Parental Investment and Tax Policy

Musab Kurnaz
Department of Economics, Koc University

Mehmet A. Soytas
Faculty of Business, Ozyegin University

November 6, 2017

Abstract

A dynastic model of household behavior is used to estimate and decompose the correlations in earnings across generations. The estimated model can explain 75% to 80% of the observed correlation in lifetime earnings between fathers and sons, mothers and daughters, and families across generations. We find that human capital accumulation in the labor market, the nonlinear return to part- versus full-time work, and the return to parental time investment in children are the main forces driving the intergenerational correlation in earnings through their effects on fertility and the division of labor within the household. Assortative mating magnifies these forces.

Keywords: Intergenerational Models, Estimation, Discrete Choice, Human Capital, PSID. **JEL classification**: C13, J13, J22, J62.

1 Introduction

Intergenerational correlation of earnings is an important measure of mobility. A large positive literature has studied the sources of the correlation. This literature has analyzed parents' time and monetary investment decisions and their impact on the human capital of children which is transformed into future earnings. However, while the positive literature documenting the importance of parents' resource allocations on the correlation of earnings is extensive, the normative work exploring the policy implications of intergenerational correlation of earnings is not. Our paper fills this gap. We explore how the tax policy should be designed considering the impact of parental investment on the correlation of earnings between parents and children. We find that

We make theoretical and quantitative contributions. On the theoretical side, we embed a dynastic framework of Barro-Becker into life cycle model. The former has been used by macro and labor economist to analyze the implication of family dynamics across generations. Embedding a life cycle into this framework allows us to study the impacts of parental investment on children's per-period attainments. On the quantitative side, we do .

The optimal tax theory largely focuses on the trade-off between efficiency and redistribution. The policy designer can reduce income inequality by a progressive tax system, in which average taxes increases in income. However, the progressive taxes will lower parents' market labor supply and lower efficiency. On the other hand, the reduction in the market labor time can be devoted to time investment for children. This indirect change can reduce inequality in children's attainment and intergenerational correlation of earnings. Therefore, the progressivity rate of a tax system can be used not only for redistribution across recent generations and but also for shaping before-tax income distribution of future generations. Empirical evidences suggest that the correlation of earnings are higher in countries with higher progressive taxes such as Scandinavian countries. Moreover, the Gini coefficient of both market and disposable income in Scandinavian countries are much lower comparing to the US. This raises an important question: What should be the optimal tax progressivity considering its impact on the future income distribution?

To answer this question, we assume that the government uses a parametric tax code, $T_n(y) = y - \lambda_n y^{1-\tau_n}$, where $T_n(y)$ is the tax liabilities of $n-$ child families who generate y income and τ_n is the progressivity rate of tax liabilities of $n-$ child families. We find that

Our model is an extension of Gayle, Golan, and Soytas (2017) who essentially estimate life cycle dynastic version of Barro and Becker (1989).

Related Literature: Stantcheva (2015)

Stantcheva (2017)

Gelber and Weinzierl (2016)
cite Heckman and his co-workers!

2 Model

We model dynasties to analyze transfers and intergenerational transmission of human capital. Married individuals form a unitary household and parents are altruistic to their children.[1] Marriage is not a choice in the model but individuals match according to a matching function that depends on their characteristics. The matching function helps us to capture the nonrandom formation of families, which might affect the degree of investment in children as well as the family labor supply response to different tax policies. We do not model divorce which is beyond the objective of this paper. Yet, divorce decision can play important role on intra-household allocation (see Chiappori, Fortin, and Lacroix (2002)). In an equilibrium model, households choose Pareto optimally.

The aim of the model is to capture the impact of changes in the tax policy on the investment on children (both in levels and also in the composition of the investment in terms of goods and time), on the household structure (number of children, spouses labor supply arrangements) and on the intergenerational income mobility from a quantitative perspective. We assume that fertility is endogenous and households choose labor supply and investment in their children (money and time). The returns from children are modeled as the discounted expected value of children's utilities from the next generation as in Barro and Becker (1989). We incorporate the life-cycle in a dynastic framework with endogenous transmission of human capital. We explore the impact of different tax policies on parental investment, and long term outcomes of children.

2.1 Environment and Choices

An economy is populated with females and males, and each gender is indexed by $g = \{f, m\}$. Agents' life-time invariant characteristics, such as their education and labor market skill, are denoted by x_g and assume that the supports of x_f and x_m are finite.

Adults live for T periods. Adults may have children one at a period until period T^f after which the household is no longer fertile. Children (ages 0 to T^e) do nothing but are raised by adults. The childhood period is divided into the early childhood period (ages 0 to 5), and the later childhood period (ages 6 to T^e).[2] In the early childhood, parents spend

[1] We particularly focus on married households to avoid potential complication of parental investment and intergenerational correlations of earnings. Single households face tighter time constraints and their children suffers non-existing parental investment (even they receive this investment, we cannot observe this from data). Analysis on marital status is an important channel which we will leave as a future study.

[2] At period t of T, for some of the kids in the household, the early childhood period might have ended already (these children are over age 5), but for some others it might be the early childhood period (these children are below age 6).

time and goods to raise their children, and parents spend only goods in the late childhood periods. Children become adults at the period $T^e + 1$ and get married to according to a marriage matching function.[3] Married individuals form unitary households.

