Using discrete choice dynamic programming to model job search and reservation wages in South Africa

by Cobus Burger

STELLENBOSCH
UNIVERSITY

Abstract

There exists a large literature of Mincerian earnings function estimates that attempt to estimate the causal effect of education and experience on earnings for African countries. Apart from suffering from well-documented endogeneity problems, these models are only very loosely based on economic theory. This paper attempts to contribute to this literature in two ways. Firstly, it models the earnings and employment outcomes as being generated in a labour market populated by heterogeneous but forward-looking expected discounted lifetime utility-maximisers who operate in an environment of uncertainty. Secondly, the individual's choice set is extended to also include whether or not to look for work and which wage offers to accept. This is, to our knowledge, the first paper that attempts to estimate a theory-consistent job search model for an African country.

The paper uses discrete choice dynamic programing techniques to estimate the model parameters for working aged African males using the Labour Force panel dataset for South Africa. Beyond providing earnings function estimates that can be transparently interpreted in an explicit theoretical framework, the model also produces explicit estimates of reservation wage and job search costs faced by the typical South African worker. While the estimated reservation wage compares favourably to that of self-reported reservation wages, the estimated returns to actual experience in our model is far higher than was suggested in the Mincerian earnings function.

1. Introduction

In a seminal article by Mincer (1974), the author showed that education and experience provide us with a useful predictor of earnings. Although the model has since been shown to be reconcilable with certain utility function specifications, this was not the initial focus of the model. The original model was primarily used to describe life cycle earnings and the returns to education.

Since we do not observe market wages, conventional Mincerian earnings function are restricted to the employed. This is problematic, since it has been shown that only including the employed in our regressions analysis may lead to inconsistent population estimates if there are unobservable determinants, which are correlated with both the selection into employment and the offered wage. In the classical example ability is taken to be positively correlated with the selection into employment and the observed wage. If this is the case, and ability is also positivly correlated with education then it can be shown that the sample selection bias will cause the returns to education to be inflated (Heckman 1979).

The Mincerian function not only ignores those who are not employed, it also ignores the job histories of of those who are employed but used to be unemployed. Since we rarely have accurate data on actual working experience, researcher are forced to use potential experience as a proxy for actual experience (where potential experience is derived from subtracting the years of experience and starting age of schooling from the actual age). While these differences between actual and potential experience may be close to negligible in developing countries it is unlikely to be permissible in South Africa, where we have both high unemployment and low participation.

In the rest of this paper we plan to formulate and estimate a model that would is more fitting to the South African labour market landscape.

2. Discrete[1] Choice Dynamic Programming Job Search Model

2.1. Model Structure

Each individual faces a finite decision horizon over their working age life. In our model we assume that an individual can start working at age 16 and not beyond 65, these represent the minimum age at which one can legally be employed and the minimum age at which pension grants can be received, respectively. To bring the model in-line with the data that will be introduced in the second portion of the paper each period will be defined as lasting six months. This leaves each individual with a maximum of 100 discrete decision periods over which to maximize their discounted net present lifetime utility.

$$U_i = \sum_{t=1}^{100} \beta^{t-1} u_{it} \tag{1}$$

We will not be modelling all 100 period for everyone. Nor will w be assuming that all individuals are homogeneous. Unlike some other DCDP models (see Keane & Wolpin, 1997; Belzil & Hanzen, 2002) our paper will regard the education attainment choice as being predetermined and exogenous to our model. Individuals will thus be heterogeneous with regard to their endowment of education. Since some learning still takes place beyond the age of 16 we will only start modelling the career decisions of those individuals that study beyond grade 10 once they are suppose to be done studying[2].

Each period, each individual will choose between three alternatives: staying at home, actively looking for work or working. Formally, the choice set, d_i, that each individual faces, is denoted as follows:

$$d_{it} = \begin{cases} 1\ household\ production \\ 2\ actively\ seeking\ work \\ 3\ working \end{cases}$$

These choices are regarded as being mutually exclusive, in that, for each period each individual can only pursue one of these choices - d_{it} gets 'locked in' for the entiryty of the period.

