Complementarities in High School and College Investments

John Eric Humphries (Yale University)
Juanna Schrøter Joensen (University of Chicago & SSE)
Gregory F. Veramendi (LMU Munich)

Dynamic Structural Econometrics 2021
University of Bonn

August 19, 2021

Do Skills Beget Skills?

- Large literature on early childhood skill formation asking:
  - Are there complementarities between skills and investment?
  - Are dynamic complementarities important for understanding inequality?

Do Skills Beget Skills?

- Large literature on early childhood skill formation asking:
  - Are there complementarities between skills and investment?
  - Are dynamic complementarities important for understanding inequality?

This paper: Are complementarities important later in the life cycle?

- Focused on secondary and post-secondary investments
- Consider three components: ability, HS investments, and college investments
- Are investments specialized or heterogeneous?

Contribution

1. We study complementarities in schooling investments in Sweden

- **Construct a novel administrative dataset of ∼100k males:**
    - ▷ High quality ability measures from military enlistment data
    - ▷ Enrollment and grade data from centralized education system
    - ▷ Rich background data on family, health records, birth records
    - ▷ Labor Market records

Contribution

1. We study complementarities in schooling investments in Sweden

2. Non-parametric evidence on complementarities between ability and schooling

- **Estimate non-parametric or linear models with latent ability to show:**
  - Absolute and differential sorting into HS track, college enrollment, and graduation
  - Non-parametric earnings variance decomposition: three components important but highly dependent
  - Strong complementarities between abilities and majors in earnings

Contribution

1. We study complementarities in schooling investments in Sweden

2. Non-parametric evidence on complementarities between ability and schooling

3. Develop a Roy model of high school investment, college investments, and labor market outcomes

- **The model:**
  - Jointly model a sequence of education decisions and long-run outcomes
  - Use modeled latent heterogeneity as well as choice-specific sources of exogenous variation
  - Accounts for additional unobserved heterogeneity

Contribution

1. We study complementarities in schooling investments in Sweden

2. Non-parametric evidence on complementarities between ability and schooling

3. Develop a Roy model of high school investment, college investments, and labor market outcomes

4. Use the model to study complementarities between high school and college investments ▸ Literature

- **Using the model, we find:**
  - More challenging HS tracks increase college enrollment and graduation
  - Strong complementarities between HS track and abilities
  - Find both positive and negative dynamic complementarities between HS and college investments
  - We consider two policies: Marginal incentives for more STEM courses and eliminating vocational tracks

Brief Review of the Literature

1. **Quasi-experimental literature:**
   - **College:** Kirkebøen, Leuven, and Mogstad (2016); Hastings, Neilson, and Zimmerman (2013)
   - **High School**: Altonji (1995), Levine and Zimmerman (1995), Rose and Betts (2004), Goodman (2012), Joensen and Nielsen (2009, 2016)
2. **Structural Dynamic Discrete Choice Literature:**
   - Arcidiacono (2004), Beffy, Fougere and Maurel (2012), Kinsler and Pavan (2014)
3. **Reduced-form literature:**
   - Berger (1988), Altonji (1993), see Altonji et al (2012, 2015) for review
4. **Literature on non-cognitive abilities:**
   - Heckman and Rubinstein (2001); Heckman et al. (2006); Lindqvist and Vestman (2011); Heckman et al. (2014); Weinberger (2014); Borghans et al. (2016); Deming (2017)

Data and Outcomes

Sample

- 96,949 Swedish males born in 1974-76 who graduated high school

Conscription Exam

- Measures of IQ, psychological aptitude , physical ability, and health

Education

- 9th grade, High school, college enrollment, credit and degree registers
- HS track categorized into three levels: Vocational, Non-STEM Academic, STEM.
- Majors are categorized into 12 choices for academic programs
  - 2-3 year: non-STEM, STEM, Business, Health
  - 4+ year: Education; Humanities; Social Science; Science, Math, and Computer Science; Engineering; Medicine; Business; Law
- Initial major choice: First enrollment ▸ Figure
- Final major choice: major/level of last degree, last enrollment if no degree

Labor Market Outcomes

- Average wages 2010-2013 (34-39 years old)
- Present value of after-tax income

High School Investment

High School: Different tracks have different outcomes

College and Labor Market Outcomes by High School Track

	High School Track		
	Vocational	Academic	STEM
College Outcomes:			
Enroll College (academic)	0.15	0.53	0.82
Enroll 4-year STEM	0.15	0.13	0.46
Grad rate (4yr, cond on enroll)	0.52	0.60	0.67
Labor Market Outcomes:			
Monthly Wages at 36 (USD)	$4,611	$5,839	$6,363

Notes: Authors' calculations using Swedish administrative data. Data include Swedish men born in 1974-1976.