Consider a household (f, m). Starting from T^e, the household choose a discrete choice vector a and a continuous choice c in each period. The discrete choice vector is given by $a = (h_f, h_m, d_f, d_m, b)$ which consists of household market work time $h = (h_f, h_m)$, household time with children $d = (d_f, d_m)$, and whether to have a child b until the period T^f.[4] Let A represent the feasible set of action vectors. The continuous choice is the level of consumption.

For each period, t, we state a vector of state variables, which consists of the history of past choices, time invariant characteristics, and the gender of each child denoted by $z_t = (a_{T^e+1}, ..., a_{t-1}, \zeta_{T^e+1}, .., \zeta_{t-1}, x_f, x_m)$ where ζ is a dummy variable and denotes whether a child is a female when $b = 1$.

The gender of a child is modeled as random with the probability of having either gender equally likely. However, the exact gender composition of a family is somewhat endogenous in our environment. The age and gender composition of existing children – captured by $\zeta_{T^e+1}, .., \zeta_{t-1}$ in the state variable specified above – are in the state space. Therefore the gender composition of existing children can affect the decision to have another child or not. This will capture the well-known empirical finding that parents have a preference for gender balanced in the sex composition of their children (see Angrist and Evans (1998)).

Human Capital and Earnings Life-Cycle Dynamics The *realized* earnings, w_{gt}, of gender g at period t is decomposed into four categories: the interaction of the labor productivity with labor hours, work experience, innate ability (fixed effect), and idiosyncratic error term.

$$\ln w_{gt} = W_{gt}(e, h_{gt}) + H_{gt}(h_{gT^e+1}, ..., h_{gt-1}) + \eta_g + \epsilon_{gt} \quad \text{for} \quad g \in \{f, m\}. \tag{1}$$

The first component, $W_g(x, h_{gt})$, captures the interaction between the market labor hours, h_{gt}, and agent's labor productivity which is related with agent's education level e. This term is in the center of the mainstream of public finance literature which assumes that labor productivities are exogenous. In this paper, we allow *endogenous* labor pro-

[3]Marriage is not a choice, but the marriage matching function is designed to recover empirical moments related to marriage decisions.

[4]One can consider $b = 0$ after the period T^f.

ductivity via *endogenous* education decided by parents.[5] We allow $W_g(x, h_{gt})$ depend on h_{gt} in a nonlinear manner, for example, full-time work pays more than twice as such as part-time work.[6]

The second component, $H_g(h_{gT^e+1}, ..., h_{gt-1})$, represents the return of the experience on earnings and depends on the type of experience – part-time versus full-time – and how recent experience was obtained. This specification specifically captures both depreciation of human capital and differential returns to part-time versus full-time, both of which are gender-specific. One important benefit of the decomposition of realized earnings lies in this component. Empirical works show that the depreciation of experience is essential for females and the decision on childcare can draw back the return to experience.

The third component, η_g, captures the gender specific unobserved ability to earn income. This component provides rationale for those who have same education levels and work histories to have different earnings. This component also includes potential reward for adults' occupations.[7]

The fourth component, ϵ_{gt}, is an i.i.d. idiosyncratic error term.

The earnings dynamics specified above distinguish between endogenous state dependence through the return to experience and persistent productivity heterogeneity via education and unobserved ability, which form the characteristics of an agent: $x \equiv (e, \eta)$. The process of experience accumulation is central to our analysis because it captures the potential gender differences in the career interruptions and the effect of fewer labor market hours on the earnings of women and men. This may help rationalize some of the specialization patterns observed in the data. Moreover, market work decision can be deterred and hence not only the earning component but also the experience component is impacted by the tax rates which is discussed next.

Taxes and Budget Constraint The (joint) income of households are taxed according to the following tax code used by Guner, Kaygusuz, and Ventura (2014):

$$t(\tilde{y}) = \alpha + \beta \log \tilde{y} \tag{2}$$

[5]Starting from Mirrlees (1971), the literature on optimal income taxation assume that the earnings are multiplication of the labor productivity and hours: $w = \theta h$, where θ represents exogenous labor productivity (see also Saez (2001)). There are few studies on the impact of endogenous labor productivities on tax policies (see Ales, Kurnaz, and Sleet (2015) and references there). We let labor productivities are determined by education level determined by parents and we show the interaction between market hours and education in the empirical part.

[6]See Altug and Miller (1998), Gayle, Golan, and Miller (2012), and Gayle and Miller (2012), who document these features of the recent labor market.

[7]Empirically, η_g is fixed effect of a particular gender on earnings.

where $t(\tilde{y})$ shows average tax rate and $\tilde{y}$ is the normalized income (the ratio of household income to the mean household income). Guner, Kaygusuz, and Ventura (2014) show that this specification fits the US tax code using the official IRS data. We enhance this specification to family size, which is an important component of the US tax code. Tax code parameters, α, β, are estimated based on family sizes.[8] Equation (2) also implies that households' marginal tax rates are:

$$m(\tilde{y}) = \beta + t(\tilde{y}) = \alpha + \beta(1 + \log \tilde{y}).$$

This specification is consistent with *balanced growth* as stated by Guner, Kaygusuz, and Ventura (2014), which we follow to name this specification *Log* form.