The instantaneous returns differ depending on what the individual chooses to do that period. To encapsulate the relative costs and benefits associated with each of these alternatives, the net instantaneous utility derived from each decision is captured by the rewards, denoted $r_{d,t}$.

$$u_{it} = I(d_{it} = 1) r_{1,it} + I(d_t = 2) r_{2,it} + I(d_t = 3) r_{3,it} \tag{2}$$

The discounted lifetime utility can be obtained by substituting equation (1) into equation (2):

$$U_i = \sum_{t=1}^{100} \beta^{t-1} \{ I(d_{it} = 1) r_{1,it} + I(d_{it} = 2) r_{2,it} + I(d_{it} = 3) r_{3,it} \} \tag{3}$$

[1] Throughout our paper time is treated as discrete rather than a continues variable - a simplifying assumption that is standard in all discrete choice dynamic programming models

[2] We will assume that everyone finishes schooling within the allotted amount of years and that once someone is done studying they are unable to return to school.

Clearly the utility derived each period will depend on how the returns associated with the chosen activity. We proceed to model these rewards under each of the alternative decisions.

a) Home Production

In our model, we assume that everyone faces the same expected level of utility from home production regardless of how educated or experienced they are. This average level of home production is captured by γ_0. Actual utility is allowed to deviate from this level, by the idiosyncratic error term, ε_{it}^{h} , which is said to capture household productivity shocks. ε_{it}^{h} is assumed to be i.i.d. normally distributed with mean 0 and deviation σ_h.

$$r_{1,it} = \gamma_0 + \varepsilon_{it}^{h} \tag{4}$$

b) Reward for Actively Seeking Work

Individuals occur a cost in searching for work, these cost could be direct (e.g. traveling costs or phoning up potential employees) or indirect (e.g. less time for housework or leisure). In our model the utility derived from spending a period in the job search states is assumed to be the identical across all individuals. Although our model does not strictly assume it, we would expect that the instantaneous utility derived from spending a period searching, γ_1, to be lower than the utility derived from performing household chores, γ_0.

$$r_{2,it} = \gamma_1 \tag{5}$$

c) Reward for Working (Wage)

The current period reward associated with employment follows the conventional Mincerian shape, which combines innate ability with time spent in school and time spent working[3]. Where a_1 and a_2 measure the effect of schooling and a_3 and a_4 represents the effect of experience. The error term, ε_{it}^{w}, can be regarded as an idiosyncratic productivity shock, which is assumed to be i.i.d. normally distributed with mean zero and standard deviation σ_w. The second source of uncertainty is the realised wage draw (u_{it}^{d}) , which is obtained from the job offer draw discussed below.

$$r_{3,it} = w_i = \alpha_0 + a_1 educ_{it} + a_2 educ_{it}^2 + a_3 exp_{it} + a_4 exp_{it}^2 + u_{it}^{d} + \varepsilon_{it}^{w} \tag{6}$$

d) The job offer and the remuneration draw

The remuneration draw is i.i.d. normally distributed with mean zero and standard deviation σ_d.

$$\varepsilon_{it}^{d} \sim N(0, \sigma_d^2)$$

[3]The original Mincerian model is augmented. An education-squared term is introduced to capture the concavity in the returns to education found in South Africa and other developing countries.

McCall (1970) showed that prospective employees will keep on searching (redrawing ε_{it}^{d}) as long as the future benefits of waiting for a better draw outweighs the present cost of searching and the forfeited wage that they could currently be earning. In our model, we will assume that each period an individual has a chance to draw an error which will impact their potential numeration. Individuals are able to get work both through actively searching and through non-searching, but the probability of being offered a job is allowed to differ between the two. While an individual that is currently at home, receives a job offer with probability ϕ_1, an individual who is actively searching for receives a job offer with probability ϕ_2. Although we do not explicitly restrict ϕ_2 to be higher than ϕ_1, we would expect this to be the case - that looking for work increases your likelihood of being offered work.

After the job offer is made, an individual has one period to decide whether they are content and want to take the job. If they decide to take the job offer the remuneration draw gets 'locked in' .

$$u_{it}^{d} = \varepsilon_{it}^{d}$$

This error stays 'locked in' until the employee changes jobs, quits or looses their job. As of yet, our model is unable to distinguish between these three alternatives, consequently we are forced to model them together. Collectively an individual who is employed has a probability of ϕ_3 to forfeit their renumeration draw.

