincerian earnings regressions (for example, Mincer (1974), Becker (1964) and Psacharopoulos (1973, 1985, 1994)). However, a number of recent studies suggest that the schooling-earnings profile has turned increasingly convex, both in the US (Mincer (1996), Heckman, Lochner & Todd (2008), Lemieux (2006)) as well as in a number of developing countries (Appleton, Hoddinott, & MacKinnon, 1996; Carnoy, 1995; Nielsen & Westergaard-Nielsen, 2001; Siphambe, 2000; Teal, 2001; Whaba, 2000), including South Africa (Keswell & Poswell, 2004).

The continued popularity of the linearity or concavity assumption partly derives from its appeal in explaining observed schooling investment decisions. In most countries only a small share of individuals choose schooling corner solutions – obtaining no schooling at all or the maximum number of schooling years – and the prevalence of interior solutions is easily explained if the marginal cost of schooling is an increasing function of schooling year while a concave schooling-earnings profile is maintained. Assumptions of this nature form the basis for many prominent models of the schooling-earnings nexus (for example, Becker (1967), Card (1999)). If the schooling return is actually low initially, and then increases as a student progresses up the schooling ladder, that makes it more difficult to explain the observed distribution of schooling outcomes, since most individuals then obtain the low-yielding schooling levels but drop out just as the benefits start to materialise.

Furthermore, the definition of the “value” of an additional schooling year also changes when schooling investment is viewed as a sequence of decisions about whether or not to enrol in school for another year, rather than a once-off choice of how many years of schooling to complete. Heckman, Lochner and Todd (2006) demonstrate that the returns to schooling can include an option value if there is either uncertainty about future wage offers or schooling costs, or if the schooling-earnings profile is convex. Standard Mincerian earnings functions that assume earnings to be log-linear in schooling while ignoring the role of uncertainty will not be able to capture the option value of schooling, and will therefore be unable to reconcile the apparent benefit of an additional schooling year with the revealed schooling preferences of individuals. Even structural models that estimate the schooling returns while allowing for a dynamic sequence of school investment decisions have so far neglected to separately identify the different option values associated with an additional schooling year.

Discrete choice dynamic programming provides a modelling framework in which the above issues can be potentially addressed. It aims to estimate the policy invariant “deep parameters” (Heckman, 2010) that appear in the individuals’ utility and cost functions rather than some policy-dependent weighted average of the marginal returns. Although this approach requires making strong assumptions about the behaviour of individuals, these assumptions allow the econometrician to investigate the precise motivations for their decisions and to simulate the effects of hitherto unobserved policies. Furthermore, whereas most recent IV studies choose to interpret the 2SLS estimates as local average treatment effects that are not informative about the shape of the schooling-earnings profile, structural models can estimate the shape of this profile.

3. A MODEL OF EDUCATION INVESTMENT

3.1 MODEL ASSUMPTIONS

We now develop a simple model of individual schooling decisions and labour market outcomes that will allow us to identify the different benefits of school investment. The Belzil & Hansen (2002) model is used as a point of departure. A young individual i is endowed with household income h_i and then has to make a sequence of schooling investment decisions in which the benefit of entering the labour market is weighed up against the benefit of enrolling in another year of schooling.

We assume that in period t , individual i with s_{it} years of completed schooling derives utility according to

$$u_{it}^1 = h_i \psi_1 + \psi_{21} 1(7 < s_{it} + 1 \leq 12) + \psi_{22} 1(12 < s_{it} + 1) + \eta^s + \varepsilon_{it}^s \quad [1]$$

where η^s is the schooling utility intercept and ε_{it}^s is a stochastic schooling utility shock. Years of completed schooling enter this utility function through a piecewise constant function with discontinuous jumps between primary, secondary and tertiary education⁴. The ψ_{21} and ψ_{22} coefficients represent the marginal utility that individuals expect to experience while in secondary and tertiary education respectively (relative to being in primary school), and are perhaps most plausibly interpreted as representing the differential schooling costs – either monetary or psychic – associated with shifting between different schooling phases. Household income acts as a schooling shifter either by increasing parental transfers to those enrolled in school or by alleviating credit constraints.