We also compare our benchmark with Heathcote, Storesletten, and Violante (2017):

$$t(\tilde{y}) = 1 - \lambda \tilde{y}^{-\tau} \tag{3}$$

where λ determines the tax rate and τ determines the curvature of the tax code. Note that a tax system is called progressive whether the average tax rate is less than the marginal tax rate for all income: $\tau = 1 - \frac{1-m(y)}{1-t(y)}$. Therefore, if $\tau > 0$, then the tax code becomes progressive. Progressivity plays important role on the income distribution. In addition, the empirical literature shows that the correlation between parents' and children's income is lower in countries whose tax code is more progressive.[9] We provide rationale for this empirical finding in the result section. Moreover, the tax parameters λ, τ depends on family size and hence the progressivity in the taxes for a particular size would also play important role on the labor decision of parents and the income correlation.

Raising children requires parental time, d, and market expenditure. There is a per-period cost of the expenditures of raising a child, which is assumed to be proportional to the household's current earnings and the number of children. The budget constraint is described by the following equation:

$$c_t + \alpha(z_t)(N_t + b_t)w_t(z_t, h_t) \leq w_t(z_t, h_t)(1 - t_t(w_t)) \tag{4}$$

where $w_t(z_t, h_t)$ is total household earnings which is the sum of the earnings of the female, $w_{ft}(z_{ft}, h_{ft})$, and the earnings of the male, $w_{mt}(z_{mt}, h_{mt})$. N_t is the number of children at the beginning of period t, and b_t is the decision on having a child in period t. $\alpha(z_t)$ is

[8] Table 1 has the tax code parameters for families with up to four children.

[9] Jantti, Bratsberg, Roed, Raaum, Naylor, Osterbacka, Bjorklund, and Eriksson (2006) shows that the correlation in Nordic countries is almost half of the correlation in the US. Kleven (2014) shows that the progressivity of the tax code in Scandinavia is higher than the one of the US.

the proportion of household earnings spent per child.[10] Note that there is no saving in the model, so any after tax income plus the lump-sum tans fer is either spent on parental consumption or on children and the education subsidy works directly on the amount spent on children.

Utility Empirical Specification The following structure can be can be obtained. The within- generation utility, $u_{a_t}(z_t)$, can be written as a function of only the discrete actions by substituting the binding budget constraint for consumption. This is described by the following equation:

$$u_{a_t}(z_t) = \theta_{a_t}(z_t) + u_t\left[w_t(z_t, h_t)(1 - t_t(w_t) - \alpha(z_t)(N_t + b_t)), z_t\right] + \varepsilon_{a_t}$$

where $\theta_a(z)$ is dis/utility from taking discrete action a and $u_t[., z_t]$ is the utility from consumption. Associated with each possible discrete action is a per-period additive state specific error ε_a.

Assming a linear utlity function (wealth maximization?), the utility from period consumption and lesisure can be written as follows:

$$u_{a_t}(z_t) = \theta_{a_t}(z_t) + w_t(z_t, h_t)[1 - \tau_{Lt} - \alpha(z_t)(1 - \tau_{Et})(N_t + b_t)] + \rho(z_t, b_t)G_t + \varepsilon_{a_t}$$

The empirical specification of the period utility will rely on this theoreticall budget contraint, only it will be differentiated for the different household types. Let's denote the choice made in period t by k now out of the possible set of choices K. The utility and the income functions will be written with a choice subscript and their dependence on the state z_t will be made implicit. The birth decision b_{tk} will denote the corresponding birth choice associated with the choice k. This abuse of notation will help to better describe the identification in the tedious derivations.

[10] PSID do not include the information on expenditures for children. By this assumption, we capture the differential expenditures on children made by households with different incomes and characteristics since α be a function of z. Moreover, expenditures for children, and the household income, can create substantial difference in cognitive skills (see Dahl and Lochner (2012)).

$$
\begin{aligned}
u_{tk} &= \theta_k + \alpha_0 w_{tk} + \alpha_1 w_{tk}(N_t{+}b_{tk}) + \alpha_2 w_{tk}(N_t{+}b_{tk})HS_{mot} \\
&\quad + \alpha_3 w_{tk}(N_t{+}b_{tk})SC_{mot} + \alpha_4 w_{tk}(N_t{+}b_{tk})COL_{mot} \\
&\quad + \alpha_5 w_{tk}(N_t{+}b_{tk})HS_{fat} + \alpha_6 w_{tk}(N_t{+}b_{tk})SC_{fat} \\
&\quad + \alpha_7 w_{tk}(N_t{+}b_{tk})COL_{fat} + \alpha_8 G_t + \varepsilon_k
\end{aligned}
$$

where $\alpha_0 = 1 - \tau_{Lt}$, $\alpha(z_t)(1-\tau_{Et}) \approx [\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7]$, $G_t \approx (N_t{+}b_{tk})$, so $\alpha_8 \approx \rho(z_t, b_t)$. With this parametrization the coeffcients of the utility combined with the fact that estimation will be based on moment conditions depending on value function differences, the period lump sum amount of G_t can not be separetly identified. This basically means the θ_j, the dis/utility from taking discrete action j corresponing to $(N_t{+}b_{tj})$ captures the effect of this lump sum transfer in the estimation. Therefore theis coeffients should be interpreted accordingly in terms of level effects differences in the final estimation ouputs. Next we will form the moment conditions that will allow us the estimation of the parameter set $\Theta = (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7)$.

