The Career Decisions of Young Men, JPE, 1997

Wenzhi Wang [*]

March 17, 2023

Introduction

1 A Basic Human Capital Model

Each individual has a finite decision horizon beginning at age 16 and ending at age A. At each age a, an individual chooses among five mutually exclusive and exhaustive alternatives: work in either a blue- or white- collar occupation, work in the military, attend school, or engage in home production. Let $d_m(a) = 1$ if alternative m is chosen $(m = 1,2,3,4,5)$ at age a and zero otherwise. The reward per period at any age a is given by

$$R(a) = \sum_{m=1}^{5} R_m(a) d_m(a), \tag{1}$$

where $R_m(a)$ is the reward per period associated with the mth alternative. These rewards contain all the benefits and costs associated with each alternative.

1.1 Working Alternatives $(d_m(a) = 1; m = 1,2,3)$

The current-period reward for working in occupation m is the wage $w_m(a)$. An individual's wage in an occupation is the product of the occupation-specific market (equilibrium) rental price (r_m) times the number of occupation-specific skill units possessed by the individual, $e_m(a)$. The latter will depend on the technology of skill production. In a standard human capital formulation, the level of skill accumulated up to any age in an occupation depends on the number of years of schooling completed, $g(a)$, and on work experience in that occupation, $x_m(a)$, which typically takes a quadratic form. Letting $e_m(a)$ be the number of skill unites possessed ate age a, $e_m(16)$ the skill "endowment" at age 16, and $\varepsilon_m(a)$ a skill technology shock, we get

$$e_m(a) = \exp\{e_m(16) + e_{m1} g(a) + e_{m2} x_m(a) - e_{m3} x_m^2(a) + \varepsilon_m(a)\} \tag{2}$$

where $m = 1,2,3; a = 16,\ldots,A$. This specification leads to a standard ln wage equation in which the constant term is $\ln(r_m) + e_m(16)$, the sum of the ln rental price and the age 16 skill endowment. Notice that the exponential form implies that the higher the endowment, the more skill units are "produced" per additional year of schooling or work experience.

[*]This note is written down during my M.phil. period at the University of Oxford.

1.2 Nonwork Alternatives: Attending School ($d_4(a) = 1$) or Remaining at Home ($d_5(a) = 1$)

In terms of the current-period reward for school attendance $R_4(a)$, we allow for an indirect cost of schooling associated with effort. Specifically, it is the effort cost at age a, which has a component that is a fixed endowment at age 16 ($e_4(16)$) and a component that fluctuates randomly with age ($\varepsilon_4(a)$) minus direct schooling costs of attending college (tc_1)or of attending graduate school (tc_2). Although adding an effort cost implies that $R(a)$ is interpreted as utility, given the additive nature of rewards in Equation 1, effort cost is denominated in wage units (dollars). Regarding the choice of staying at home, we allow for home production (leisure). The reward for remaining home $R_5(a)$, which is also denominated in dollars, consists of the value of a fixed skill endowment at age 16 ($e_4(16)$) and a component that fluctuates randomly with age ($\varepsilon_5(a)$).

Thus, the structure of rewards is given by

$$\begin{aligned} R_m(a) &= w_m(a) = r_m \exp\left[e_m(16) + e_{m1}g(a) + e_{m2}x_m(a) - e_{m3}x_m^2(a) + \epsilon_m(a)\right], \quad m = 1,2,3, \\ R_4(a) &= e_4(16) - tc_1 \cdot I[g(a) \geq 12] - tc_2 \cdot I[g(a) \geq 16] + \epsilon_4(a), \\ R_5(a) &= e_5(16) + \epsilon_5(a). \end{aligned} \tag{3}$$