Due to the rich set of assumptions regarding an agent's behaviour, the parameters that underlie a DCDP model can not be cannot be solved with conventional regression methods. In the next session we discuss how the agent's optimisation problem is solved and how the parameters that determine these decisions are estimated.

2.2 Solving the agent's dynamic optimisation problem

Finding the optimal path is complicated by the fact that our current decisions have implications on our future rewards and choices. For instance, if we do not work this period then we do not gain any experience, and decrease the returns associated with future emloyment. Similarly, when looking for a job offer you have to way up the present costs associated with searching for work against the possible gains in future remuneration if you are able to secure a more favourable job offer (McCall, 1970).

In deciding the optimal path agents need to be both aware of the cost and benefits of their current actions on future returns as well as the role of uncertainties. While agents are assumed to be unaware of what their future shocks will be, they are assumed to have perfect knowledge with regard to the probability distribution from which these future shocks will be drawn. In the presence of these shocks what was previously envisioned to be the optimal decision path might turn out to be an inferior strategy once an array of shocks are realised. Due to these complexities econometricians make use of the discrete choice dynamic programming (DCDP) models. These models have the desirable attribute of being both complicated and flexible at the same time. Complicated enough to allow us to make an optimal choice while considering future shocks that could occur, their likelihood of occurring and how they impact our future returns, but flexible enough to allow our agent to renege on his optimal path when shocks get revealed in the future.

In its current form ε_{it}^{h}, ε_{it}^{w} and ε_{it}^{d} all follow normal distributions. In order to allow for all the possible values that each of these errors can take on we are required to solve an integral. Follow an approximation method developed by Keane and Wolpin (1994), we proceed to approximate the continuous values under the normal distribution with a set of discrete values, with a probabilities that mirrors that of the normal distribution.

In order to keep track of the sequential decisions faced by the individual we introduce the state space, denoted as Ω_t, that contains all the information from the past that could influence the current-period decision d_t. The current state $\Omega_t = (educ_t, exp_{t-1}, \varepsilon_{t-1}^{d})$ is determined by a combination of current endowments (experience and education) as well as the last remuneration draw[4].

When considering what to do individuals are faced with the problem of maximising lifetime utility. Defining $V(\Omega_a, a)$ as the value function as the maximum expected present value of lifetime utility at period a, given the state space Ω_a and the discount factor β.

$$V(\Omega_a, a) = \max_{\{d\}} E\left[\sum_{t=a}^{100} \beta^{t-1}\left[I(d_t = 1)r_{1,t} + I(d_t = 2)r_{2,t} + I(d_t = 3)r_{3,t}\right]|\Omega_t\right] \qquad (7)$$

In its current form the evolution of the decisions faced by our agents follows that the specifications of a finite Bellman equation[5]. Consequently, we are able to rewrite our value function as being equal to the highest of the 3 alternative value functions coinciding with the three choices contained in our decision space.

$$V(\Omega_a, a) = \max_{\{d\}}[V_1(\Omega_a, a), V_2(\Omega_a, a), V_3(\Omega_a, a)] \qquad (8)$$

Since each of these value functions (V_1, V_2 and V_3) also obey the Bellman principles, they can each be rewritten as as the sum of the instantaneous returns plus the discounted sum of another value function from the next period. The state space faced in the following period differs depending on what decision we follow today (whether $d_a = 1,2$ or 3).

$$V_i(\Omega_a, a) = r_{i,a} + \beta E[V(\Omega_{a+1}, a+1)|\Omega_a, d_a = i] \qquad (9)$$

Following this approach we can expand our initial value function all the way to the last period. At $t = 100$, all errors are observable and no uncertainty remains. Consequently, the optimal decision under each possible state can be solved. Once we have knowledge regarding what our optimal decision will be taken in period $t = 100$ under each state, our agent is able to move one step back and solve their optimal decisions under all possible states in the second last period, $t = 99$. Following this backward-recursion approach we are able to the optimal decision path under each state, once we have done so for the period that follows. Walking back from $t = 100$ to $t = a$, we are able to calculate the value function $V(\Omega_a, a)$ and more importantly, the optimal decision path at period a.