The family can choose among K possible choice alternatives in each period during their life-cycles. As in Gayle et al. (2015), this choice set is restricted to the possible actions the family can can choose depending on the which stage of the life-cyle the family is. For instance, the family will only be able to invest in their children if they already have children or chooses to have one in the current period.

In the analysis, the period earnings w_{tk} depends on the participation decision, the past labor market attachement, skills. The family can allocate all of its adult individuals to market work or can choose to differentiate between who will spend more in labor work versus household work. However in the paper, we assume it is the total household income that is taxed through the tax function parameter τ_{Lt}, which is depending on the household income and number of children. Therefore the optimizing family should decide on this labor/home time trade-off taking into account the implications of the tax policy on the final household income. The tax code we use is based on the specifications suggested by Guner et al. (2015) and Heathcode et. al. (2017).

Children Outcomes To capture the impact of parents characteristics and inputs on children education, we specify an "education production function" that accounts for parental time and monetary investment in children, as well as the parents characteristics and skills. We denote child's education by e', and child's ability by η'. The characteristics of children in the next generation $x' \equiv (e', \eta')$, are affected by their parents' characteristics $x \equiv (e, \eta)$, early childhood time investments, monetary investments in her childhood, and the presence and timing of siblings in early childhood. We further index the variables by gender; for instance e'_f represents the education outcome of a daughter in the next generation. Formally, the child education is determined by the following sets of equations:

$$
\begin{aligned}
e'_g &= \Gamma_g[x, d^{(0)}, ..., d^{(5)}, w^{(0)}, ..., w^{(5)}, S_{-5})] + \omega'_g && \text{(5a)}\\
\eta'_g &= \Gamma_{g\eta}(e'_g) + \widetilde{\eta}'_g && \text{(5b)}\\
\Pr(\widetilde{\eta}' &= \widetilde{\eta}_g) = F_g(e_f, e_m, \eta_f, \eta_m). && \text{(5c)}
\end{aligned}
$$

In the empirical implementation, Γ_g and $\Gamma_{g\eta}$ are both linear functions. The vector $d^{(j)} = (d_f^{(j)}, d_m^{(j)})$ is the parental time investment at age j of the child, $w^{(j)}$ is the household earnings at age j of the child, S_{-5} is the gender-adjusted number of young siblings present in the household during early childhood, and ω'_g is the gender-specific luck component that determines the educational outcome of the offspring. Child's ability, η'_g, is determined once the education level is determined as the sum of systematic, $\Gamma_{g\eta}(e'_g)$, and random components, $\widetilde{\eta}'_g$ which is assumed to have finite support and to be independent of ω'_g with probability distribution function, $F_g(e_f, e_m, \eta_f, \eta_m)$.

Equations (5a)-(5c) imply that child's ability into a component determined by parental inputs through the effect of the educational outcome and a random component which is determined by the interaction of parents' education and ability. This specification allow us to capture the effect of parental monetary and time investment on children's outcome. Through the life-cycle of adults, income taxation plays important role on labor supply decision and hence household income. Through the life-cycle of parents, the income taxation plays also important role on time investment in children. The reduction in labor supply also implies reduction in household income and hence monetary investment. Yet, the time amount of reduced from market can be transferred to investment in children which can surpass the negative effects of the reduction in monetary investment. In the result section, we show the significance of income taxation on parental investment and show the interaction between monetary time investment on children's earnings.

Investment in Children and Stock of Human Capital in the Life-cycle To capture the life-cycle early childhood human capital accumulation, we alllow a flexible functional form for the evolution of the human capital of the child in the childhood. The literature on skill accumulation generally model the age specific investment as an input in children's cognitive skills (and sometimes noncognitive skills also as a separate set of skills along with cognitive). The output at period t is the skill of the child which depends on the period $t-1$ skills and the investment input in period t. To capture the dynamics of human capital accumulation, we allow the accumulation:

$$k_t = [\pi \left(d^{(t)}\right)^{\phi} + (1-\pi)\left(\alpha(z_t)(N_t+b_t)w^{(t)}\right)^{\phi}]^{\frac{1}{\phi}}$$

With this formulation, the aggregated investment k_t should matter for the skill accumulation of the child[11]. Therefore, we can replace k_t in equation (5a) and the remaining equations will follow:

$$e'_{f(m)} = \Gamma_{f(m)}[x, k^{(0)}, ..., k^{(5)}, w^{(0)}, ..., w^{(5)}, S_{-5})] + \omega'_{f(m)}$$

Shocks There are four main shocks that characterize the model, the timing of the realization of these shocks are crucial understanding the predicts of the model. The first shock is embedded in the matching probability, $G(x_m, x_f)$, is realized at the beginning of adulthood, T^e+1. The second shocks are on the per-period time allocation and the fertility preferences. Both shocks are realized at the beginning of each period during the adulthood and are i.i.d. across households and time. The third shock is the shock to unobserved ability in the labor market, $\widetilde{\eta}_g$. This shock is realized at the beginning of adulthood, T^e+1, and is persistent over the household life-cycle but independent across parents and children. The final shock is on the children's educational outcome, ω'_g, which is realized at the end of childhood, T^e, and is independent across generation.