Labour market participants derive utility according to $u_{it}^0 = m_{it} \ln W_{it}$ where m_{it} is a binary employment status variable and W_{it} measures annual labour market earnings. The data generating process for earnings has the same specification as a Mincerian earnings function with a quadratic schooling term

$$\ln W_{it} = w_{it} = \chi_1 s_{it} + \chi_2 s_{it}^2 + \chi_3 x_{it} + \chi_4 x_{it}^2 + \eta^w + \varepsilon_{it}^w \quad [2]$$

where x_{it} is potential years of work experience, η^w is the earnings intercept and ε_{it}^w represents a stochastic earnings shock. Employment is determined according to

$$m_{it} = 1(\kappa_1 s_{it} + \kappa_2 s_{it}^2 + \kappa_3 x_{it} + \kappa_4 x_{it}^2 + \eta^m + \varepsilon_{it}^m > 0) \quad [3]$$

where η^m represents the employment intercept and ε_{it}^m is a stochastic employment shock. Unlike in the Belzil-Hansen model, employment status is a binary variable that may depend non-linearly on years of completed schooling.

Individuals choose whether to enrol in school ($d_{it} = 1$) or to enter the labour market ($d_{it} = 0$), and this decision is guided by their desire to maximise discounted expected intertemporal utility

$$E \left[\sum_{t=\tau}^{T-t} \beta^j U(d_{i,t+\tau}, \mathbf{q}_{i,t+\tau}, \boldsymbol{\varepsilon}_{i,t+\tau}) \mid d_{it}, \mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it} \right]$$

where $U(d_{it}, \mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}) = (1 - d_{it})u_{it}^0(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}) + d_{it}u_{it}^1(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it})$, β is the subjective annual discount factor, \mathbf{q}_{it} is a vector of state variables containing accumulated schooling, potential experience, household income and whether the individual has left school or not, and $\boldsymbol{\varepsilon}_{it}$ is a vector of current period stochastic shocks.

The transition of the schooling state variable is determined as $s_{i,t+1} = s_{it} + d_{it}$. The schooling variable is restricted to lie between 0 and 16, and individuals automatically leave the school system upon completing

⁴ The South African school system consists of 7 years of primary education followed by 5 years of secondary school. The LFS panel data (discussed in section 4) allow us to distinguish between post-graduate diplomas or certificates (the holders of which are assigned 13 years of completed schooling), an undergraduate university degree (15 years) and post-graduate degree (16 years).

16 years of schooling. No individuals are allowed to enrol in school before the age of 6 and no wage offers are made to individuals below the age of 16. Once individuals have entered the labour market, they remain there until the retirement age of $T = 65$ and are not allowed to re-enter the school system. The error terms are assumed to be identically and independently normally distributed: $\varepsilon_{it}^s \sim n(0, \sigma_s^2)$, $\varepsilon_{it}^w \sim n(0, \sigma_w^2)$ and $\varepsilon_{it}^m \sim n(0, 1)$. It is straight-forward to also allow for permanent unobserved heterogeneity or stochastic school interruption, but these issues are ignored in order to simplify the decomposition technique below.

3.2 IDENTIFICATION AND ESTIMATION

The assumptions above allow us to express the choice-specific value function associated with state variables \mathbf{q}_{it} as

$$V_{it}^d(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}) = u^d(\mathbf{q}_{it}) + e_{it}^d + \beta \int \bar{V}_{i,t+1}(\mathbf{q}_{i,t+1}) dF(\mathbf{q}_{i,t+1} | \mathbf{q}_{it}, d) = \bar{V}_{it}^d(\mathbf{q}_{it}) + e_{it}^d$$

where $\bar{V}_{it}^d(\mathbf{q}_{it}) = \int V_{it}^d(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}) d\Phi(\boldsymbol{\varepsilon}_{it})$ is the integrated value (or Emax) function, $u^d(\mathbf{q}_{it})$ is the deterministic component of contemporaneous utility, e_{it}^d is a stochastic error term, and $F(\cdot)$ and $\Phi(\cdot)$ are the cumulative distribution functions for the state variables and error terms, respectively. Additive separability⁵ in the choice-specific current period utility function means that the choice-specific value function can also be written as the sum of a deterministic component $\bar{V}_{it}^d(\mathbf{q}_{it})$ and the current period stochastic utility shock. The dynamic programming problem can then be fully characterised by the integrated Bellman equation:

$$\bar{V}_{it}^d(\mathbf{q}_{it}) = \text{Emax}(V_{it}^0(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}), V_{it}^1(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}) | \mathbf{q}_{it}) = \int \max_{d_{it}} \{V_{it}^d(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it})\} d\Phi(\boldsymbol{\varepsilon}_{it})$$

and the conditional choice probability expressed in terms of the value functions:

$$P(d_{it} = 1 | \mathbf{q}_{it}) = P(V_{it}^1(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}) > V_{it}^0(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}) | \mathbf{q}_{it})$$