[4] Past realisations of the two productivity shocks (ε_t^{w} and ε_t^{h}) are not included since their realisaion are assumed to be serialy independent.

[5] See Rust (1992) for a complete discussion of the principles that need to be satisfied in order to use a Finite Markov Bellman model

2.3 Estimating the Structural Parameters

Having solved the optimal decision of process of each agent model, we now proceed to find the set of parameters, Θ, under which the decision made by our optimizing agents would best fit that of the decisions we have data on. In order to solve this model we first need to set up a likelihood function that allows us to measure how confident we are in any set of parameters. Once the likelihood function is properly specified, we use numerical optimization to find the set of parameters that are most likely to have produced the outcomes observed in the South African labour market.

Assuming that we have observable data on an individual's education and age, as well as whether their decision in period t, and their wage we can back out the following likelihood function.

$$L_i(\Theta) = \prod_{\mathrm{t=t_1}}^{\mathrm{t_2}} L_{it}(\Theta)$$

$$= \prod_{\mathrm{t=t_1}}^{\mathrm{t_2}} [\mathrm{P}(d_t = 1|\Psi_t)]^{1(\mathrm{d_t=1})} \, [\mathrm{P}(d_t = 2|\Psi_t)]^{1(\mathrm{d_t=2})} [\mathrm{P}(w_t|d_t = 3, \Psi_t)\mathrm{P}(d_t = 3|\Psi_t)]^{1(\mathrm{d_t=3})} \quad (10)$$

where $\Psi_t = \begin{cases} educ_i, age_{it} & \text{if this is the } 1^{\mathrm{st}}\text{observation of individual } i \\ educ_i, age_{it}, d_{it-1}, w_{it-1} & \text{if this is not the } 1^{\mathrm{st}}\text{observation of individual } i \end{cases}$

From the above equation it is clear that our likelihood function not only encapsulates the likelihood of participation and employment, but also the wage conditional on being employed. For a model to do well it needs to accurately predict which decision state we are likely to observe an individual in and which wage bin they will fall in if they are employed.

In making this predictions we use the underlying parameters of our model. If we have observed the individual in an earlier wave we also use their previous observed decision states and wage. All the choices and actions that predates the previous period is irrelevant to our model since our model evolves in Markovian manner. Consequently, we only need to condition on observables from one period back and can ignore everything before that. This gives us the likelihood of having observed the decisions based on an individual's current characteristics (age and education) and their previous observable choice (d_{it-1}) as well as their previous wage, had they worked the previous period.

In solving the parameters researchers often prefer to work with the logged likelihood rather than the likelihood. The logged likelihood function (equation 12) is obtained by summing the separate likelihood functions rather than multiplying (equation 11) them as one would have done if we were working with likelihoods.

$$L_i(\Theta) = L_{i,t_2}(\Theta) \times \ldots \times L_{i,t_1+1}(\Theta) \times L_{i,t_1}(\Theta) \quad (11)$$

$$l_i(\Theta) = l_{it_2}(\Theta) + \cdots + l_{i,t_1+1}(\Theta) + l_{i,t_1}(\Theta) \quad (12)$$

3. Empirical Analysis

3.1. Data

Data is taken from the six consecutive bi-annual Labour Force Surveys; the first was captured in the second semester of 2001 and last was captured in the first semester of 2004. As mentioned before, we limit our sample to that of working aged African males. Using the rotating panel within the surveys we were able to identify 28216 unique individuals. Although 22902 of these individuals are observed more than once over the six waves, only 431 individuals were observed in all six waves. For each of the periods that these individuals were surveyed, we obtained their level of schooling, their age, as well as information regarding employment status: whether they were currently working, seeking work or inactive. We also obtained the wages of those who were said to be currently working.

