(**Not for Publication**)

Supplementary Material to: "Estimation of Dynastic Life-Cycle Discrete Choice Models"

George-Levi Gayle
Department of Economics, Washington University in St. Louis and Federal Reserve Bank of St. Louis.

Limor Golan
Department of Economics, Washington University in St. Louis and Federal Reserve Bank of St. Louis.

Mehmet A. Soytas
Faculty of Business, Ozyegin University

December 15, 2017

Abstract

These supplementary materials contain the details of the estimation algorithm, explanations for the first and second stages of the estimation results not presented in the main text, as well as the description of the data sets and programs supplied with the paper.

1 Estimation details

Our two step estimator can be described in 4 steps below. Of those 4 steps below, first 3 steps are first stage of the estimator and step 4 (the structural estimation) can be thought as the second step. The estimation proceeds in 4 steps. In **step 1** the (i) earnings equation (Table 3), (ii) intergenerational education production function (Table 4 and and Table 1-A), and (iii) the marriage market matching function at age 25 (Table 2-A, Table 3-A, Table 4-A, Table 5-A, Table 6-A, Table 7-A) are estimated. In **step 2** CCP for household choices (Table 8-A, Table 9-A, Table 10-A, Table 11-A, Table 12-A, Table 13-A, Table 14-A) are estimated. These estimated functions use the actual variables (wages, education outcomes of the children, spouse outcomes and the parental choices for the respective equations) in the data as the dependent variables. Also the explanatory variables of those functions include family (individual) characteristics and past and current choices. Therefore in the structural model, those explanatory variables are endogenously generated and supplied to these functions. The outcomes are just calculated by using the estimated coefficients of these functions whenever they are required. In **step 3** the alternative value function representation, the estimates from steps 1 and 2, and the Hotz et al. (1994)'s forward simulation technique are used to estimate the household continuation value for each age in the life-cycle. In **step 4** the Hotz-Miller (1993) inversion is used to form moment conditions for a generalized method of moment (GMM) estimation of the utility function parameters and discount factors.

1.1 Step 1

1.1.1 Estimation of the Education Production Function

Estimation of the education production function (Table 4 and Table 1-A) is conducted as a system estimation. We use an instrumental variable identification strategy with a linear probability model (IV-LPM).[1] There are three other methods of estimating discrete choice models with endogenous regressors[2] However, given the other issues (discussed below) in estimating the intergenerational production functions, the IV-LPM is the most straightforward method for simultaneously dealing with all these issues. We assumed that observed ability in the labor market is a monotonic transformation of academic ability; therefore by using the panel structure of our data we are able to estimate fixed effects for both parents and children using data on earnings.[3] This estimated fixed effect is then used in the estimation of the education production function to mitigate the ability bias. Second, we include fathers' education and time with children in the education production function while explicitly accounting for household interactions.

However, this leads to a third problem: the simultaneity of the inputs of both fathers and mothers and the endogeneity of which parent and type of parent spends time with a given child (by type of parent we refer to the education and skills of the parent and their spouse). The output of the intergenerational education production function (i.e., completed education level) is determined across generations, while the inputs, such as parental time investment, are determined over the life-cycle of each generation. Therefore, we treat these inputs as predetermined and use instruments from within the system to estimate the production function. This leads to a system of equations that need to be estimated simultaneously. The system of equations are the education (discrete) outcomes for the children, as well as labor supply, income, time spent with children. To estimate our system we need a number of exclusion restrictions. The first is the sex composition of siblings; it enters the education production function but not the labor supply equations. This is similar to the siblings-sex ratio, first used by Angrist and Evans (1998), and is justified on the basis that the sex composition of the children does

[1] One reason to use a linear specification is that the nonlinearity in the intergenerational production function itself can generate persistence in earnings across generations. However, we wanted to focus on the economic mechanism that generates persistence of earnings across generations.

[2] The three are: maximum likelihood, control variable, and special regressor approaches. See Lewbel, Dang, and Yang (2012) for a comparison of the different approaches.

[3] For complete estimation details and results of the fixed effects in the labor markets see Gayle, Golan and Soytas (2014).

not have a direct effect on labor supply, or the outcome of the child (again the outcome of the child depends on its own gender and number of siblings but not the siblings sex composition). However, sex composition has an indirect effect on parental time investment in children, and their outcomes through its effect on time investment in children (parents potentially spend different amount of time with boys and girls) and fertility (parents might have preferences, for example, for a balanced sex composition of their children). The second set of instruments – the difference in the age-earnings profile by education – is used to provide quasi-experimental variation in income, labor hours, and subsequent fertility.