The model parameters are estimated in a two-step procedure that consists of an inner and an outer iteration loop⁶. Firstly the model is solved for an initial parameter vector $\boldsymbol{\theta}_0$. This step entails calculating the values of $\bar{V}_{it}^d(\mathbf{q}_{it}, \boldsymbol{\theta}_0)$ and $P(d_{it} = 1 | \mathbf{q}_{it}, \boldsymbol{\theta}_0)$ using backwards induction. In order to calculate these values the household per capita income variable h_i is discretised into 10 deciles and an annual discount factor of $\beta = 0.9$ is assumed.

The likelihood function for individual i in period t consists of five components: the probability of having completed s_{it} years of schooling by period t ; the probability of enrolling in school for another year; the probability of entering the labour market for the observed wage offer; the conditional employment

⁵ The additive error term e_{it}^d is not the same as the error terms in equations [1], [2] or [3], and is therefore not generally independent of the state variables and does not follow a common distribution. This means that the optimisation problem does not have a closed-form solution, and the Monte Carlo integration approach pioneered by Keane and Wolpin (1994) is used to approximate these values instead.

⁶ A more comprehensive discussion of these steps are contained in the Technical Appendix.

probability, and the conditional wage density. The outer iteration loop uses numerical optimisation techniques to find the parameter vector that maximises this likelihood function.

3.3 THE OPTION VALUES OF SCHOOLING INVESTMENT

Agents in the estimated model make human capital investment decisions in a way that is rational, dynamically optimal and consistent with observed schooling, wage and employment outcomes for young black South African men. One of the primary benefits of deriving the estimable model directly from a choice-theoretic framework is that it allows us to investigate the economic reasoning behind individual schooling decisions. In fact, the structural model allows us to explicitly quantify the perceived benefits associated with different schooling years, and to decompose this into a myopic benefit and two types of option values.

Heckman, Lochner and Todd (2006) demonstrate how both convexity in the earnings profile and uncertainty about future schooling and labour market shocks can mean that there is an option value to educational investment. Completing lower levels of schooling allows access to later and higher yielding schooling levels, and also provides learners with more information about the stochastic elements of schooling utility and wage offers. Our model uses a different estimation strategy than the one suggested in Heckman, Lochner and Todd (2006), but similarly allows us to quantify each of these effects separately for the various schooling years.

The expected value of enrolling in schooling year $s_{it} + 1$ is expressed⁷ as $V_{it}^1(s_{it}) - V_{it}^0(s_{it})$. This perceived benefit can be expanded into four additive components. *The myopic net benefit*

$$u^1(s_{it}) + \beta \bar{V}_{i,t+1}^0(s_{it} + 1) - \bar{V}_{it}^0(s_{it})$$

is the net expected benefit of completing a single additional year of schooling before entering the labour market. This corresponds to the only systematic determinant of schooling investment in human capital models that do not make allowance for uncertainty or convex schooling-earnings profiles, and consists of three terms. Firstly, individuals who enrol in school can expect to experience instantaneous utility (or cost), $u^1(s_{it})$. Secondly, individuals will expect to earn higher wages and be more employable after successfully completing an additional year of schooling: $\beta [\bar{V}_{i,t+1}^0(s_{it} + 1) - \bar{V}_{i,t+1}^0(s_{it})]$. Thirdly, enrolling in school entails an implicit cost in terms of foregone labour market earnings, $\beta \bar{V}_{i,t+1}^0(s_{it}) - \bar{V}_{it}^0(s_{it})$.

The second schooling benefit is the *convexity option value*

$$\beta \left[\max \left(\bar{V}_{i,t+1}^0(s_{it} + 1), \bar{V}_{i,t+1}^1(s_{it} + 1) \right) - \bar{V}_{i,t+1}^0(s_{it} + 1) \right]$$

and measures the value of gaining access to more advanced and higher yielding schooling years after completing the current year. Individuals who enrol in school will receive a positive option value if the expected value of enrolling again after completing the current school year exceeds the expected value of entering the labour market; otherwise the value of this component is zero.

