Parental Investment and Child Development

Duc Thanh Nguyen

June 10, 2024

Objectives

- Formulate a dynamic production function model for the various dimensions of human capital with endogenous parental investments to examine dynamic complementarities and interactions among different inputs in forming child human capital.
- Use a maximum likelihood approach by the minorization-maximization algorithm to estimate the distribution of the latent factors and dynamic CES production functions of human capital.

Contributions

- This research is the first attempt to analyze the determinants and interactions of three important dimensions of human capital: cognitive skills, noncognitive skills and health.
- It provides rare evidence in developing country settings on parental investments and child development
- It uses a latent factor approach to identify the latent, unobserved factors instead of (noisy) proxy variables to correct measurement error problems and explore the endogeneity of investments.

Model: Dynamics of Skill Formation

- The framework builds on the dynamic factor models of Cunha et al. (2010), Cunha and Heckman (2007), Attanasio et al. (2017), Agostinelli and Wiswall (2016) and Aucejo and James (2021):

$$\Theta^k_{i,t} = f(\Theta^C_{i,t-1}, \Theta^{NC}_{i,t-1}, \Theta^H_{i,t-1}, I_{i,t}, P^C_i, P^{NC}_i, X_{i,t}, A^k_t, v_{i,t}, \varepsilon^k_{i,t}) \tag{1}$$

$$k \in \{C, NC, H\}$$

- I use a Constant Elasticity of Substitution (CES) production function

$$\begin{aligned}\Theta^k_{i,t} = \; & \left[\gamma^k_{1t}(\Theta^C_{i,t-1})^{\rho^{tk}} + \gamma^k_{2t}(\Theta^{NC}_{i,t-1})^{\rho^{tk}} + \gamma^k_{3t}(\Theta^H_{i,t-1})^{\rho^{tk}} + \gamma^k_{4t}(I_{i,t})^{\rho^{tk}}\right. \\ & \left.\gamma^k_{5t}(P^C_i)^{\rho^{tk}} + \gamma^k_{6t}(P^{NC}_i)^{\rho^{tk}}\right]^{1/\rho^{tk}} e^{X'_{i,t}\delta^k_t + A^k_t + \mu^k v_{i,t} + \varepsilon^k_{i,t}}\end{aligned} \tag{2}$$

Where $\gamma^k_{1t} + \gamma^k_{2t} + \gamma^k_{3t} + \gamma^k_{4t} + \gamma^k_{5t} + \gamma^k_{6t} = 1$.

Measurement System

- The latent factors ($\Theta^{C}_{i,t}$, $\Theta^{NC}_{i,t}$, $\Theta^{H}_{i,t}$, $\Theta^{C}_{i,t-1}$, $\Theta^{NC}_{i,t-1}$, $\Theta^{H}_{i,t-1}$, P^{C}_{i}, P^{NC}_{i}, and $I_{i,t}$) are not directly measured, the factor model approach is used to extract these unobserved variables from a large set of observed data.
- Define the natural log of the factors as $\theta^{k}_{i,t} = ln(\Theta^{k}_{i,t})$, $\theta^{k}_{i,t-1} = ln(\Theta^{k}_{i,t-1})$, $\mathcal{P}^{C}_{i} = ln(P^{C}_{i})$, $\mathcal{P}^{NC}_{i} = ln(P^{NC}_{i})$ and $\mathcal{I}_{i,t} = ln(I_{i,t})$
- The observed measures proxy the natural log of the factors.

Measurement System

- There are two types of the observed measures, continuous and binary measures.
- The continuous measures are described by:

$$T^k_{i,j,\tau} = \alpha^k_{j,\tau} + \beta^k_{j,\tau}\theta^k_{i,\tau} + u^k_{i,j,\tau} \tag{3}$$

- The binary measures are described by:

$$T^k_{i,j,\tau} = \mathbb{1}[\alpha^k_{j,\tau} + \beta^k_{j,\tau}\theta^k_{i,\tau} + u^k_{i,j,\tau} > 0] \tag{4}$$

where:

- $T^k_{i,j,\tau}$: the jth measure relating to latent factor k for individual i at time τ, $\tau \in \{t-1, t\}$.
- $\theta^k_{i,\tau}$: unobserved/latent factors: $\theta^k_{i,t}$, $\theta^k_{i,t-1}$ $\mathcal{P}^C_i$, $\mathcal{P}^{NC}_i$ and $\mathcal{I}_{i,t}$
- $u_{i,h,t}$: error terms.