1.1.2 Estimation of the Earnings Equation

The earnings equation (Table 3) is estimated using a GMM dynamic panel data using choices as instruments, as in Altug and Miller (1998), Blundell and Bond (1998), among others. The earnings for a male (female) individual depends on education, experience, and innate ability. Observed earnings are assumed to be noisy measures of individual's actual earnings. Taking logs and defining the functions $W_{\sigma t}$ and $H_{\sigma t}$:

$$\ln w_{\sigma t} = W_{\sigma t}(z_{\sigma t}) + H_{\sigma t}(h_{\sigma t}, ..., h_{\sigma t-4}) + \eta_{\sigma} + \epsilon_{\sigma t} \tag{1}$$

The error term in (1) is conditionally independent over individuals, the covariates in the earnings equation and the current and past part-time full-time labour supply decisions defined in $W_{f(m)t}$ and $H_{f(m)t}$. Since the error term enters log earnings linearly and is independent of any variables known at time t, including current labour supply, this representation for earnings allows us to circumvent the selection. In the estimation the past labor market experience is only relevant for the past 4 years given that the effect of experience with higher lags is insignificant.

The functions $W_{f(m)t}$ and $H_{f(m)t}$ are linear in their arguments and there first differencing equation (1) gives

$$\Delta\epsilon_{\sigma t} = \Delta \ln w_{\sigma t} - \Delta x'_{1,\sigma t}\Pi_1 - \Delta x'_{2,\sigma t}\Pi_2 \tag{2}$$

where the vectors x_1 and x_2 collects the variables age squared, age and education interactions, and current labor supply decision (full-time dummy) in the former and the past full-time and part-time decisions for the male and female separately for the later. This equation is estimated using generalized method moments (GMM) estimation with the optimal instruments, using the conditional independence of the disturbances with the covariates. The $M \times 1$ parameter vector to be estimated can be written as $\Pi = (\Pi'_1, \Pi'_2)'$. Define the $T - T^e$ dimensional vector Y_n, the $(T - T^e) \times M$ dimensional matrix X_n , and the square $T - T^e$ matrix W_n as:

$$\begin{aligned}
Y_n &= (\Delta \ln w_{f(m),T^e+2}, .., \Delta \ln w_{f(m)T})' && \text{(3a)}\\
X_n &= \begin{bmatrix} \Delta x_{1,f(m),T^e+2} & \Delta x_{2,f(m),T^e+2} \\ \vdots & \vdots \\ \Delta x_{1,f(m)T} & \Delta x_{2,f(m)T} \end{bmatrix} && \text{(3b)}\\
\Delta\epsilon_{f(m)} &= (\Delta\epsilon_{f(m),T^e+2}, ..., \Delta\epsilon_{f(m)T})' && \text{(3c)}\\
W_n &= E[(Y_n - X'_n\Pi)(Y_n - X'_n\Pi)' \mid X_n] && \text{(3d)}
\end{aligned}$$

The GMM estimator with lowest asymptotic covariance within this class is obtained as:

$$\hat{\Pi}_{GMM} = \left[N^{-1}\textstyle\sum_{n=1}^{N} X'_n\hat{W}_n X_n\right]^{-1}\left[N^{-1}\textstyle\sum_{n=1}^{N} X'_n\hat{W}_n Y_n\right] \tag{4}$$

where $\hat{W}_n$ is any consistent estimate of W_n (See Robinson (1987) for example).

1.1.3 Marriage Market Matching Function

Our model is a unitary household model. Parents decide on the number and also the spacing of the children in their life-cycle. For the completeness of the model in terms of generations of households, we need to construct the type of families formed for the children's generation. However this is an endogenous outcome in the model as it depends on the existence of the child in the first place. Secondly the gender and education of the child matters in the type of family he/she will form. The matching equations are required to match the children to their spouses in the second generation which completes the dynastic model.