▸ Figure: College Enrollment and Graduation ▸ Figure: College Major

High School: Different tracks invest in different skills

Curriculum of Academic High School Tracks

	High School Courses		
High School Track	Math/Sci/Tech	Social Sci	Languages/Arts
Academic non-STEM			
Business line	0.125	0.156	0.313
Social Science line	0.203	0.297	0.391
Humanities line	0.141	0.297	0.453
Academic STEM			
Technical line	0.563	0.109	0.219
Science line	0.406	0.172	0.313

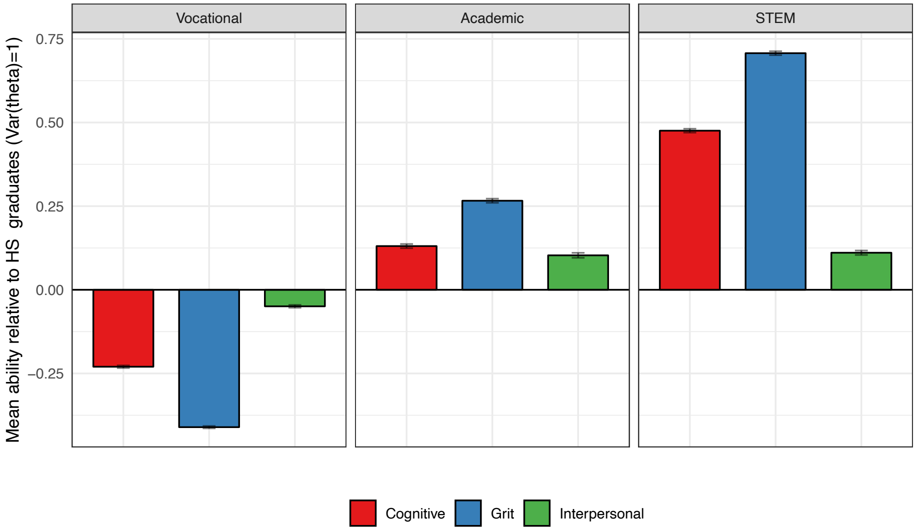
Notes: The average fraction of time devoted to each set of courses in the core curricula over the 3 year duration of each academic high school line.

Ability Sorting in High School and College

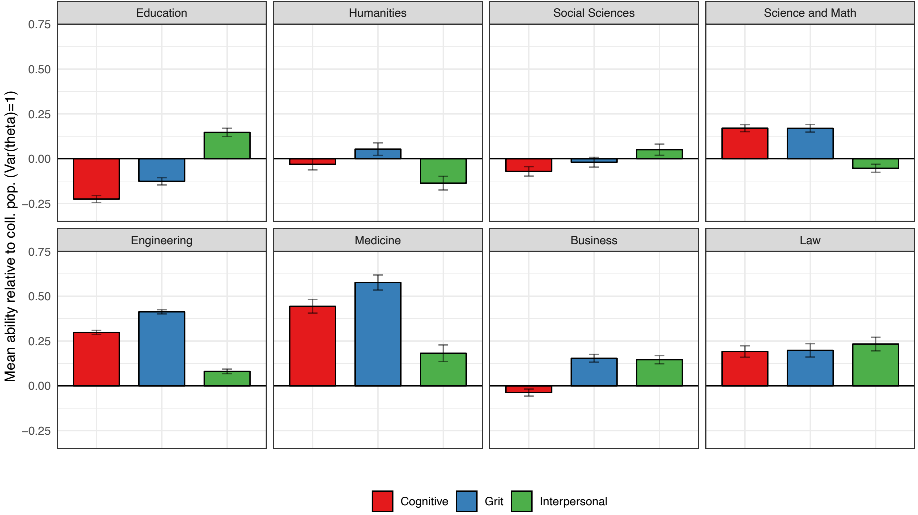
Baseline Latent Ability Factors

- Estimate three 9th grade latent ability factors:
  - cognitive
  - interpersonal
  - grit
- Based on military conscription exams and 9th grade educational outcomes
- Account for schooling and background in measurement system
- Identification based on an extension of Heckman, Hansen & Mullen (2004)
- Validate interpretation using auxiliary measures from survey data