3.2 Descriptive Statistics

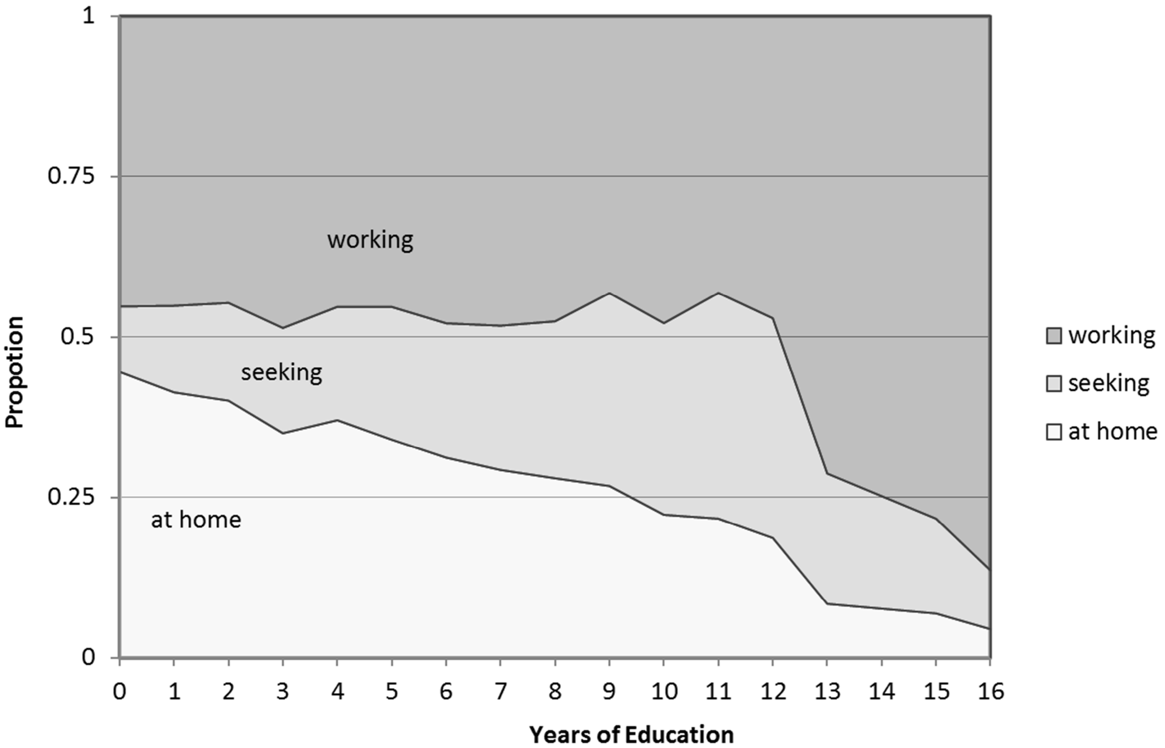
Although our model does not attempt to explain educational attainment, it is important to be aware of the fact that a large portion of working aged individuals are still in school. According to figure A1 in the appendix, school attendance is especially high for those aged 20 and younger, but drops of quickly. Beyond the age of age 26 almost no-one is in school. Our model ignores these individuals that are full time students.

The choice between working, staying at home or actively seeking employment is central to our discussion throughout. The following graph shows how the likelihood of each state differed by age.

Figure 2: Choice Distribution by Age, for those who are not currently studying

Among those who are not studying and younger than 20 less than 10% appear to be working. Most are either inactive or searching for work. If we limit our sample to only those aged 40 and younger, we observe that individuals are more likely to be employed the older they are. Similarly, the likelihood of being at home decreases from 70% to 20% over the same sample. Interestingly, the likelihood of being actively looking for work stays relatively stable, round 20% over the sampled age range. From 40 onwards we see a sharp decline in both the search for work, and the likelihood of working. It is unclear whether the decrease is driven by retirement or by discouraged job seeking.

Figure 3: Choice Distribution by Educational Attainment, for those who are not currently studying



Those individuals with higher education are more likely to be employed and less likely to be busy with household production. The likelihood to be actively searching for employed seems to be the highest among those individuals with moderate levels of education.

In our formal model we are interested in the level of churning between the three decision states and the factors that give raise to these changes. The following table shows the transition between the three states in two consecutives periods for the subsample of individuals who were observed in two subsequent periods[6].