The marriage market matching function is estimated at age 25 to abstract from the completed education considerations. Therefore depending on the characteristics of the adult child prior to age 25, the following equations are estimated: (i) Probability of husband's labor market history (Table 2-A); (ii) Probability of husband's education (Table 3-A); (iii) Probability of husband's age group (Table 4-A); (iv) Probability of wife's labor market history (Table 5-A); (v) Probability of wife's education (Table 6-A); (vi) Probability of wife's age group (Table 7-A). The estimation results are given in online appendix of the paper. Obviously if the child is a female, only (i), (ii) and (iii) are estimated, and (iv), (v) and (vi) are estimated if the child is a male. Of those equations, the husband (wife)'s labor market history is estimated using multinomial logit with not working, working part-time and working full-time as the discrete choices of the spouse. The variables used in the specification are age, education, number of all and only female children of the wife (husband), ages of the first 4 children of the wife (husband), wife (husband)'s time spent to the first 4 children and past 4 years of labor supply of the wife (husband); age group and education of the husband (wife), and also depending on which labor supply history is being estimated, the husband (wife)'s labor supply decisions of the previous years are used. For instance in the estimation of labor supply in year $(t-2)$, the labor supply decisions of the husband (wife) in years $(t-3)$ and $(t-4)$ are used. This is a consequence of the conditioning used in the estimation. Instead of estimating a big joint distribution of the husband (wife)'s characteristics, each of them are estimated as conditional probability distributions. Therefore husband (wife)'s labor market history is estimated conditional on husband (wife)'s age and education as can be seen from the variable list mentioned. Similarly, husband (wife)'s education is estimated using multinomial logit with less than high school, high school, some college and college as the outcomes conditional on the same list of variables of the wife (husband) and education of the husband (wife). Finally husband (wife)'s age is estimated as a group variable using multinomial logit with the age span from 18 to 60 grouped into 8 age brackets conditional on the same list of variables of the wife (husband).

1.2 Step 2

Conditional choice probabilities are estimated using multinomial logit specification. The estimation results are given in the online appendix of the paper. The choice set that is feasible for the husband (wife) in a particular age can vary in the model depending on the state variable (i.e. wife (husband) can only spend time with their child if they have one already or if they choose to have one in the current period in case they don't have one already). Therefore all possible choice sets that wife (husband) can encounter in the life-cycle are estimated conditional on the state. This leads us to estimate 2 specifications for the husband; (i) husband with young children (9 choices) and (ii) husband without young children (3 choices which are just the labor supply decisions). In the estimation of the probabilities of husband with young children (Table 8-A), husband has 9 choices as a combination of 3 discrete labor supply and 3 discrete time investment choices. Table 9-A estimates the multinomial logit specification with no work, part-time and full-time work as the choices, this is the estimation for the husband without young children case. The estimation equations use age, education, number of all and only female children of the husband, ages of the first 4 children of the husband, husband's time spent to the first 4 children, past 4 years of labor supply of the husband; age, education of the wife, wife's time spent to the first 4 children and past 4 years of labor supply of the wife as the covariates.

For the wife, there are 4 possible states in the life-cycle being in which leads to different choice sets. Those specifications of the states for the wife are; (i) infertile wife without young children (3 choices, Table 10-A), (ii) infertile wife with young children (6 choices, Table 11-A), (iii) fertile wife without young children (10 choices, Table 12-A), (iv) fertile wife with young children (16 choices, Table 13-A). Our model is based on households so we actually need to estimate the CCP for the household choices. Again as we did in the estimation of the matching function, we condition the wife CCP equations on the husband's choices. Therefore instead of estimating the joint distribution, we estimate the marginal (husband) and the conditional (wife) CCP equations. This is the reason for having husband choices in the wife CCP estimations. We also estimated one final CCP equation; fertile wife without young children (10 choices, Table 14-A). This equation is not conditioned on the husband choices since this is used in the first period of the adult female child. Same set of covariates are used in all equations whenever they are feasible.

1.3 Step 3

In **step 3** the alternative value function representation, the estimates from steps 1 and 2, and the Hotz et al. (1994)'s forward simulation technique are used to construct the household continuation value for each age in the life-cycle. Step 3 has some important intermediate steps that allows us to construct the continuation value. Starting with the alternative value function representation, those steps will be explained below.