Timing of Choices Within each period the model uses a timing for choices made by adults. The timing is:

[11]Obviously, we can write a full life-cyle dynamics for the skill accumulation at the early childhood as the period $t+1$ skill σ_{t+1} should be affected by the period t investments (aggregated money and time) as well as the period t level of the period t skill σ_t.

$$\sigma_{t+1} = \exp(\xi_t)[\kappa k_t^{\psi} + (1-\kappa)\sigma_t^{\psi}]^{\frac{1}{\psi}}$$

Of course, in the empirical implementation, this will imply at the end of age 5 of the child:

$$\ln \sigma_6 = \frac{1}{\psi}\ln[\kappa k_5^{\psi} + (1-\kappa)\sigma_5^{\psi}] + \xi_6$$

Education outcome of the child in the next generation will be a function of parental traits,investment and the level of the child skill as in Cunha et. al (2006). However the education production function in the Gayle et. al (2015) is estimated outside of the structural model depending on the realized histories taken from the data. Therefore, the paremeters governing the "structural skill formation" can only be identified up to some parameters. We will disscuss this further in the empirical implementation.

1. The preference shocks for the period is chosen by nature and observed my the households.
2. The household makes fertility, time allocation, and consumption decisions.

Preferences and Household Optimization Adult households care about consumption, leisure, the number of children, and the future household utility of their children. Extending the original formulation of Barro and Becker (1989) to unitary households, we assume that the life-time utility for a couple (f,m) household at age T^e+1 is as follows:

$$U^i(f,m) = V^i(f,m) + \beta^{T-T^e-1}\lambda E_{T^e+1}\left[N_{T^f}^{1-v}\overline{U}^{i+1}|f,m\right], \tag{6}$$

where $U^i(f,m)$ represents the full value of the utility of a household at age T^e+1 in generation i from that point forward; $V^i(f,m)$ is the utility the household derives from its own path of consumption and discrete actions; N_{T^f} is the number of children in the household at the end of the fertile period ,and $\overline{U}^{i+1}$ is the expected utility of the household to which their typical child will be assigned.

Let $I^o_{a_t}$ be the indicator variable of the optimal discrete choice of a type-(f,m) household of age T^e+1+t. We assume that the utility from the life-time of own action and consumption is of the form

$$V^i(f,m) = E_{T^e+1}\left[\textstyle\sum_{t=T^e+1}^{T}\beta^{t-T^e-1}\sum_{a_t\in A_t} I^o_{a_t}\{u_{a_t}(z_t)+\varepsilon_{a_t}\}\right]. \tag{7}$$

We distinguish between the time preference, β, and the degree of altruism between generations, λ. Thus, $\lambda=1$ means that a household cares as much about their children's household utility as they care about their own. Also, households discount the utility of each additional child by a factor of $1-v$, where $0<v<1$ because we assume diminishing marginal returns from children. The within- generation utility, $u_{a_t}(z_t)$, can be written as a function of only the discrete actions by substituting the binding budget constraint for consumption. This is described by the following equation:

$$u_{a_t}(z_t) = \theta_{a_t}(z_t) + u_t[w_t(z_t,h_t)(1-\alpha(z_t)(N_t+b_t)), z_t]$$

where $\theta_a(z)$ is dis/utility from taking discrete action a and $u_t[.,z_t]$ is the utility from consumption. Associated with each possible discrete action is a per-period additive state specific error ε_a.

Similar to equation (6), we can define expected future utility for a young adult in generation $i+1$ at age T^e+1. Therefore, recursively $\overline{U}^{i+1}$ is described by the following equation:

$$\overline{U}^{i+1}(f,m) = \frac{1}{N_{Tf}} \sum_{n=1}^{N_{Tf}} \sum_{f'=1}^{F} \sum_{m'=1}^{M} G(f',m') U_n^{i+1}(f',m'), \tag{8}$$

where N_{Tf} is total number of children in the household during the fertile period, and $U_n(f',m')$ is the expected utility of the household of child n.

3 Empirical Results

3.1 Discussion

In Barro and Becker (1989) model with endogenous fertility there is no persistence in income. However, several features of our model can lead to intergenerational persistence in income. These are (i) the non-linearity in the cost of transferring human across generations, (ii) non-separability in the feasible set across generations, (iii) specialization in housework and labor market work within households, and (iv) assortative mating.[12]

The per-period cost of raising children and transferring human capital across generations is described in the budget constraint in equation as well as the opportunity cost of time investment input in children, which is the forgone earnings. Time investment and labor supply are modeled as discrete choices which introduces non-linearity in the cost of raising children and transferring human capital. Specifically, the fact that labor supply is discrete and that the returns to part- and full-time work are nonlinear allows for the cost of transferring human capital to each child to be increasing in the number of children. As a result it can generate persistence in income across generations.

We incorporate the dynamic elements of the life-cycle, which involve age effect and experience. The opportunity cost of time with children therefore incorporates returns to experience, which are also nonlinear (depends on the level of labor supply). The non-linearity involved in labor supply is realistic; parents' labor market time is often not proportional to the number of children they have and hours in the labor market. For a given wage rate, these are not always flexible and depend on the occupation and type of job. Furthermore, fertility decisions are made sequentially, and due to age effects, the cost of a child varies over the life-cycle. Mookherjee, Prina, and Ray (2012) develop a model with most of these characteristics. They show that by incorporating a dynamic analysis of the return to human capital can help generate persistence in a dynastic Barro-Becker model.