Table 1: Transition between Decision States

		Future State			
		At home	Seeking	working	Total
Current State	At home	64% *(5222)*	24% *(1928)*	13% *(1031)*	100% *(8181)*
	Seeking	26% *(2090)*	53% *(4283)*	21% *(1666)*	100% *(8039)*
	Working	7% *(1169)*	10% *(1512)*	83% *(13044)*	100% *(15725)*
	Total	27% *(8481)*	24% *(7723)*	49% *(15741)*	100% *(31945)*

The table suggests that individuals are most likely to stay in the state they are in. Compared to those who are not actively searching, those individuals who are looking for work are more likely to be employed. The difference is however not as large as one would have expected. Those who are actively seeking work have a 21% chance of being employed the next period, while those who are not looking have a 13% chance of finding work. From these discriptives it appears as though actively seeking labour only increases once likelihood of being employed in the future by 50%. 83% of those who are currently employed will be employed in the future. Of the 17% who stopped working, 7% do not look for work the next period, and 10% seek employment.

[6] Only those individuals for whom we had consecutive observations are included in this table. Some people were not observed in consecutive periods, while others were observed multiple times.

3.3 Fitting the Data to the Model

Unfortunately our data exhibits a high rate of attrition. The table below gives an idea of the level of attrition and balance in our data.

Table 2: Amount of Observed Waves per Individual

Waves	Frequency	Proportion
1	5314	18.8%
2	10520	37.3%
3	6616	23.4%
4	3539	12.5%
5	1796	6.4%
6	431	1.5%
total	28216	100%

Since not every individual is observed in every period, we are unable to use the likelihood function that we specified earlier, which was designed for a balanced panel with no missing values. Ignoring those individuals that we do not observe for all six periods will not only compromise our sample size, but could also aggravate attrition bias if drop out is non-random. Consequently, we expand our model so as to allow us to also incorporate these individuals for which we have missing values. The model is augmented in two ways. Firstly, we will assume that $L_{i,t}(\Theta) = 1$ if missing and secondly since observations can go unobserved for up to four periods before being observed again, we need the model to have a five period memory.

To allow for these missing values we expand our model so that Ψ_t has a τ-period memory, where we allow the value τ to go up to five, but to depend on how long ago a specific individual was observed.

$$L_i(\Theta) = \prod_{\mathrm{t=t_1}}^{\mathrm{t_2}} \left[\mathrm{P}\left(d_t = 1|\Psi_{t,\tau}\right)\right]^{1(\mathrm{d_t}=1)} \left[\mathrm{P}\left(d_t = 2|\Psi_{t,\tau}\right)\right]^{1(\mathrm{d_t}=2)} \left[\mathrm{P}\left(w_t|d_t = 3, \Psi_{t,\tau}\right)\mathrm{P}\left(d_t = 3|\Psi_{t,\tau}\right)\right]^{1(\mathrm{d_t}=3)} \quad (13)$$

$$\text{where } \Psi_{t,\tau} = \begin{cases} educ_i, age_{it} & \text{if this is the } 1^{\text{st}} \text{observation of individual } i \\ educ_i, age_{it}, d_{it-\tau}, w_{it-\tau} & \text{if this is not the } 1^{\text{st}} \text{observation of individual } i \end{cases}$$

Using this notation we are able to derive the conditional probability of observing d_t and w_t, even if we did not observe an individual in the previous period. For instance if we only observed person i in period 1 and then again in period 6 then the likelihood function would look as follows:

$$L_i(\Theta) = L_{i,t_2}(\Theta) \times 1 \times 1 \times 1 \times 1 \times L_{i,t_1}(\Theta).$$

In our model we proceed to simulate the decision process of 1,7 million (a 100 000 thousand individuals for each education level) over their entire working age life. Using the likelihood approach proposed above we compare the outcomes of our simulated model to those of the actual surveyed data, to see how believable our parameters are. The optimal parameters were derived using the Powell searching algorithm[7] (Brent, 1973).

