Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA

Model and Identification

Attanasio (2011) paper presents three models. The first one is the Education Choice model, which is used to identify the effect of the PROGRESA grant on school participation. This model is also the main structural model of this paper. The other two models are two reduced form equations to model the potential wage and initial education attainment of the individuals in the sample because these two variables are unobservable from the data.

Compared to the original model in the Attanasio (2011) paper, the model I presented here is a simplified version while at the same time keep the essence of the original one. The simplification mainly involves changing a T periods dynamic problem into a two periods one and substituting the vector covariates into a scalar one. Meanwhile, I revise the notation of the original paper to make it compatible with the simplified model. Moreover, some of the words and descriptions in the original paper are replaced with equations and formal assumptions.

1 Main Model: Education Choice Model

In this simplified version of education choice model, there are only two periods. In period one, students choose either go to school or to work and receive an instantaneous utility. In period two, they reap a terminal utility defined by the terminal value function. Assume that all students will choose in the way to maximize their total present value of the utility.

1.1 Period one: Instantaneous utilities from schooling and work

Let u_i^s represent the instantaneous utilities student i gets if he/she chooses to go to school in period one. Then we have:

$$u_i^s = Y_i^s + \alpha g_i,$$

$$Y_i^s = \mu_i^s + \beta_1^s z_i + \beta_2^s \text{ed}_i + 1(p_i = 1)\beta_3 x_i^p + 1(s_i = 1)\beta_4 x_i^s + \epsilon_i^s, \ (1)$$

where g_i is the amount of the grant an individual is entitled to; it will be equal to zero for non-eligible individuals and control localities. Y_i^s represents the individual's remaining utility from attending school. z_i is a scalar of taste shifter variables. The variable $1(p_i = 1)$ denotes attendance in primary school, while the variable $1(s_i = 1)$ denotes attendance in secondary school. x_i^p and x_i^s represent factors affecting the costs of attending primary school and secondary school, respectively. The term ϵ_i^s represents the preference shock. Each individual knows his/her own preference shock but this shock is unobservable to econometricians. ed_i represents individuals' current education attainment and this variable is unobservable from the data. μ_i^s represents individual's unobserved heterogeneity that could influence both the current and past schooling choice of individuals.

Similarly, the instantaneous utility of going to work is denoted by u_i^w, with the following expressions:

$$u_i^{\text{w}} = Y_i^{\text{w}} + \delta \omega_i,$$

$$Y_i^{\text{w}} = \mu_i^{\text{w}} + \beta_1^w z_i + \beta_2^{\text{w}} \text{ed}_i + \epsilon_i^{\text{w}} \ , \ (2)$$

where ω_i are the potential wage when individual i is out of school. The wage, especially the wage for those students attending the school is not observable, so it has to be estimated.

Since the numerical value of the utility does not matter, so without loss of generality, set $Y_i^w = 0$, then we can rewrite equations (1) and (2) as follows:

$$u_i^{\text{s}} = \gamma \delta g_i + \mu_i + \beta_1 z_i + \beta_2 \text{ed}_i + 1(p_i = 1)\beta_3 x_{it}^{\text{p}} + 1(s_i = 1)\beta_4 x_{it}^{\text{s}} + \epsilon_i, (3)$$
$$u_i^{\text{w}} = \delta \omega_i, (4)$$

,where $\beta_1 = \beta_1^s - \beta_1^w, \beta_2 = \beta_2^s - \beta_2^w, \gamma = \alpha/\delta, \mu_i = \mu_i^s - \mu_i^w$, and $\epsilon_i = \epsilon_i^s - \epsilon_{it}^w$. ϵ_i is the difference between two random preference shock and we assume it is distributed as a standard logistic across the population. The model also assumes that μ_i is a discrete random variable whose points of support and probabilities will be estimated empirically. The coefficient γ measures the impact of the grant as a proportion of the impact of the wage on the education decision. If $\gamma = 1$, the effect of the grant on utility and therefore on schooling choices would be the same as that of the wage.

1.2   Period Two: Terminal Value Function and Total Utility

In period two, the dynamic problem ends and students reap a terminal utility defined by the terminal value function $V(.)$. The paper models terminal value function in the following way:

$$V(\text{ed}_{i,2}) = \frac{\alpha_1}{1 + exp(-\alpha_2 * \text{ed}_{i,2})},$$

where $\text{ed}_{i,2}$ is the education level achieved by period 2. α_1 and α_2 are parameters and are constrained to be non-negative.