Parental Investments

$$
\begin{aligned}
lnI_{i,t} = \quad & \alpha_{1,t} + \alpha_{2,t} ln\Theta^{C}_{i,t-1} + \alpha_{3,t} ln\Theta^{NC}_{i,t-1} + C\alpha_{4,t} ln\Theta^{H}_{i,t-1} \\
& + \alpha_{5,t} lnP^{C}_{i} + \alpha_{6,t} lnP^{NC}_{i} + \alpha_{7,t} X_{i,t} + \alpha_{8,t} Z_{i,t} + v_{i,t}
\end{aligned} \tag{5}
$$

where:

- $X_{i,t}$: child gender, urban/rural residence and the number of siblings.
- $Z_{i,t}$: instrumental variables including the log of wealth index reflecting parental resources, household economic shocks and the log of regional prices.
- $v_{i,t}$: an error term.

Estimation

- Measurement System:

$$\begin{aligned} T^k_{i,j,\tau} &= \alpha^k_{j,\tau} + \beta^k_{j,\tau}\theta^k_{i,t-1} + u^k_{i,j,\tau} \\ T^k_{i,j,\tau} &= \mathbb{1}[\alpha^k_{j,\tau} + \beta^k_{j,\tau}\theta^k_{i,\tau} + u^k_{i,j,\tau} > 0] \end{aligned} \tag{6}$$

- Parental Investments:

$$\begin{aligned} lnI_{i,t} = \quad & \alpha_{1,t} + \alpha_{2,t} ln\Theta^C_{i,t-1} + \alpha_{3,t} ln\Theta^{NC}_{i,t-1} + C\alpha_{4,t} ln\Theta^H_{i,t-1} \\ & + \alpha_{5,t} lnP^C_i + \alpha_{6,t} lnP^{NC}_i + \alpha_{7,t} X_{i,t} + \alpha_{8,t} Z_{i,t} + v_{i,t} \end{aligned} \tag{7}$$

- Production functions:

$$\begin{aligned} \Theta^k_{i,t} = \quad & [\gamma^k_{1t}(\Theta^C_{i,t-1})^{\rho^{tk}} + \gamma^k_{2t}(\Theta^{NC}_{i,t-1})^{\rho^{tk}} + \gamma^k_{3t}(\Theta^H_{i,t-1})^{\rho^{tk}} + \gamma^k_{4t}(I_{i,t})^{\rho^{tk}} \\ & \gamma^k_{5t}(P^C_i)^{\rho^{tk}} + \gamma^k_{6t}(P^{NC}_i)^{\rho^{tk}}]^{1/\rho^{tk}} e^{X'_{i,t}\delta^k_t + A^k_t + \mu^k v_{i,t} + \varepsilon^k_{i,t}} \end{aligned} \tag{8}$$

Estimation

- The distributions of the log factors, $f(\theta), f(\mathcal{P}), f(\mathcal{I})$, are assumed to be distributed as a mixture of two normals.
- Model estimation consists of two steps.
  - First step: estimate the measurement system to recover the parameters $\beta^{C}_{j,\tau}$, $\beta^{NC}_{j,\tau}$, $\beta^{H}_{j,\tau}$, β^{PC}_{j}, β^{PNC}_{j}, $\beta^{I}_{j,t}$, α^{C}_{j}, α^{NC}_{k}, α^{PC}_{j}, α^{PNC}_{j}, $\alpha^{I}_{j,t}$ and the latent factor distributions
  - Second step: use the estimated parameters from the first step to take individual-specific draws and use these draws as observable data to estimate investment and production functions.