[12]See Alvarez (1999) for similar conditions that can generate persistence in income and wealth across generations in dynastic models with endogenous fertility. Also see Doepke (2005) and Jones, Schoonbroodt and Tertilt (2008) for other discussion of these conditions.

The feasible set across generations is non-separable in our model because the wages of the children (and therefore, their opportunity cost of time) depend on their education and labor market skills. However, education and labor market skills of children are linked to their parents' skills and education through the production function of education. This is one of the most natural ways of generating persistence in the standard dynastic model.

Incorporating two household members into the model captures important issues of the degree of specialization in housework and labor market work in households with different composition of education. The importance of which spouse spends time with the children (and the levels of time) depends on the production function of education of children and whether the time of spouses is complement or substitute. To the best of our knowledge, ours is the first paper to explicitly analyze this mechanism as a potential source of intergenerational persistence in earnings.

Finally, patterns of assortative mating may amplify the persistence of income across generations relative to a more random matching pattern. In our model, there is potential correlation of the cost of transfers to children (time input) with both parents' characteristics and assortative mating patterns. This implies that if children of more educated parents are more likely to be more educated, they are also more likely to have a more educated spouse, which increases the family resources and their children's educational outcomes. Several recent papers have highlighted the importance of this mechanism for explaining cross-sectional inequality. For example, Fernandez and Rogerson (2001); Fernandez, Guner, and Knowles (2005); and Geenwood, Guner, Kocharkov, and Santas, (2014, 2016). While these papers do not directly analyze intergenerational persistence of earnings, they do use dynastic models with household behavior which similar to the one used here.

3.2 Results

3.2.1 Discount Factors and the Direct Costs of Raising Children

This section presents estimates of the intergenerational and intertemporal discount factors, the preference parameters, and child care cost parameters. Table 4 describes the utility function estimates including the discount factors. It shows that the intergenerational discount factor, λ, is 0.795 (for complete table of the estimates see table B-1). This implies that in the second to last period of a parent's life, the parent's valuation of their child's utility is 79.5% of their own utility. The estimated value is in the same range of values obtained in the literature calibrating dynastic model (Rios-Rull and Sanchez-Marcos, 2002; Greenwood, Guner, and Knowles, 2003). However, these models do not the include life-cycle. The estimated discount factor, β, is 0.813. The discount factor is smaller than typical calibrated values; however, few papers that estimate it find lower values (for example, Arcidiacono, Sieg, and Sloan, 2007, find it to be 0.8).[13] Lastly, the discount factor associated with the number of children, υ, is 0.111. This implies that the marginal increase in value from the second child is 0.68 and of the third child is 0.60.

Table 3 also presents the marginal utility of income. Utility from income declines in the number of children; for a person with less than a high school diploma and a spouse with less than a high school diploma, the coefficient on the interaction of children and family income is -0.309, implying that the net costs of raising children increase with the number of children as well as the family income[14]. The costs decline with own and spouse education. However, for all households the net utility from children is negative and declining in family income, capturing the increase in spending on children for wealthier families. For families with the same income and number of children, the costs of children increase in income for all types of households. In our model, fertility decisions depend, therefore, on education and income through the costs in the utility function. The costs of children are lower in households with higher education; however, these costs increase in income and income is higher for more educated households. The earnings equations capture the increase in earnings and, therefore, the increase in opportunity costs of time for more educated households. In the Barro-Becker model, the neutrality result holds - that is, wealthier people have more children, so the investment per child is the same and there is no intergenerational persistence. In our model, however, several other channels are correlated with education creating persistence. Whether wealthier households have more or fewer children and whether investment per child increases in more educated households is an empirical question.

3.2.2 Model Fit and Explanatory Power

There are many criteria for assessing the fit of a model; in this paper we used two such criteria. The first is the statistical overidentifying J-test. We cannot reject the overidentifying test at the 5% level. The other criteria require us to solve the model numerically. As such, we numerically solve the model and simulate 10,000 synthetic generations. Then given the synthetic dataset, we fist calculate the labor supply, time investment, and fertility patterns of families and compare it with the estimates from the data. This is the simulation colunterpart of the J-test we conducted. Secondly, we calculate the intergener-

[13]We are not aware of dynastic models in which the time discount factor is estimated.

[14] Note that the coefficients on children in the utility represent net utility because we cannot observe expenditure on children directly.

ational correlation of earnings and compare the results to the estimates of the correlations from the data. This is an independent source of model validations as these correlations are not moments that are targeted in the estimation. Table XX provides the results of this latter exercise.