1.3.1 Alternative value function representation

Relying on the representations derived in the paper, we form alternative valuation functions for the dynastic life cycle problem in terms of the utility function parameters, discount factors and the estimated quantities from step 1 and step 2 above. In the estimation, we need to construct the conditional valuation function for every single observation from the data. Let be clear about what is meant by a single observation. There are N individuals we observe over years. The number of years we observe the individuals can vary. This means that we may be observing an individual from age 25 to 30 (6 years), and another one from age 25 to 45 (21 years). We set the number of years of data from age 25 to 55 (31 years). Therefore from a theoretical perspective, we may have an individual who we observe from age 25 to 55. However in reality we will have observations fitting into some part of this age range for a particular individual. A single observation is a particular age observation for the nth individual. This observation will include the state vector z_{nt} (state vector now also indexed by n to represent its dependence on the nth individual). Given z_{nt}, we will construct the conditional valuation functions for the individual using the forward simulation. The sample used for the structural estimation is supplied as stata files in the data set and programs for the paper.

1.3.2 Forward simulation

The utility function is given as $u_{kt}(z_{nt}) + \varepsilon_{nt}$ for period t (which corresponds to the particular adulthood age). The first part of the utility is known if the state vector z_{nt} and the choice k are known up to an unknown parameter vector α. Moreover the utility function is linear in the parameter vector α by the specification. The model parameters are $\theta = (\beta, \lambda, \nu, \alpha)$, We need to calculate the expected value of future utilities in order to construct the utility from the life time of own action and consumption (which depends on the unknown parameter vector. We denote the period t counterpart of the utility from the remaining life time of own action and consumption for the ith generation individual as:

$$V_t^i(z_{nt}) = E_t\left[\sum_{s=t}^{T} \beta^{s-t} \sum_{k=0}^{17} I_{ks}^o\{u_{ks}(z_s, A_s^o) + \varepsilon_{ks}\}|z_{nt}\right]. \quad (5)$$

However in the above form, calculating the expectation requires calculating all the optimal choices in any future period conditional on choosing optimally in the periods before. Therefore this is quite cumbersome.

We use the forward simulation method of Hotz et al. (1994) to construct the conditional valuation functions by simulating a set of future paths starting from the value of z_{nt} in period t. In this algorithm we generate a realized state variable z_{ns} for each future period conditional on the accumulated information up to that point. Then conditional on the simulated state variable, we calculate the CCPs (probabilities of choosing each feasible action k_s in period s) for the household and simulate the optimal choice k_s^o for that period. In this way, we have $(T - t)$ choices and state vectors generated for the household starting with the state variable z_{nt} in period t. Then we repeat this algorithm n_g times to create more simulated outcomes for the same household. Using these n_g paths we construct the simulated counterpart of the equation in (5).

Similarly we denote the period t counterpart of the life time utility for a household with state z_{nt} at t as follows. Our problem is a dynastic life-cycle problem, therefore includes discounted value of generation $i + 1$ life time utility):

$$U_t^i(z_{nt}) = V_t^i(z_{nt}) + \beta^{T-t}\lambda E_t\left[N^{1-\nu}\overline{U}^{i+1}|z_{nt}\right]. \tag{6}$$

Observe that forward simulation algorithm described above helps us to construct the first expression on the right-hand side of equation (6). However we still need to construct the object $\overline{U}^{i+1}$. This is achieved through applying a matrix inversion implied by the stationarity assumption as described in the paper. Then relying on the representation obtained in the paper, a simulated version of the life time utility can be obtained using the forward simulation algorithm.

1.4 Step 4

Finally, in **step 4** the Hotz-Miller (1993) inversion is used to form moment conditions for a generalized method of moment (GMM) estimation of the utility function parameters and discount factors. Since the period utility function is linear in parameters α, moment conditions can be partially linearized conditional on some parameters of the model. This will further ease the estimation of the model. Therefore the GMM estimation can be conducted as a combination of a nonlinear search for the parameters β, λ, ν and a linear OLS solution for the parameters α for a given set of parameters $\beta^o, \lambda^o, \nu^o$.

2 Structural Estimation and Counterfactuals

Table 5 gives the structural estimation results for the model parameters. Tables 5 and 6 are obtained from the Fortran 90 programs that conducts the GMM version of the estimator using the estimations obtained from the previous steps as inputs. The stata files produced from the main estimation and from the counterfactuals to analyze the intergenerational effects of using the proposed estimator are given as a separate file. Since these are the essential contributions of the proposed estimator (the estimator gives one the ability to do full blown counterfactual analysis), we believe that those results are more informative for replication. The Fortran 90 codes implements the estimation using the forward simulation (described above) combined with a simplex algorithm for the parameter search for β, λ, ν. For each set of these parameters, the rest of the parameters in Table 5 are obtained as an OLS solution.