Estimation

- The likelihood of all the observed measures conditional on θ_i:

$$L(T_i|\theta_i) = \prod_{k=1}^{K} \prod_{j=1}^{J} \prod_{\tau=1}^{T} f(T_{i,j,\tau}^{k}|\theta_i) \tag{9}$$

Where $\theta_i = [\theta_{i,\tau}^{k}]' = [\theta_{i,\tau}^{C}, \theta_{i,\tau}^{NC}, \theta_{i,\tau}^{H}, \mathcal{P}_i^{C}, \mathcal{P}_i^{NC}, \mathcal{I}_{i,\tau}]'$

- The log-likelihood function:

$$\mathcal{L}(\Psi) = \sum_{i=1}^{N} lnL(T_i) = \sum_{i=1}^{N} ln\left(\int L(T_i|\theta_i)dF(\theta)\right) = \sum_{i=1}^{N} ln\left(\int L(T_i|\theta_i)f(\theta)d\theta\right) \tag{10}$$

Where:

  - $f(\theta) = \sum_{c=1}^{2} \tau_c f(\theta|\mu_c, \Omega_c)$, where μ_c, Ω_c and τ_c are the mean, covariance and the mixture probability of the two normals.
  - Ψ is all the parameters of the model, $\Psi = \{\alpha, \beta, \sigma, \tau_c, \mu_c, \Omega_c\}$.

Estimation

- The log-likelihood function is estimated using the minorization-maximization algorithm.
- Given the unobservable nature of the factors, the log-likelihood function is constructed by integrating over the distributions of the unobservable factors.

Data and variables

- Data is from Round 2 (at age 12) and Round 3 (at age 15) of the Young Lives survey for the Older cohort in Vietnam.

Latent factors		**Observed variables**
Child's cognitive skills - Round 2	θ_2^C	1. PPVT test 2. Math Test 3. Reading level 4. Writing level
Child's cognitive skills - Round 3	θ_3^C	1. PPVT test 2. Math Test 3. Cloze test
Child's noncognitive skills - Round 2 and Round 3	θ_2^{NC}, θ_3^{NC}	1. Self-esteem score 2. Self-efficacy score 3. Self-respect and inclusion score
Child's health - Round 2 and Round 3	θ_2^H, θ_3^H	1. Child height for age z-score 2. Child weight 3. How is child health?

Data and variables

Latent factors		Observed variables
Parental cognitive skills	$\mathcal{P}^C$	1. Mother's years of education 2. Father's years of education
Parental noncognitive skills	$\mathcal{P}^{NC}$	1. Self-esteem score 2. Self-efficacy score 3. self-respect and inclusion score
Parental Investments	$\mathcal{I}_3$	1. Expenditure on the Young Lives child 2. Number of hours studying outside school as a proxy for the time that parents dedicate to the child 3. Quality of relationship between child and parents

Estimates: Determinants of Parental Investments

Variables	Parental investments
Child's cognitive skills, age 12	-0.024
	(0.030)
Child's noncognitive skills, age 12	-0.102
	(0.076)
Child's health, age 12	0.218***
	(0.032)
Parental cognitive skills	0.402***
	(0.057)
Parental noncognitive skills	0.172***
	(0.066)
Wealth index	0.177***
	(0.045)
Observations	*961*

Note: Standard errors in parentheses based on 100 bootstrap replications of the entire estimation process; *** p<0.01, ** p<0.05, * p<0.1.
- Selected covariates are reported.