TABLE 5: STRUCTURAL ESTIMATES OF DISCOUNT FACTORS AND UTILITY PARAMETER

Variable	Estimates	Variable		Estimates
Discount factors		Disutility/Utility of Choices		
β	0.816	Wife	Husband	
	(0.002)	Labor supply		
λ	0.795	No work	Part -time	-0.512
	(0.200)			(0.005)
υ	0.248	No work	Full-time	0.207
	(0.168)			(0.009)
Marginal Utility of Income		Part-time	No work	-2.023
Family labor income	0.480			(0.003)
	(0.004)	Part-time	Part-time	-1.168
Children x Family labor income	-0.466			(0.009)
	(0.066)	Part-time	Full-time	-0.605
Children x HS x Family labor income	1.216			(0.008)
	(0.065)	Full-time	No work	-0.408
Children x SC x Family labor income	1.279			(0.007)
	(0.066)	Full-time	Part-time	-1.24532
Children x COL x Family labor income	1.300			(0.011)
	(0.065)	Full-time	Full-time	0.001
Children x HS spouse x Family labor income	-1.017			(0.010)
	(0.066)	Time with children		
Children x SC spouse x Family labor income	-0.995	Low	Medium	0.502
	(0.066)			(0.014)
Children x COL Sspouse x Family labor income	-0.992	Low	High	0.564
	(0.066)			(0.013)
Children x Black x Family Labor Income	-0.108	Medium	Low	-0.169
	(0.004)			(0.008)
		Medium	Medium	0.129
				(0.010)
		Medium	High	0.593
				(0.013)
		High	Low	-0.364
				(0.007)
		High	Medium	0.353
				(0.011)
		High	High	-0.140
				(0.012)
		Birth		0.701
				(0.025)

3.2.3 Effect of Labor Tax on the Intergenerational Persistence in Earnings

This secton explores the effect of the tax schedule on the dynamic decison making behavior of the households. The taxes in our model are basically distortions to the labor income of the households, however with the dynamic structure in the lifecycle involving labor supply and time investment decisions, taxes can affect household decisons intertemporally. One of the major benefit of our structural econometric approach is the ability to conduct counterfactual analysis. An alternative approach which has been used in empirical structural models to incorporate intergenerational/altruistic concerns is to modify the standard dynamics structural estimation methods by introducing an approximation for the value parents place on their children's adults outcome as a function of some state variables, normally the educational outcome or test scores (See example Bernal (2008), Brown and Flinn (2011), and Del Boca, Flinn, and Wiswall (2013) among others.). The main advantage of this alternate approach is that the estimation is easier and standard techniques in the literature can be used. The major disadvantage is that welfare /counterfactual analysis is subject to the Lucas's Critique. That is, in a counterfactual environment the value parents place on their children quality changes in two ways, the value of the state variables and the functional form of the mapping between the state variables and the utility derived from those state variables (See Gayle, Golan and Soytas 2015). This alternative approach does not allow the functional form of the mapping to change. This is in fact can be crucial in aswering questions in tax literature. Since changes in taxes alter the behavior of the households in the current generations, so do it for the future generations. The implications of a change in the tax rate/policy, therefore can have an across generations effect, both due to re-optimizing parents (who will reallocate resources through time and money investment) and re-optimizing children (who will face a different initial endowment to start their lifecycle). For instance a tax schedule with increasing marginal rates reduces both the returns to working more hours and the returns to acquiring human capital (e.g., Heckman, Lochner, and Taber 1998; Krueger and Ludwig 2013; Guvenen, Kuruscu, and Ozkan 2014). Of the two aferomentioned effects, the former attracted more attention in the tax literature (Gayle and Shephard, 2017). Studying the second effect quantitatively obviosuly requires a dynastic model with children's behavior in the second generation modeled explicitly. Therefore a dynastic model as the one introduced in this paper with a life-cycle component can be used as a natural framework to study the effect of taxation in altering generational behavior and the consequently to quantify the welfare implications of tax policy in terms of intergenerational income mobility.

We will conduct several counterfactual anaysis to identify the effect of income taxation on the intergenerational transimission. Our first analysis eliminates the main important feature of U.S. taxation scheme. We set the tax code to be indepenedent of the number of children. With this, families still will pay their tax according to their earned income, however having more children will not decrease the effective tax rate for the family given the income. In the second counterfactual analysis we set a flat average tax rate for households given the number of children. Therefore in the second counterfactual we take into account this important feature of U.S. income tax schedule but make it available for any income level using the dyanstic model's estimated preference parameters and transitions. In each counterfactual we simulate outcomes for a syntetic cohort of 10,000 individuals. We calculate the the intergenerational correlation of income across parents and children, and report them. We name these two counterfactual analyses as TWOC (tax without children) and TWC (tax with children) respectively where the abbrevations relate the counterfactuals dependence on the number of children.

One can see that our framework can nest the issues in the afferomentioned alternative approach in the literature, since the approximation used in the alternative approach is equivalent to conducting a welfare/counterfactual analysis holding fixed the CCPs and transition used in the calculation of the value of a child. To illustrate the possible bias induced by ignoring the fact that the children themselves will re-optimize when we change the economic environment we asked the same counterfactual questions of how much the mobility across generation will be affected by the dependency of the tax system on the number of children and to the differences in the average tax rate at different income levels. To do these in the empirical implementation, we set the tax function of each household equal to the tax function of a family without any children for the TWOC and set the tax function equal to an tax rate of an average income earning family with a given number of children for the TWC respectively. We also report the results from the estimated model as MODEL to compare the results from the countefactuals.