To close the model, productivity shocks are assumed to be joint normal, $N(0,\Omega)$, and serially uncorrelated. Initial conditions are the given level of schooling at age 16, $g(16)$, and the accumulated work experience ate age 16 in each occupation, assumed to be zero $x_m(16) = 0$. It is convenient to define the age 16 endowment vector $\mathbf{e}(16) = \{e_1(16), e_2(16), e_3(16), e_4(16), e_5(16)\}$, the work experience vector $\mathbf{x}(16) = \{x_1(16), x_2(16), x_3(16), x_4(16), x_5(16)\}$, and the technology shock vector $\varepsilon(16) = \{\varepsilon_1(16), \varepsilon_2(16), \varepsilon_3(16), \varepsilon_4(16), \varepsilon_5(16)\}$. Further, we denote $\mathbf{S}(a) = \{\mathbf{e}(16), g(a), \mathbf{x}(a), \varepsilon(a)\}$.

At any age the individual's objective is to maximize the expected present value of remaining lifetime rewards. Defining $V(\mathbf{S}(a), a)$, the value function, to be the maximal expected present value of lifetime rewards at age a given the individual's state $\mathbf{S}(a)$ and discount factor δ, we get

$$V(\mathbf{S}(a), a) = \max_{d_m(a)} E\left[\sum_{T=a}^{A} \delta^{T-a} \sum_{m=1}^{5} R_m(a) d_m(a) \mid \mathbf{S}(a)\right]. \tag{4}$$

Note that $\mathbf{S}(a)$ contains the relevant history of choices that enter the current-period rewards, the endowment vector, and the realizations of all shocks at a, $\varepsilon_m(a)$ for $m = 1,...,5$. In addition, the individual knows all relevant prices and functions (occupation-specific rental prices, the reward functions, the skill technology functions, direct schooling costs, and the distribution of shocks). The maximization in Problem 4 is achieved by choice of the optimal sequence of control variables $\{d_m(a) : m = 1,...,5\}$ for $a = 16,..., A$. The value function can be written as the maximum over alternative-specific value functions, each of which obeys the Bellman equation:

$$V(\mathbf{S}(a), a) = \max_{m \in M}\{V_m(\mathbf{S}(a), a)\}, \tag{5}$$

where $V(\mathbf{S}(a), a)$, the alternative-specific value functions, are given by

$$\begin{gathered} V_m(\mathbf{S}(a), a) = R_m(\mathbf{S}(a), a) + \delta E\left[V(\mathbf{S}(a+1), a+1) \mid \mathbf{S}(a), d_m(a) = 1\right], \quad a < A, \\ V_m(\mathbf{S}(A), A) = R_m(\mathbf{S}(A), A). \end{gathered} \tag{6}$$

The expectation in Equation 6 is taken over the distribution of the random components of $\mathbf{S}(a+1)$ conditional on $\mathbf{S}(a)$, that is, over the unconditional distribution of $\varepsilon(a+1)$ given serial independence. The predetermined state variables such as schooling occupation-specific work experience evolve in a Markovian manner that is conditionally independent of the shocks $x_m(a+1) = x_m(a) + d_m(a)$ ($m = 1,2,3$) in the case of occupation-specific work experience and $g(a+1) = g(a) + d_4(a)$ in the case of schooling ($g(a) \leq \bar{G}$), where $\bar{G}$ is the highest attainable level of schooling.

The individual's decision process is described as follows: beginning at age 16, given $\mathbf{e}(16)$ and $g(16)$, the individual draws five random shocks from the joint $\varepsilon(16)$ distribution, uses them to calculate the realized current rewards and thus the (five) alternative-specific value functions, and chooses the alternative that yields the highest value. The state space is then updated according to the alternative chosen. The process is repeated. The solution of the optimization problem at each age a can be represented by the set of regions in the five-dimensional $\varepsilon(a)$ space over which each of the alternatives would be optimal, that is, would have the highest alternative-specific value function.

The solution of the optimization problem serves as the input into estimating the parameters of the model given data on choices and possibly some of the rewards. Although the solution is deterministic for the individual, it is probabilistic from our view because we do not observe the contemporaneous shocks, that is, $\varepsilon(a)$.