Sorting into HS Track



Sorting into Final College Major



Complementarities Between Abilities, Education, and Earnings

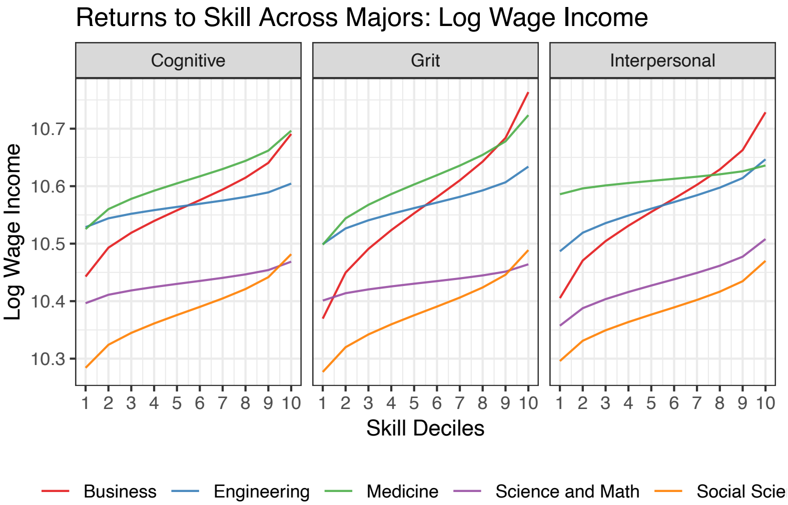
Earnings Equations

- Estimate earnings equations for 15 final schooling states:

$$Y_{isk} = \beta_{sk}^{Y} \boldsymbol{X}_i + \boldsymbol{\lambda}_{sk}^{Y} \boldsymbol{\theta}_i + \eta_{isk}.$$

- $\boldsymbol{X}$ are observables on demographics and family background
- $\boldsymbol{\theta}$ latent abilities
- Parameter of interest: $\boldsymbol{\lambda}_{sk}^{Y}$

Expected Earnings by Ability



PV Disposable Income

THE ECONOMETRIC MODEL

The Econometric Model

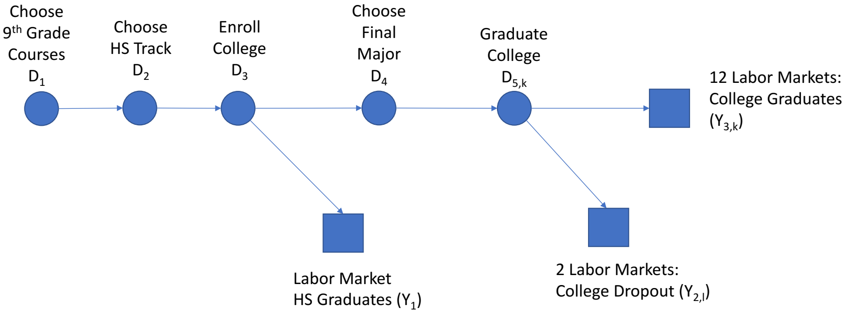
- Goal: Estimate model with dynamic complementarities
  - ▷ Requires causal inference of sequence of educational choices
  - ▷ Do not take a stand on optimization problem of agents. (*ex post* treatment effects)
  - ▷ Consider only policies that change educational choices
  - ▷ Characterize conditional choice probabilities and identify marginal individuals at decision nodes

The Econometric Model

- Goal: Estimate model with dynamic complementarities
  - Requires causal inference of sequence of educational choices
  - Do not take a stand on optimization problem of agents. (*ex post* treatment effects)
  - Consider only policies that change educational choices
  - Characterize conditional choice probabilities and identify marginal individuals at decision nodes

- Develop Generalized Roy Model
  - For each individual, there are 15 potential outcomes Y_{ks}
  - Approximate educational decisions using observables, latent factors, and random effect
    - Assumes that there is a function $f(\boldsymbol{X}, \boldsymbol{\theta}, \upsilon, \epsilon)$ that appoximates agents decisions and state space
  - Account for unobserved heterogeneity beyond latent factors
  - Estimation includes exclusion restrictions at each margin