If a student chooses to go to school in period one, he/she may not be successful in completing the grade and the level of education does not increase in this case. This model assumes that the probability of failing a grade is exogenous and is only correlated with the grade to be completed and the age of the individual. To simplify, I denote the probability of individual i completing the next grade by p_i. Therefore, the total utility V_i^s for student i who chooses to go to school in period one is:

$$V_i^s = u_i^s + r[p_i V(ed_i + 1) + (1 - p_i)V(ed_i)]$$

Here r is the discount rate and $V(.)$ is the terminal function defined above. r is also a parameter that should be estimated.

The total utility value for student i who chooses to work in period one is:

$$V_i^w = u_i^w + rV(ed_i)$$

Let D_i represent the choice of individual i in period one, $D_i = 1$ if the individual chooses to go to school and equals to 0 if the individual chooses to work. Then we have:

$$\begin{aligned}
D_i &= 1[V_i^s > V_i^w] \\
&= 1[u_i^s + r[p_i V(ed_i + 1) + (1 - p_i)V(ed_i)] > u_i^w + rV(ed_i)] \\
&= 1[(u_i^s - u_i^w) + rp_i(V(ed_i + 1) - V(ed_i)) > 0]
\end{aligned}$$

2   Two Minor Models

As is mentioned above, individuals' current level of educational attainment and the potential wage are not observable, so Attanasio (2011) constructs the following two models to estimate them.

2.1   Initial Condition Model

The initial condition model modelled the level of schooling already attained (ed_i) by student i by an ordered probit with index function $\beta_5 h_i + \beta_6 \mu_i + \epsilon_i^{ed}$, where h_i is a scalar observed variables affecting schooling decisions. μ_i is the unobserved heterogeneity defined in section 1.1. ϵ_i^{ed} is the random shock, which is assumed to be distributed as standard normal across the population conditional on μ_i. The model allows for thresholds that change with students' age and further assumes that the probability of $ed_i = e$ and of student i attending school ($D_i = 1$) to be:

$$P(ed_i = e, D_i = 1 | z_i, \ x_i^p, \ x_i^s, \ h_i, \ wage_i, \ \mu_i)$$
$$= P(D_i = 1 | z_i, \ x_i^p, \ x_i^s, \ ed_i, \ wage_i, \ \mu_i) * P(ed_i = e | z_i, \ x_i^p, \ x_i^s, \ h_i, \ wage_i, \ \mu_i)$$

The parameters β_5, β_6 of the initial condition model will be estimated jointly with the main model above.

2.2   Potential Wage Model

The original paper uses the following models to estimate the potential wages of students. First, they specify a standard Mincer-type wage equation, where the wage (ω_{ij}) of a student i living in community j is determined by his age and education according to

$$\ln \omega_{ij} = q_j + a_1 \text{age}_i + a_2 \text{educ}_i + \bar{\omega}_{ij}, \ (4)$$

where q_j represents the log price of human capital in the locality and $\bar{\omega}_{ij}$ is the residual. The paper follows the Heckman (1979) selection correction approach to specify equation (4). The inverse Mills ratio ($Mills_i$) is added into the regression and is constructed by estimating a reduced-form probit for school attendance as a function of the variables including in the main model. These include measures of the availability and cost of schools in the locality where the child lives, which controls among others for whether this is an experimental locality or not. Finally, they model q_j in equation (4) as a function of the male agricultural wage ω_j^{ag} in community j and whether the program was implemented in that community (P_j).

This model is estimated separately from the main model and the author assumes that individuals in the main model will use the point estimation of wages and ignore any variance around them.The resulting wage equation and estimation of wage for student i living in community j has the form:

$$\ln \omega_{ij} = \underset{(0.384)}{-0.983} + \underset{(0.028)}{0.0605} P_j + \underset{(0.049)}{0.883} \ln \omega_j^{\text{ag}} + \underset{(0.027)}{0.066} \text{age}_i$$

$$+ \underset{(0.0065)}{0.0116} \text{educ}_i - \underset{(0.053)}{0.056} \text{Mills}_i + \bar{\omega}_{ij}, \ (5)$$

3 Identification

The main identification problem of this paper is to jointly identify the parameters of the Education Choice Model and the initial condition model. In the Education Choice model, the parameters to be identified are: γ, δ, β_1, β_2, β_3, β_4, α_1, α_2, r and the distribution of μ_i. In the initial condition model, the parameters to be identified are: β_5, β_6.

Identification Proof:I am still working on the proof of Identification).