The third component represents the *uncertainty option value*

⁷ Our notation suppresses the dependence of the choice-specific value functions on the stochastic error terms and the state variables other than schooling.

$$\beta \left[Emax \left(V_{i,t+1}^0(s_{it} + 1), V_{i,t+1}^1(s_{it} + 1) \right) - max \left(\bar{V}_{i,t+1}^0(s_{it} + 1), \bar{V}_{i,t+1}^1(s_{it} + 1) \right) \right]$$

and captures the benefit of staying in school for another year so as to observe the wage and schooling error draws for another year. This is especially beneficial for those at the margin of the school-work decision, for whom a slightly larger (smaller) than expected schooling utility (wage offer) in period $t + 1$ could make it beneficial to invest in another schooling year.

The fourth component is the *stochastic net benefit*, $e_{it}^1 - e_{it}^0$, which represents the difference between the current period school utility and wage offer shocks. This indicates the magnitude with which the current expected benefit of school enrolment exceeds its expected value from the previous period. Since individuals form expectations rationally by assumption, this component is orthogonal to all the state variables, and does not help account for systematic differences in individual behaviour.

4. DATA

We estimate the model using the panel component of the Labour Force Survey (LFS) dataset gathered by Statistics South Africa. The LFS was a nationally representative household survey conducted twice a year between 2000 and 2007. The households sampled between September 2001 and March 2004 were allocated according to a rotating panel design, with 20% of the original sample being replaced by new households in the next round of the survey (apart from September 2002). It is worth taking note of a few problems with this particular dataset. Statistics South Africa did not publish survey weights for this data (Vermaak, 2010), and our estimates are therefore all taken from unweighted regressions. Furthermore, the data was collected as a rotating panel of dwelling units, which implies that individuals were dropped from the sample once they left the dwelling. This may result in non-random attrition which could potentially bias our estimates. However, the LFS panel is the only South African panel datasets with enough observations to estimate this sort of model. In order to avoid issues of racial and gender discrimination, as well as the need to model fertility decisions, we restrict our estimation sample to black males between the ages of 6 and 30. This gives us 27,331 individuals and a total of 54,501 observations.

5. RESULTS

Table 1 reports the coefficient estimates from the dynamic programming model. The schooling profiles (for earnings and employment) are both convex, whereas the experience profiles are concave. The school utility coefficients associated with being enrolled in the various schooling levels suggest that being in secondary school is more costly (in terms of either monetary or psychic costs) than being in primary school, and that tertiary education is much more costly than secondary school. Household income is shown to be positively correlated to school utility and hence the educational attainment of school-going children. Finally, the standard deviation of the schooling utility error term is much higher than those of the wage or employment processes, which suggest that unobservable schooling utility determinants are more important in determining schooling enrolment than the variation in wage offers.

Table 1: Coefficient estimates

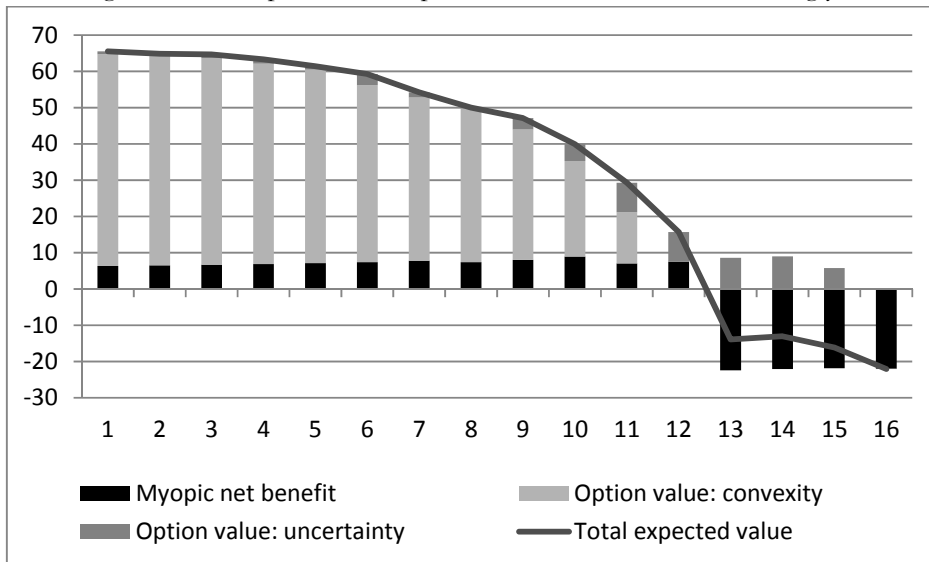
	Wage		Employment		School utility	
Schooling	-0.034***	(0.0014)	-0.154***	(0.0014)		
Schooling ²	0.012***	(0.0001)	0.012***	(0.0001)		
Experience	0.102***	(0.0017)	0.079***	(0.0009)		
Experience ²	-0.002***	(0.0001)	-0.002***	(0.0000)		
Secondary school					-0.903***	(0.0159)
Tertiary education					-31.417***	(0.2062)
Household income					0.715***	(0.0046)
Constant	4.997***	(0.0225)	-0.408***	(0.0060)	3.060***	(0.0263)
Std. dev.	0.852***	(0.0038)	1		38.889***	(0.0041)