Sequential Decision Model



Control Variables and Instruments Used in the Analysis

Variables	Measurement Equations	Choice	Income
Mother's Education (indicators)	x	x	x
Father's Education (indicators)	x	x	x
Mother's Family Income (quadratic)	x	x	x
Parent's Married	x	x	x
Mother's age at childbirth	x	x	x
Birth cohort[a]	x	x	x
Strength	x	x	x
Fitness	x	x	x
9th grade and High School track	x	x	x
High School GPA		x	
Enrollment Major		x	
Instruments			
Within-School-Across-Cohort		x	
College Distance		x	

CAUSAL EFFECT OF EDUCATION

The Effects of Education

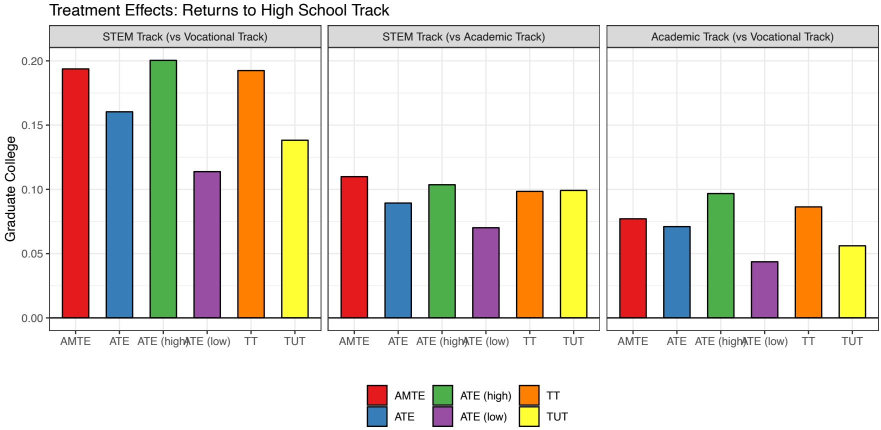
- We use this model to generate mean treatment parameters
- **ATE**: average effect of the treatment

$$\Delta^{ATE} \equiv \int \int \mathbb{E}[Y_{s'} - Y_s | X = x, \theta = t] dF_{X,\theta}(x, t)$$

- We also estimate TT, TUT, AMTE
- TE by final schooling level compared to HS graduates
- Estimate heterogeneous TE depending on latent abilities