Estimates of Production Functions

	Cognitive skills, age 15	Noncognitive skills, age 15	Health, age 15
Child's cognitive skills at age 12	0.613***	0.001	0.064**
	(0.044)	(0.039)	(0.029)
Child's noncognitive skills at age 12	0.145***	0.339***	0.130**
	(0.055)	(0.077)	(0.053)
Child's health at age 12	0.042*	0.065	0.670***
	(0.025)	(0.053)	(0.073)
Parental Investments	0.261***	0.505***	0.346**
	(0.075)	(0.167)	(0.170)
Parental cognitive skills	0.017	-0.005	-0.184
	(0.041)	(0.122)	(0.193)
Parental noncognitive skills	-0.079	0.096	-0.026
	(0.050)	(0.081)	(0.065)
A_t	-0.010	0.211***	0.205***
	(0.024)	(0.031)	(0.024)
Control Function (investment residuals)	-0.126*	-0.132	-0.420**
	(0.068)	(0.262)	(0.179)
Complementarity(ρ)	-0.168	-1.851**	0.034
	(0.117)	(0.757)	(0.327)
Elasticity of substitution	0.856***	0.351*	1.036***
	(0.141)	(0.211)	(0.195)
Observations		*961*	

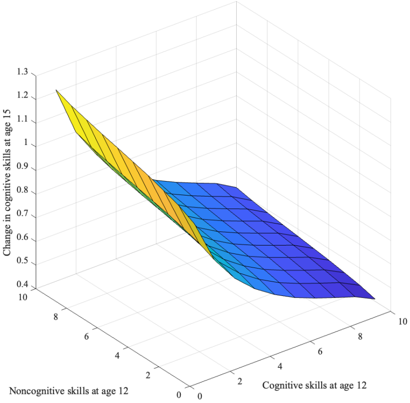
Note: Standard errors in parentheses based on 100 bootstrap replications of the entire estimation process; *** p<0.01, ** p<0.05, * p<0.1.

Estimates of Production Functions - Marginal Effects

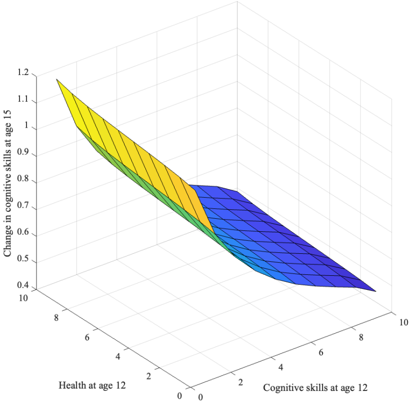
	Cognitive skills, age 15	Noncognitive skills, age 15	Health, age 15
Child's cognitive skills at age 12	0.614***	0.002	0.064**
	(0.043)	(0.039)	(0.029)
Child's noncognitive skills at age 12	0.145***	0.316***	0.130**
	(0.055)	(0.076)	(0.052)
Child's health at age 12	0.042*	0.085	0.670***
	(0.025)	(0.062)	(0.065)
Parental investments	0.261***	0.513***	0.346**
	(0.075)	(0.162)	(0.169)
Parental cognitive skills	0.017	-0.008	-0.184
	(0.042)	(0.123)	(0.194)
Parental noncognitive skills	-0.079	0.093	-0.026
	(0.050)	(0.078)	(0.067)
Observations		*961*	

Note: Standard errors in parentheses based on 100 bootstrap replications of the entire estimation process; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Cognitive Skills: Self-productivity ($\partial\Theta_t^C/\partial\Theta_{t-1}^C$)

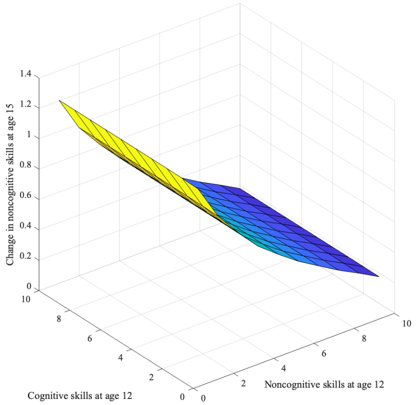


(a)

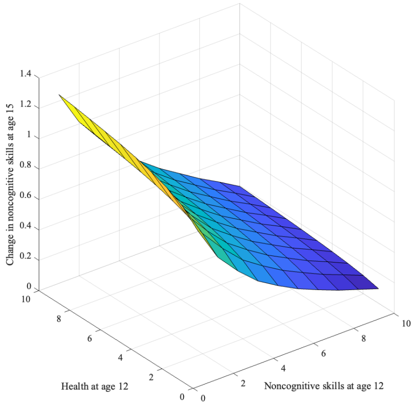


(b)