Notes: *Statistically significant at the .10 level; **at the .05 level; ***at the .01 level.

Standard errors calculated using numerical differentiation, and the delta method for the various error term standard deviations.

Plotting the three deterministic components of the net schooling benefit across the different schooling levels gives us the graph in Figure 1 (plotted for an individual from the fifth household income decile). The line indicates the height of the stacked components, which represents the total discounted expected value of investing in a particular schooling year (expressed in terms of current period log earnings). Each individual's stochastic shock component $e_{it}^1 - e_{it}^0$ is added to this value, and those for whom this sum is positive then choose to enrol in the next year of school. The higher the net expected value of a specific schooling year, the more likely potential learners will be to enrol in this year of schooling.

Figure 1: Decomposition of expected value of additional schooling year



Source: Author's calculations from Labour Force Survey panel (Statistics South Africa: various years).

In our model, the myopic benefit of additional schooling is initially low, but then gradually increases with additional years of primary and secondary schooling. There are small decreases in the myopic benefit of

additional schooling between schooling years 7 and 8 (as individuals incur the higher estimated cost of secondary schooling) and again between grades 10 and 11 (when individuals turn 16 and start receiving wage offers, albeit very infrequently). The expected myopic benefit of enrolling in school turns negative between years 12 and 13, where the relatively high wage and employment benefits are more than offset by the large costs of tertiary education.

The relatively small myopic net benefit of completing early schooling years is dominated by the larger positive option value associated with the convexity of the schooling-earnings and schooling-employment profiles. By far the most important reason why young black South African men choose to complete primary and early secondary school, despite the relatively low returns associated with these investments, is the promise of higher returns further up the schooling ladder. The fact that the expected net benefit of tertiary education is negative means that individuals attach no convexity option value to schooling years 12 to 15.

During early schooling levels the option value attached to uncertainty regarding future wage offers and schooling utility is quite small. This is because most agents are relatively certain that they will enrol in school again after the current schooling year – given the large value associated with these schooling levels – and therefore attach little value to the benefit of observing next year’s wage offer and schooling utility before choosing whether to drop out of school or not. As the individual approaches completed secondary school and the convexity option value starts shrinking, individuals move closer to the margin where costs exceed benefits, which increases the uncertainty regarding future school investment decisions. This is reflected in the increasing uncertainty option value: individuals now attach more value to the opportunity to observe next period’s schooling utility and wage offers before deciding whether or not to leave the school system in the current period.

6. CONCLUSION

In this paper we have sought to develop a novel decomposition technique that allows the identification of the myopic net benefit of schooling, as well as the option values that arise due to convexity in the earnings profile and uncertainty about future school utility and wage offers. We have then applied this technique to data for young black South African men whose earnings function may be highly convex. Most agents in the model choose to complete primary and early secondary schooling years despite the relatively low returns associated with these investments, because this offers them the option of proceeding on to the high yielding final years of secondary school. Thus the observed schooling decision can be rationalised within the model as a result of the large option values of earlier schooling levels and a schooling cost function that increases steeply between schooling phases.