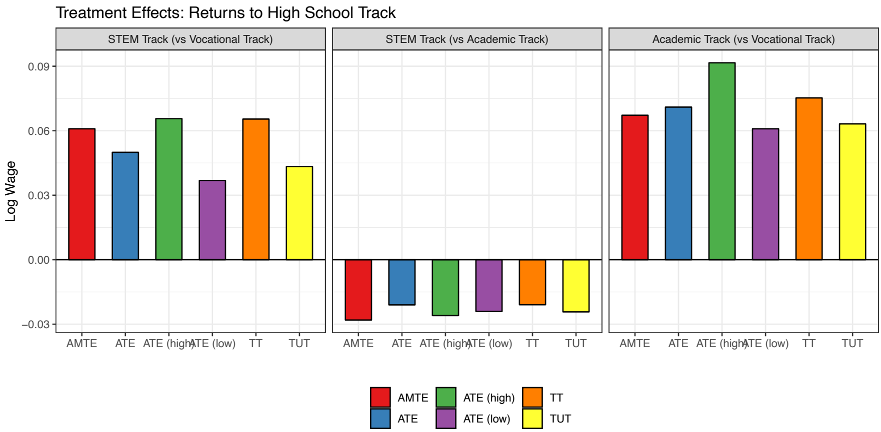
HIGH SCHOOL TRACK

Treatment Effect on College Graduation by HS Track



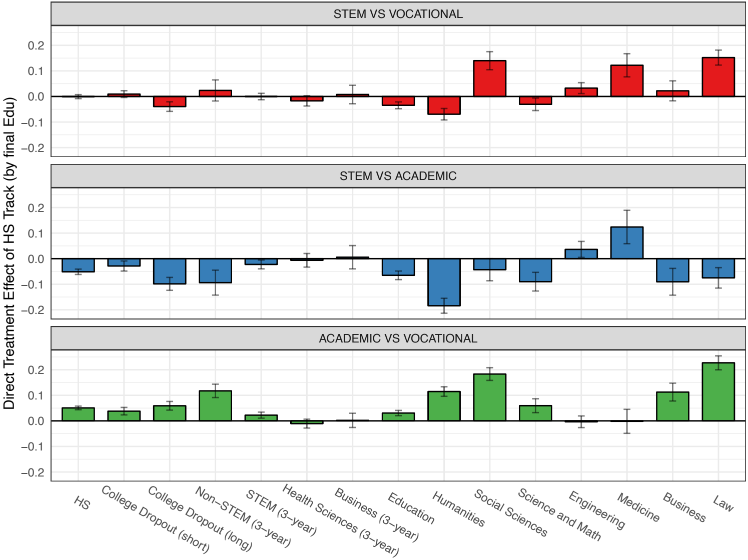
▶ TE by Cognitive ▶ TE by Interpersonal ▶ TE by Grit

Treatment effect on Wage by HS Track



▶ TE by Cognitive ▶ TE by Interpersonal ▶ TE by Grit

Complementarities between HS Track and Final Schooling State



Conclusions

Conclusions:

- Find strong complementarities between abilities, HS investments and college investments
  - Specialization of investments can lead to positive and negative dynamic complementarities

Conclusions

Conclusions:

- Find strong complementarities between abilities, HS investments and college investments
  - Specialization of investments can lead to positive and negative dynamic complementarities

- Implications for Policy
  - Policies that target specific populations likely have highest returns
  - *e.g.* Joensen and Nielsen (2009) and Joensen and Nielsen (2016)

Conclusions

Conclusions:

- Find strong complementarities between abilities, HS investments and college investments
  - Specialization of investments can lead to positive and negative dynamic complementarities
- Implications for Policy
  - Policies that target specific populations likely have highest returns
  - *e.g.* Joensen and Nielsen (2009) and Joensen and Nielsen (2016)
- Implications for broader secondary and post-secondary literature
  - Sequential choice model changes interpretion of LATE and RD designs
- Current work: adding college application process to model

The Econometric Model: Schooling

- Students choose from among hundreds of programs
- Preferences are likely heterogeneous by geographic region and scholastic aptitude
- Students list up to 12 major-college alternatives on their applications

$$D_i^1(\mathcal{L}_i) = \arg\max_{l \in \mathcal{L}_i^1} \{I_{il}\}$$

$$D_i^2(\mathcal{L}_i) = \arg\max_{l \in \mathcal{L}_i^2} \{I_{il}\}$$

$$\ldots$$

  where $D_i^j(\mathcal{L}_i)$ denotes individual i's jth ranked choice given their choice set $\mathcal{L}_i^j$
- Goal: Estimate exploded nested logit model where latent utility of choice l is:

$$I_{il} = f_k(\boldsymbol{X}_i, \boldsymbol{Z}_i, \boldsymbol{\theta}_i) + \delta_{il} + \varepsilon_{il}.$$

Nested Logit of College Choices

- Choice probability can be decomposed into marginal and conditional probabilities

$$P\left[D_i^1 = l\right] = P\left[D_i^1 = l \middle| D_i^1 \in B_{ik}\right] P\left[D_i^1 \in B_{ik}\right],$$

where

$$P\left[D_i^1 \in B_{ik}\right] = \frac{e^{f_k(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\theta})+\lambda_k H_{ik}}}{\sum_{j=1}^{K} e^{f_j(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\theta})+\lambda_j H_{ij}}} \tag{1}$$

$$P\left[D_i^1 = l \middle| D_i^1 \in B_{ik}\right] = \begin{cases} \frac{e^{\delta_{il}/\lambda_k}}{\sum_{j\in B_{ik}} e^{\delta_{ij}/\lambda_k}} & \text{if } l \in \mathcal{L}_i \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $H_{ik} = \ln \sum_{j\in B_{ik}} e^{\delta_{ij}/\lambda_k}$ is the scaled expected utility of nest k and $\lambda_k \in (0, 1]$ is a parameter that describes the amount of correlation between ε_{il} within nest k.