Noncognitive Skills: Self-productivity $(\partial\Theta_t^{NC}/\partial\Theta_{t-1}^{NC})$

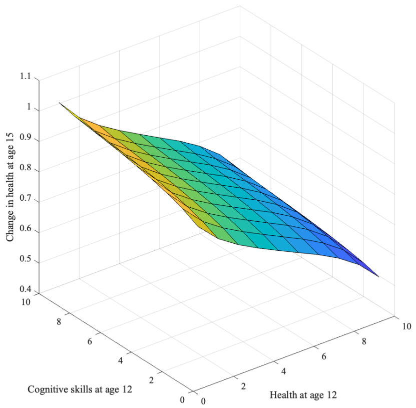


(a)

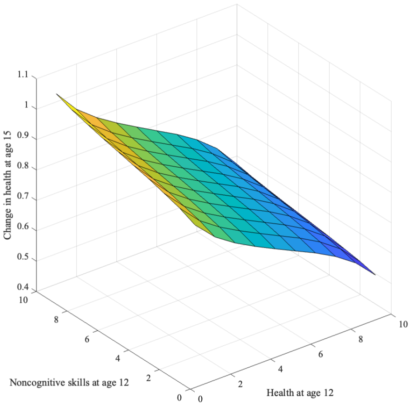


(b)

Health: Self-productivity ($\partial\Theta_t^H/\partial\Theta_{t-1}^H$)



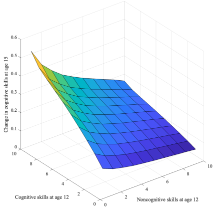
(a)

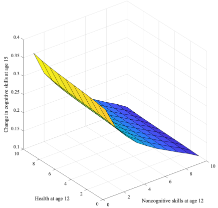


(b)

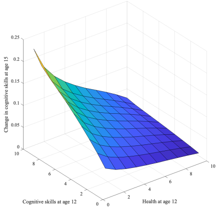
Cognitive skill: Cross-productivity

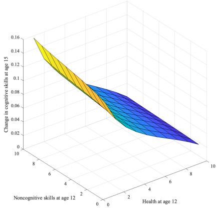
- Cross-productivity from Noncognitive Skills to Cognitive Skills $(\partial\Theta^C_t/\partial\Theta^{NC}_{t-1})$





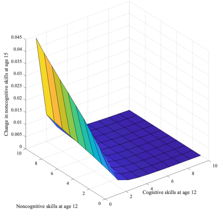
- Cognitive Skills: Cross-productivity from Health to Cognitive Skills $(\partial\Theta^C_t/\partial\Theta^{H}_{t-1})$

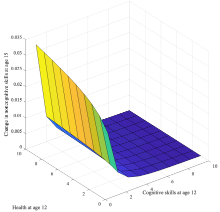




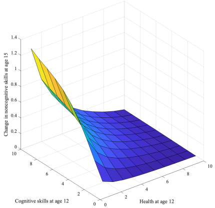
Noncognitive skill: Cross-productivity

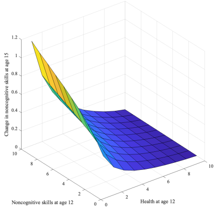
- Cross-productivity from Cognitive Skills to Noncognitive Skills $(\partial\Theta_t^{NC}/\partial\Theta_{t-1}^{C})$





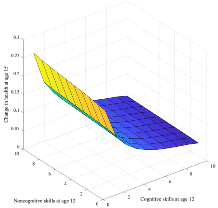
- Cross-productivity from Health to Noncognitive Skills $(\partial\Theta_t^{NC}/\partial\Theta_{t-1}^{H})$