A. TECHNICAL APPENDIX

The empirical model outlined in section 3 is solved and estimated in MATLAB using a two-step procedure. An inner iteration loop solves the individual's dynamic optimisation problem using backwards induction for an initial parameter vector θ_0 . An outer iteration loop then calculates the likelihood function associated with this solution, and uses numerical optimisation to find the parameter vector that maximises this likelihood. These two steps are discussed in more detail below. Before solving or estimating the model the per capita household income of the sample is discretised by assigning to each individual the income of the individual at the midpoint of his income decile. Next, R error terms are drawn for school utility, wages and employment from a standard normal distribution for each combination of the state values, and these same error terms are used to calculate the model solution for each of the outer iteration loops. The coefficients in Table 1 were estimated using $R = 1,000$ error draws. The state vector consists of household income $h_i \in \{h_1, \dots, h_{10}\}$, age $a_{it} \in \{6, \dots, 65\}$, completed schooling years, $s_{it} \in \{0, \dots, 16\}$, and whether or not the individual has entered the labour market $i_{it} \in \{0, 1\}$ ⁸.

The parameter vector θ_0 is used to calculate the expected instantaneous schooling utility $E(u_{it}^1 | \mathbf{q}_{it})$ for ages from 6 to 21⁹, as well as the expected employment status $E(m_{it} | \mathbf{q}_{it})$ and earnings $E(w_{it} | \mathbf{q}_{it})$ for all ages between 16 and 65. The last two matrices are used to calculate the expected instantaneous labour market utility: $E(u_{it}^0 | \mathbf{q}_{it}) = E(m_{it} | \mathbf{q}_{it}) \cdot E(w_{it} | \mathbf{q}_{it})$. Next, the choice-specific integrated value function is calculated using backwards induction and Monte Carlo integration. At the retirement age of $a_{it} = 65$, the integrated value function is simply the expected instantaneous labour market utility $\bar{V}_{it}(\mathbf{q}_{it}) = \bar{V}_{it}^0(\mathbf{q}_{it}) = E(u_{it}^0 | \mathbf{q}_{it})$. Similarly, for values of $a_{it} \in \{22, \dots, 64\}$, the integrated value function is $\bar{V}_{it}(\mathbf{q}_{it}) = \sum_{\tau=t}^{65} \beta^{\tau-t} E(u_{it}^0 | \mathbf{q}_{it})$.

Individuals of school-going age have to decide whether or not to leave school by comparing the choice-specific value functions for school attendance $V_{it}^1(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it})$ and labour market participation $V_{it}^0(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it})$. The model error terms $\boldsymbol{\varepsilon}_{it}$ are calculated by multiplying the standard normally distributed draws by the variance parameters in θ_0 and these are then used to calculate R values of $u_{it}^0(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}) = m_{it}(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}) \cdot w_{it}(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it})$ and $u_{it}^1(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it})$ for each combination of \mathbf{q}_{it} .

The choice-specific value function associated the state vector \mathbf{q}_{it} and the decision to leave school is $V_{it}^0(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}) = u_{it}^0(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}) + \beta \bar{V}_{it}^0(\mathbf{q}_{it})$ whereas that associated with staying in school is $V_{it}^1(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}) = u_{it}^1(\mathbf{q}_{it}, \boldsymbol{\varepsilon}_{it}) + \beta \bar{V}_{i,t+1}(\mathbf{q}_{i,t+1})$, where $\bar{V}_{i,t+1}(\mathbf{q}_{i,t+1})$ is the integrated value function that represents the expected value obtained by making the optimal enrolment decision in future periods.

The integrated value function $\bar{V}_{i,t+1}(\mathbf{q}_{i,t+1})$ is calculated by averaging the maximum value of $V_{i,t+1}^0(\mathbf{q}_{i,t+1}, \boldsymbol{\varepsilon}_{i,t+1})$ and $V_{i,t+1}^1(\mathbf{q}_{i,t+1}, \boldsymbol{\varepsilon}_{i,t+1})$ across all R error draws. At the same time, the proportion of individuals for whom enrolling in school is the preferred decision is also calculated, and used as the estimate for the solution for the conditional choice probability $P(d_{it} = 1 | \mathbf{q}_{it})$. The integrated value function is then used to calculate the values of the choice-specific and integrated value functions for the

⁸ i_{it} takes on a value of 1 if the individual has (permanently) left the school system, otherwise $i_{it} = 0$.

⁹ Our simplifying assumption that schooling cannot be interrupted means that school-going age effectively ranges from 6 to 21.

preceding period t . This process is repeated via backwards induction, starting at age 21 and finishing at the age of first school enrolment, 6.