Table XY presents the summary of labor supply, time investment, and birth rate by gender and race for the data, baseline model and counterfactual simulations. It shows that if we eliminate the portion of parental education that is transmitted automatically across generation then parents will re-optimize and change labor supply, time investment, and fertility behaviors. Therefore, a pure statistical decomposition would be inappropriate for answer the question of how much mobility would change if there were no automatic (Nature) transmission of education from parents to children. The columns $NN1$ presents the counterfactual estimates of our model and $NN1'$ the estimation results of the approximated model; similarly, $NN2$ presents the results of the counterfactual from our model with ($\beta = 0.90$ and $\lambda = 0.95$) and $NN2'$ the results of the counterfactual of the approximated model. The columns $NN1'$ and $NN2'$ show that not taking into account that the all subsequent generations will also re-optimize induces significant bias with the bias being greater the larger the discount factors.

To obtain a number that summarize the impact on mobility, Figure 1 presents the probability that a children born in a family in the bottom 20 percent of the family income distribution will end up a family in with family income above the median of the next generation family income distribution. It shows that in the baseline model, i.e. model1, that only 30 percent of children born in the bottom 20 percent will end up with families earning above the median. However, if the automatic transition of education was eliminated that probability would increase by about 20 percent to about 40 percent. However, we not account for the fact that subsequent generation will re-optimize we would over estimate it impact by about 25 percent. Model (2) shows the similar qualitative patterns but shows that the overestimate of the impact of 'Nature' on mobility could be as high as 90 percent. Which illustrate the gain from using the approach outline in this paper.

3.2.4 Effect of Labor Tax on

Individuals differ exante with respect to two characteristics: learning ability and the disutility of work effort. Those with higher learning ability invest more in skills prior to entering the labor market, whereas more diligent individuals work and earn more.

4 Taxes

Knowles (1999)

References

Ales, L., M. Kurnaz, and C. Sleet (2015). Technical Change, Wage Inequality, and Taxes. *American Economic Review 105*(10), 3061 – 3101.

Altug, S. and R. A. Miller (1998). The Effect of Work Experience on Female Wages and Labour Supply. *The Review of Economic Studies 65*(1), 45–85.

Angrist, J. D. and W. N. Evans (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *The American Economic Review 88*(3), 450–477.

Barro, R. J. and G. S. Becker (1989). Fertility Choice in a Model of Economic Growth. *Econometrica 57*(2), 481–501.

Chiappori, P.-A., B. Fortin, and G. Lacroix (2002). Marriage Market, Divorce Legislation, and Household Labor Supply. *Journal of Political Economy 110*(1), 37–72.

Dahl, G. B. and L. Lochner (2012, May). The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit. *American Economic Review 102*(5), 1927–56.

Gayle, G., L. Golan, and M. Soytas (2017). Estimation of Dynastic Life-cycle Discrete Choice Models. *Quantitative Economics forthcoming*.

Gayle, G.-L., L. Golan, and R. A. Miller (2012). Gender Differences in Executive Compensation and Job Mobility. *Journal of Labor Economics 30*(4), 829 – 872.

Gayle, G.-L. and R. A. Miller (2012). Life-Cycle Fertility and Human Capital Accumulation. Working paper.

Gelber, A. and M. Weinzierl (2016). Optimal Taxation When Children's Abilities Depend on Parents' Resources. *National Tax Journal 69*(1), 11–40.

Guner, N., R. Kaygusuz, and G. Ventura (2014). Income Taxation of US Households: Facts and Parametric Estimates. *Review of Economic Dynamics 17*(4), 559–581.

Heathcote, J., K. Storesletten, and G. L. Violante (2017). Optimal Tax Progressivity: An Analytical Framework. *Quarterly Journal of Economics*, 1–62.

Jantti, M., B. Bratsberg, K. Roed, O. Raaum, R. Naylor, E. Osterbacka, A. Bjorklund, and T. Eriksson (2006). American exceptionalism in a new light: a comparison of intergenerational earnings mobility in the nordic countries, the united kingdom and the united states.

Kleven, H. J. (2014). How can Scandinavians tax so much? *The Journal of Economic Perspectives 28*(4), 77–98.

Knowles, J. (1999). Can Parental Decisions Explain U.S. Income Inequality? Mimeo, University of Pennsylvania, 1999.

Mirrlees, J. (1971). An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies 38*(2), 175 – 208.

Saez, E. (2001). Using Elasticities to Derive Optimal Income Tax Rates. *Review of Economic Studies 68*(1), 205 – 229.

Stantcheva, S. (2015). Optimal Income, Education, and Bequest Taxes in an Intergenerational Model. Technical report, National Bureau of Economic Research.

Stantcheva, S. (2017). Optimal Taxation and Human Capital Policies over the Life Cycle. *Journal of Political Economy 6*(125), null.

A Appendix

Estimates	All	0 Children	1 Child	2 Children	3 Children	Four+ Children
Log						
α	0,1525	0,1711	0,1565	0,1377	0,1225	0,1126
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
β	0,0968	0,0856	0,1070	0,1076	0,1043	0,1152
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
HSV						
λ	0,8472	0,8290	0,8440	0,8607	0,8780	0,8853
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
τ	0,1020	0,0937	0,1121	0,1129	0,0936	0,1186
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Mean income of the sample is 65,912 in 2005$s and 58,329 in 2000$s and the Guner's sample's mean income is 57k in 2000$s.

Table 1: Tax Code Parameters

Tables and Figures