Nested Logit of College Choices

- Assumption 1: Individuals in geographic-GPA bin (g_i) have same preferences within nest: $\delta_{il} \equiv \delta_l(g_i)$ and $H_{ik} \equiv H_k(g_i, GPA_i)$
- Assumption 2: Consideration set depends on individual GPA_i: $B_{ik} \equiv B_k(GPA_i)$
- The expected utility:

$$
\begin{aligned}
H_k(g_i, GPA_i) &= \ln \sum_{j \in B_k(GPA_i)} e^{\delta_j(g_i)/\lambda_k} \\
&= \ln \left[\left(\sum_{j \in B_k} e^{\delta_j(g_i)/\lambda_k} \right) \frac{\sum_{j \in B_k(GPA_i)} e^{\delta_j(g_i)/\lambda_k}}{\sum_{j \in B_k} e^{\delta_j(g_i)/\lambda_k}} \right] \\
&= \ln \left[\left(\sum_{j \in B_k} e^{\delta_j(g_i)/\lambda_k} \right) \left(P\left[D_i^1 \in B_k(GPA_i) | D_i^1 \in B_k, g_i\right] \right) \right] \\
&= H_k(g_i) + \ln \left(P\left[D_i^1 \in B_k(GPA_i) | D_i^1 \in B_k, g_i\right] \right).
\end{aligned}
$$

Exploded Nested Logit of College Choices

- Once a student adds program l to list it must be removed from choice set
- For example, second choice $H_k^2(g_i, GPA_i)$ after choosing l' in nest k' is

$$H_k^2(g_i, GPA_i) = \begin{cases} H_k(g_i, GPA_i) & \text{if} \quad k \neq k' \\ H_k(g_i, GPA_i) + \ln\left(1 - P\left[D_i^1 = l' \middle| D_i^1 \in B_{k'}(GPA_i), g_i\right]\right) & \text{if} \quad k = k' \end{cases}$$

- Given are assumptions, we can estimate $P\left[D_i^1 = l' \middle| D_i^1 \in B_{k'}(GPA_i), g_i\right]$ non-parametrically outside the model
- Finally, we can use the geographic-GPA specific program shares to estimate the outer nest only
    - Expected utility $H_k(g_i)$ is estimated non-parametrically using indicators

Motivation for Model of Education Choices

- Following Aguirregabiria and Mira(2010), consider the model where students observe state variable $\boldsymbol{s}_t$ and choose d_t to maximize expected utility:

$$\mathbb{E}\left[\sum_{k=0}^{T-k} \beta^k U(d_{t+k}, \boldsymbol{s}_{t+k} \mid d_t, \boldsymbol{s}_t)\right]$$

- The student's dynamic programming problem can then be written as:

$$V(\boldsymbol{s}_t) = \max_{d_t \in D_t} \left(U(d_t, \boldsymbol{s}_t) + \beta \int V(\boldsymbol{s}_{t+1}) dF(\boldsymbol{s}_{t+1} \mid d_t, \boldsymbol{s}_t) \right).$$

- The choice-specific value function is given by

$$v(d_t, \boldsymbol{s}_t) = U(d_t, \boldsymbol{s}_t) + \beta \int V(\boldsymbol{s}_{t+1}) dF(\boldsymbol{s}_{t+1} \mid d_t, \boldsymbol{s}_t).$$

- Assume $\boldsymbol{s}_t = \{\boldsymbol{x}_t, \theta, \boldsymbol{\epsilon}_t\}$, where $\boldsymbol{x}_t$ observed by researcher, θ is a set of persistent state variables known by the student but unobserved by researcher, and $\boldsymbol{\epsilon}_t$ are transient shocks observed by students at time t, but unobserved by researcher.
- Other observable outcomes (e.g. earnings): $y_t = Y(d_t, \boldsymbol{s}_t)$.

Model of Education Choices

- Assumptions common in the dynamic discrete choice literature:
  - Unobservables are iid over time and across students ($\boldsymbol{\epsilon} \in G_\epsilon$).
  - Transition of state variables depends on decisions and lagged state variables

$$F_x(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t, \theta, \boldsymbol{\epsilon}) = F_x(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t, \theta)$$

- Given these assumptions

$$F(\boldsymbol{x}_{t+1}, \boldsymbol{\epsilon}_{t+1}|d_t, \boldsymbol{x}_t, \boldsymbol{\epsilon}_t, \theta) = F_x(\boldsymbol{x}_{t+1}|d_t, \boldsymbol{x}_t, \theta)G_\epsilon(\boldsymbol{\epsilon}_{t+1})$$