The model solution is used to calculate various likelihoods that are used during the estimation step. Apart from the conditional choice probability $P(d_{it} = 1 | a_{it}, h_i, i_{i,t-1}, s_{it})$, we also calculate the probabilities of having attained the observed level of schooling $P(s_{it} | a_{it}, h_i, i_{it})$, of being employed $P(m_{it} | a_{it}, h_i, i_{i,t-1}, s_{it})$ and the conditional wage density $f(w_{it} | a_{it}, h_i, i_{i,t-1}, m_{it} = 1)$. The transition of age, household income and labour market entry are all deterministic by assumption, and hence provide no additional information. These likelihoods are then combined to calculate the joint individual likelihood for all the observations and periods, and the likelihood function $L = \prod_i \prod_t L_{it}(a_{it}, h_i, i_{it}, s_{it}, m_{it}, w_{it} | \theta)$. The outer iteration loop uses numerical optimisation techniques to find the parameter vector that maximises this likelihood function. This numerical optimisation was implemented in the iFit MATLAB toolbox (Farhi, 2011; Farhi, Debab, & Willendrup, 2013). Standard errors were using numerical differentiation.

7. BIBLIOGRAPHY

- Appleton, S., Hoddinott, J., & MacKinnon, J. (1996). Education and Health in Sub-Saharan Africa. *Journal of International Development*, 8(3), 307-339.
- Becker, G., S. (1964). *Human Capital*. New York: Columbia University Press.
- Becker, G. S. (1967). Human Capital and the Personal Distribution of Income: An Analytical Approach. Ann Arbor: University of Michigan, Institute of Public Administration.
- Becker, G. S., & Chiswick, B. R. (1966). Education and the Distribution of Earnings. *The American Economic Review*, 56(1/2), 358-369.
- Belzil, C., & Hansen, J. (2002). Unobserved ability and the return to schooling. *Econometrica*, 70(5), 2075-2091.
- Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 3A, pp. 1801-1863). Amsterdam: Elsevier Science.
- Carnoy, M. (1995). Rates of Return to Education *International Encyclopaedia of the Economics of Education* (pp. 364-369). Oxford: Pergamon.
- Farhi, E. (2011). The iFit data analysis library, ILL, computing for science.: ILL, Computing for Science. Retrieved from <http://ifit.mccode.org>
- Farhi, E., Debab, Y., & Willendrup, P. (2013). iFit: A new data analysis framework. Applications for data reduction and optimization of neutron scattering instrument simulations with McStas. *Journal of Neutron Research*.
- Heckman, J. J. (2010). Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy. *Journal of Economic Literature*, 48(2), 356-398.
- Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006). Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond *Handbook of the Economics of Education* (Vol. 1, pp. 307-458). Amsterdam: Elsevier.
- Heckman, J. J., Lochner, L. J., & Todd, P. E. (2008). Earnings Functions and Rates of Return. *Journal of Human Capital*, 2(1), 1-31.
- Keane, M. P., & Wolpin, K. I. (1994). The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence. *The Review of Economics and Statistics*, 76(4), 648-672.
- Keswell, M., & Poswell, L. (2004). Returns to education in South Africa: A retrospective sensitivity analysis of the available evidence. *The South African Journal of Economics*, 72(4), 834 - 860.
- Lemieux, T. (2006). The Mincer Equation Thirty Years after Schooling, Experience, and Earnings. In S. Grossbard-Shechtman (Ed.), *Jacob Mincer, A Pioneer of Modern Labor Economics*: Springer Verlag.
- Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: National Bureau of Economic Research.
- Mincer, J. (1996). Changes in Wage Inequality, 1970-1990 *NBER Working Paper No. 5823*. Cambridge: NBER.
- Nielsen, H. S., & Westergaard-Nielsen, N. (2001). Returns to Schooling in Less Developed Countries: New Evidence from Zambia. *Economic Development and Cultural Change*, 49(2), 365-394.
- Siphambe, H. K. (2000). Rates of return to education in Botswana. *Economics of Education Review*, 19(3), 291-300.
- Teal, F. (2001). Education, Incomes, Poverty and Inequality in Ghana in the 1990's. University of Oxford.
- Vermaak, C. (2010). *Low-wage mobility in the South African labour market*. Paper presented at the Paper presented at DPRU/TIPS Conference, Johannesburg, October 2010.
- Whaba, J. (2000). Returns to Education and Regional Earnings Differentials in Egypt *Discussion Papers in Economics and Econometrics*. Southampton: University of South Hampton.