[7] The matlab file was written by Argimiro R. Secchi and can be downloaded from: http://www.mathworks.com/matlabcentral/fileexchange/15072-unconstrained-optimization-using-powell/content/powell.m

3.4 Results

The following table contains the set of parameters that best fit our observed data, as well as the set of parameters that solved the Mincerian earnings function:

Table 3: Comparing the parameters obtained from the Static and Dynamic Models

	(OLS) Static Model	**(DCDP) Dynamic Model**
Constant	-0.418	0.996
	(0.032)	*(0.000)*
Education	-0.029	-0.177
	(0.005)	*(0.000)*
Education Squared x 100	1.329	3.654
	(0.033)	*(0.000)*
(Potential) Experience	0.084	0.122
	(0.002)	*(0.000)*
(Potential) Experience Squared x 100	-0.104	-0.350
	(0.003)	*(0.000)*
Remuneration Draw		0.360
		(0.000)
Employment Productivity Shock		0.501
		(0.000)
Household Production (γ_0)		11.375
		(0.000)
Cost of Searching (γ_1)		6.838
		(0.000)
Household Productivity Shock (δ_h)		7.427
		(0.000)
Probability of Job Offer if $d_i = 1$ (ϕ_1)		0.106
		(0.000)
Probability of Job Offer if $d_i = 2$ (ϕ_2)		0.482
		(0.000)
Probability of Job Offer if $d_i = 3$ (ϕ_3)		0.167
		(0.000)
N	54227	54227
R-squared	0.3024	
log likelihood		-162640

In both models the choice of education is assumed to be exogenous. The returns appear to be similar at low levels of education. Both models show low returns to education before grade 7[8] (see figure A1 in the appendix). From grade 8 onwards both models show increasing returns to education. The returns to higher levels of education appear to be 2 times higher in our dynamic model. This begs the question as to why so few people study beyond grade 12, a question we are unable to solve within the scope of our current paper, since we are not modelling education, but treating it as exogenous.

[8] The negative returns could be driven by the quadratic form of our returns function.

The returns to experience is measured differently in the two models. Not only with regard to how experience is defined, but also with regard to how the outcome, wage, over which we are measuring returns is defined. In the static model, potential experience is measured as the years of experience one would have had had one started studying at age 6, finished one's studies on time and worked ever since and it measured on observed wages. In the dynamic model experience measures the effect of actual years working (that we do not observe in our sampled data, but do observe in our simulated data) on underlying wage that individuals are offered. As predicted, the effect of actual experience in the dynamic model is larger than the effect of potential experience in the static model (see figure A2 in the appendix). The returns, however, die off earlier.

The average shadow price of household production (commonly referred to as the reservation wage) is estimated at R2201 per month[9]. This value compares favouribly with the estimated reservations wages in the *National Income and Dynamics Study.* Among a simliarly aged group of African males, the median self-reported reservation wage was found to be R2000 per month. The average shadow price for actively searching for labour was estimated at R1323 per month[10]. By comparing the the two shadow prices we are able to derive the disutility of searching relative to full time home production, which is valued at roughly R875 per month.

Our parameters suggest that the likelihood of being offered employment is larger when one is actively looking for employment. In our model, those who were actively seeking work had a 48% chance of being offered employment, while those who were residing at home had a 11% chance of being offered employment. The gap is much larger than the suggested differences in employment likelihood found in our transition matrix (see table 1), where 21% of those who are actively seeking work found work and 13% of those who were not looking for work found work. These discrepencies suggest that a lot more job offers are being turned down by individiduals who are actively seeking better offers than by people who are in household production.

5. Conclusion

In this paper we attempted to formulate a structural model that explains the decisions process faced by utility maximizing working aged agents in the face of uncertainty regarding future shocks. Using dynamic programming the parameter estimated were obtained that best what we observed among our subsample of working aged African men in Labour Force Survey.

We found that the augmented job search model gives findings that differ substantially from those proposed by the conventional Mincerian approach. Most notably the returns to higher levels education and lower levels experience appear to be higher in our dynamic model. Our model further suggest that there is a cost to searching, but that individuals are willing to forego this utility in the short run, since they are almost 3 times more likely to receive a job offer when they are actively looking for work.

---

[9] Transformed: 11.375×4.3×45 = 2201.06. *(Assuming that the average person works 45 hours a week)*

[10] Transformed: 6.838 ×4.3×45 = 1323.15

References

BELZIL, C., & HANSEN, J. 2002. "Unobserved Ability and the Return to Schooling". *Econometrica, 70*(5), 2075-2091.

BELZIL, C. 2007. "The Return to Schooling in Structural Dynamic Models: A Survey". *European Economic Review, 51*(5), 1059-1105.