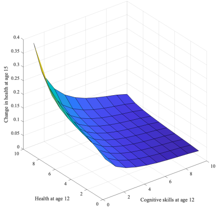




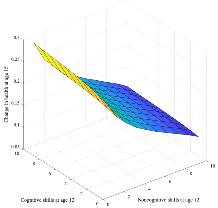
Health: Cross-productivity

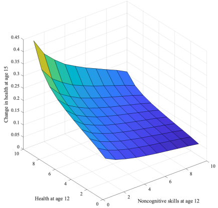
- Cross-productivity from Cognitive Skills to Health $(\partial\Theta_t^H/\partial\Theta_{t-1}^C)$



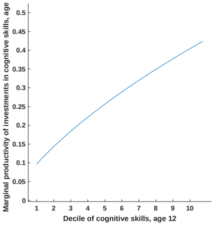


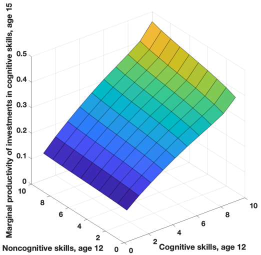
- Cross-productivity from Noncognitive Skills to Health $(\partial\Theta_t^H/\partial\Theta_{t-1}^{NC})$

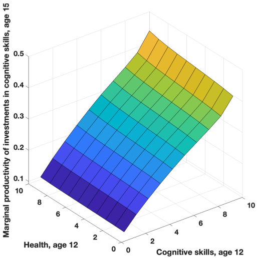




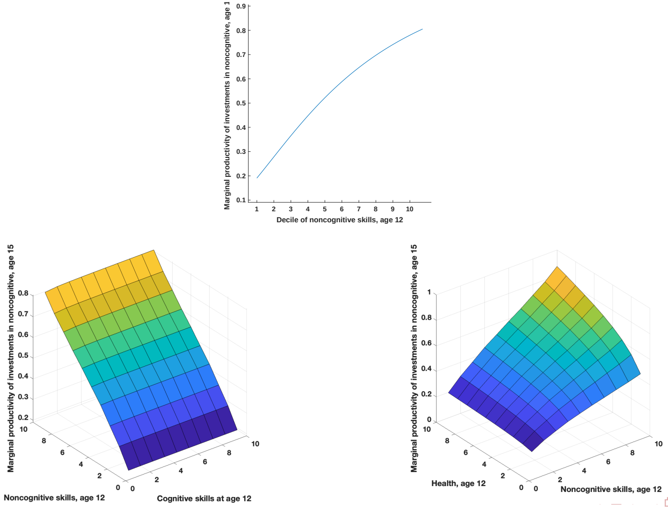
Dynamic Complementarity between Investments and Cognitive Skills



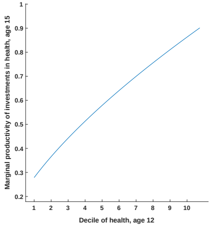


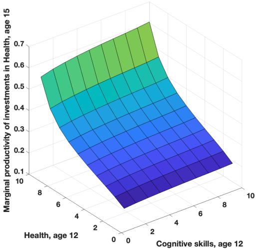


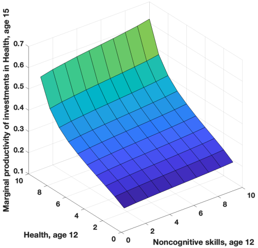
Dynamic Complementarity between Investments and Noncognitive Skills



Dynamic Complementarity between Investments and Health







Conclusions

- Self-effects are present and strong in the production of all human capital dimensions.
- There is the existence of cross-productivity from noncognitive skills and health to cognitive skills and from cognitive and noncognitive skills to health.
- Parental investments strongly and directly affect the accumulation of skills and health
- There are strong dynamic complementarities between parental investments and three dimensions of human capital.

Conclusions

- The self-effects, cross-effects and dynamic complementarities together become a dynamic multiplier effect mechanism of human capital accumulation.
- Parental skills and resources are positively associated with parental investments.
- These effects could lead to substantially different growth rate of human captial and substantial increases in inequality.
- These results indicate the importance of interventions by boosting investments at early ages that can alter child development path, especially for disadvantaged children.
- Lack of parental investments can seriously hinder the development of a child.

Thank you!