Consider having data on a sample of individuals from the same birth cohort who are assumed to be solving the model described above and for whom choices are observed over at least a part of their lifetimes. In addition, assume, as is the case, that wages are observed only in the periods in which market work is chosen and only for the occupation that is chosen. Thus, for each individual, $n = 1,,,,,N$, the data consists of the set of choices and rewards $\{d_{nm}(a), w_{nm}(a)d_{nm}(a) : m = 1,2,3\}$ and $\{d_{nm}(a) : m = 4,5\}$ for all ages in the given range $[16,\bar{a}]$. Let $c(a)$ denote the choice-reward combination at age a and let $\bar{\mathbf{S}}(a) = \{\mathbf{e}(16), g(a), \mathbf{x}(a)\}$ denote the predetermined components of the state space, that is, $\mathbf{S}(a)$ net of the technology shocks. Serial independence of the shocks implies that the probability of any sequence of choices and rewards can be written as follows:

$$\Pr[c(16),\ldots,c(\bar{a}) \mid g(16),\mathbf{e}(16)] = \prod_{a=16}^{\bar{a}} \Pr[c(a) \mid \overline{\mathbf{S}}(a)] \tag{7}$$

The sample likelihood is the product of the probabilities in Equation 7 over the N individuals. The solution to the individual's optimization problem provides the choice probabilities that appear on the RHS of Equation 7. For example, the probability that an individual chooses to attend school at age a is $\Pr\{V_4(\mathbf{S}(a),a) = \max_m V_m(\mathbf{S}(a),a)\}$, which can be viewed, for the purpose of estimation, as a function of the parameters of the model conditional on the data. Estimation involves an iterative process: solving numerically the dynamic programming problem for given parameter values and then computing the likelihood function, and so forth, until the likelihood is maximized. The likelihood function involves the calculation of multivariate integrals as in general multinomial choice problems.

The likelihood function using 7 applies to a sample that is homogeneous except for initial schooling. Clearly, the human capital investment process does not begin at age 16. The vector of age 16 endowments depends on prior human capital accumulation as well as innate talents. To allow for the possibility that individuals do not have identical age 16 endowment vectors, we define a type k individual, $k = 1,...,K$, by an endowment vector $\mathbf{e}_k(16) = \{e_{mk(16)} : m = 1,...,5\}$. Thus individuals may have comparative advantages among the different alternatives, including in acquiring schooling and in home production, that are known to them. Thus each type solves the optimization problems with different initial (age 16) conditions.

We assume that while endowment heterogeneity is unobserved by us, we do know there to be K types. Denote π_k as the proportion of the kth type in the population. In this case, the likelihood function is a mixture of the type-specific likelihoods, $\Pi_{n=1}^{N}(\sum_{k=1}^{K}\pi_k L_{nk})$, where L_{nk} is the likelihood of person n's observed choice sequence and rewards if person n is of type k, and the parameter vector is augmented to include the endowment vectors for the K types and the type probabilities.

It is unlikely that initial schooling (at age 16) is exogenous, in which case conditioning the likelihood on it as though it were non-stochastic is problematic. One remedy would be to specify the optimization problem back to the age at which initial schooling was zero for everyone (say, age 3) and solve for the correct probability distribution of attained schooling at age 16 (conditional on age 3 endowments). However, such a model would have to focus on parental decision making with respect to investments in children (including fertility decisions) and would be very demanding in many dimensions (modelling, computation, and data). The alternative we follow is to assume that initial schooling is exogenous conditional on the age 16 endowment vector. The likelihood contribution for the nth individual is

$$\Pr[c_n(16),\ldots,c_n(\bar{a}) \mid g_n(16)] = \sum_{k=1}^{K}\prod_{a=16}^{\bar{a}}\pi_{k|g_n(16)}\Pr[c_n(a) \mid g_n(16),\ \text{type} = k]. \tag{8}$$

Note that the type proportions, treated as estimable parameters, are now conditioned on initial schooling, $g(16)$. Note that here we do not need to include the endowment vector $\mathbf{e}(16)$ since it is fully characterized by an individual's type k.

2 Data