BRENT, *Algorithms for minimization without derivatives*, Prentice-Hall (1973)

HECKMAN, J., 1979. Sample Selection Bias as a Specification Error. *Econometrica.* 46, 1: 153-161

KEANE, M. P., WOLPIN, K. I. 1994. The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence. *Review of Economics and Statistics*. 76(Nov):648-672.

KEANE, M. P., WOLPIN, K. I. 1997. The Career Decisions of Young Men. *Journal of Political Economy*. 105(3):473-522.

MCCALL, J. J. 1970. Economics of Information and Job Search. *The Quarterly Journal of Economics*. 84(1):113-126.

MINCER, J. 1974. Schooling, Experience and Earnings. New York: National Bureau of Economic Research

RUST, J. 1992. Do People Behave According to Bellman's Principles of Optimallity? Stanford, CA: Hoover Institution Working Paper no. E-92-10.

Appendix

Figure A1: Choice to Study by Age

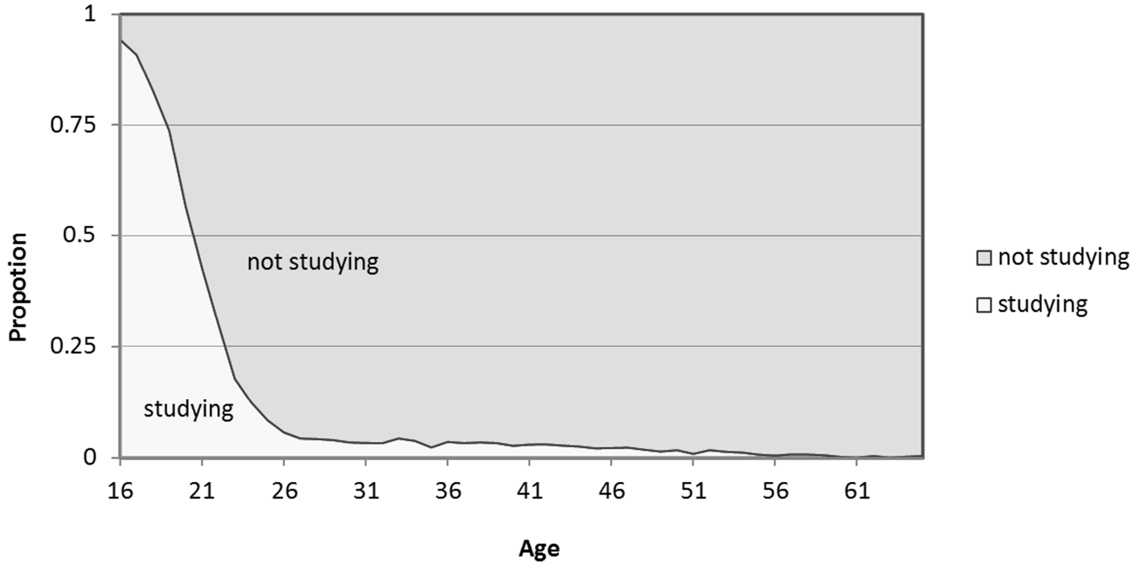


Figure A2: Estimated Returns to Education

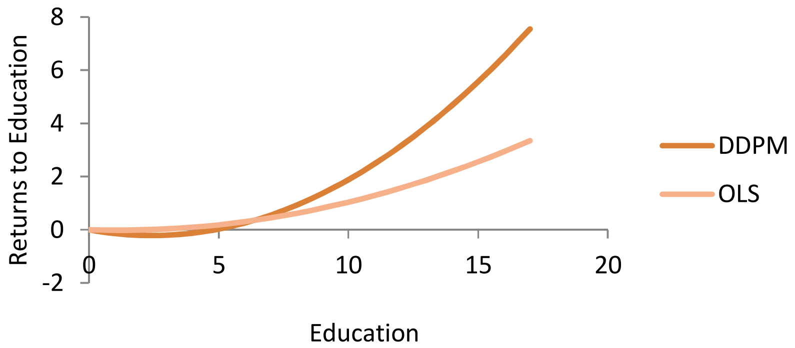


Figure A3: Estimated Returns to Experience