- The choice specific value function can be written as

$$\begin{aligned} v(d_t, \boldsymbol{s}_t) &= U(d_t, \boldsymbol{s}_t) + \beta \int \int V(\boldsymbol{s}_{t+1}) dG_\epsilon(\boldsymbol{\epsilon}_{t+1}) dF_x(\boldsymbol{x}_{t+1}|d_t, \boldsymbol{x}_t, \theta) \\ &= U(d_t, \boldsymbol{s}_t) + \beta \int \bar{V}(\boldsymbol{s}_{t+1}) dF_x(\boldsymbol{x}_{t+1}|d_t, \boldsymbol{x}_t, \theta), \end{aligned}$$

where $\bar{V}(\boldsymbol{s}_{t+1})$ is the integrated value function.

Model of Education Choices

- We can write the probability than an individual chooses action $d_{t,j}$ in period t as

$$Pr(d_{j,t}|\boldsymbol{x}_t, \theta) = \int \mathbf{I}\left\{\arg\max_{d_t}[v_t(d_t, \boldsymbol{x}_t, \theta) + \epsilon_t(d_t)] = d_{j,t}\right\} dG_\epsilon(\boldsymbol{\epsilon}_t).$$

- Many economically relevant counterfactuals can be estimated through simulation w/o explicitly solving the dynamic program or functional form assumptions on utility
- Joint probability of a given set of states and set of actions can be written as:

$$Pr\left(d_0, (d_1, \boldsymbol{s}_1), \ldots, (d_T, \boldsymbol{s}_T) \mid \boldsymbol{s}_0\right) = Pr(d_T \mid \boldsymbol{s}_T)F_{\boldsymbol{s}}(\boldsymbol{s}_T|d_{T-1}, \boldsymbol{s}_{T-1})\ldots Pr(\boldsymbol{d}_1 \mid d_0, \boldsymbol{s}_0)F_{\boldsymbol{s}}(\boldsymbol{s}_1|d_0, \boldsymbol{s}_0)Pr(d_0|\boldsymbol{s}_0)$$

- E.g. we can estimate how does d_t affect d_{t+j}?

$$\begin{aligned} &Pr(d_T \mid \boldsymbol{s}_T)F_{\boldsymbol{s}}(\boldsymbol{s}_T|d_{T-1}, \boldsymbol{s}_{T-1})\ldots Pr(\boldsymbol{d}_{t+1} \mid d_t, \boldsymbol{s}_t)F_{\boldsymbol{s}}(\boldsymbol{s}_{t+1} \mid \text{fix } d_t = 1, \boldsymbol{s}_t) \\ -&Pr(d_T \mid \boldsymbol{s}_T)F_{\boldsymbol{s}}(\boldsymbol{s}_T|d_{T-1}, \boldsymbol{s}_{T-1})\ldots Pr(\boldsymbol{d}_{t+1} \mid d_t, \boldsymbol{s}_t)F_{\boldsymbol{s}}(\boldsymbol{s}_{t+1} \mid \text{fix } d_t = 0, \boldsymbol{s}_t) \end{aligned}$$

Model of Education Choices

- Model earnings as

$$Y_t = y_t\left(d_t, \boldsymbol{x}_t, \theta\right) + \eta_t$$

and

$$\mathbb{E}[Y_t] = \int\int\int y_t(d_t, \boldsymbol{x}_t, \theta)dF_\theta(\theta)dF_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon_t})dF_{\boldsymbol{x}_t}(\boldsymbol{x}_t),$$

- We can estimate

$$\mathbb{E}[Y_t(d_{t-k} = 1)] - \mathbb{E}[Y_t(d_{t-k} = 0)]$$

where

$$\mathbb{E}[Y_t(d_{t-k} = 1)] = \int\int\int y_t(d_t, \boldsymbol{x}_t, \theta)dF_\theta(\theta)dF_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon_t})dF_{\boldsymbol{x}_t}(\boldsymbol{x}_t | \text{fix } d_{t-k} = 1).$$

and

$$\mathbb{E}[Y_t(d_{t-k} = 0)] = \int\int\int y_t(d_t, \boldsymbol{x}_t, \theta)dF_\theta(\theta)dF_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon_t})dF_{\boldsymbol{x}_t}(\boldsymbol{x}_t | \text{fix } d_{t-k} = 0).$$