Education, Lifestyles, and Health Inequality

Jesús Bueren[◇] Josep Pijoan-Mas[☆] Dante Amengual[□]

[◇]Católica Lisbon

[☆]CEMFI, CEPR

[□]CEMFI

IDB - Research Conference

(September 2026)

Introduction

- Health inequalities in the United States are large
- Two important facts:
  a) Strong connection between economic and health inequality
  Kitagawa, Hauser (1973); Pijoan-Mas, Rios-Rull (2014); Chetty et al. (2016)
  b) Growing educational gradients of health inequality
  Preston, Elo (1995); Meara et al. (2009); Montez et al. (2011); Case, Deaton (2015)
  → Reasons not well-understood
- We study to which extent differences in health behaviors across education groups can account for these two facts
  - Health behaviors are an important driver of health outcomes
  McGinnis, Foege (1993); Li et al. (2018); Zaninotto et al. (2020)
  - Individuals with higher education are more likely to adopt health-enhancing behaviors
  Lantz et al. (1998); Cutler, Lleras-Muney (2010)

Objectives

1. Measure the impact of lifestyles on health dynamics and economic outcomes
   - Lifestyle → propensity to engage in different health-related behaviors
2. Understand the joint determination of education and lifestyles
   → Why is there an education gradient of lifestyles?
   a) Do higher educated individuals have higher incentives to adopt protective lifestyles?
   b) Do individuals raised in healthier environments self-select into higher education?
3. Understand the increase in the education gradient of life expectancy
   → Quantify the role played by the increase in the education wage premium

What we do: Data

- There are many indicators of health behavior in HRS and PSID
  (preventive tests, substance abuse, exercise)

- Problems
  - They are imperfectly correlated across individuals and over time
  - Hard to identify their long-run effect on health outcomes

- We identify patterns in health behavior types that we interpret as lifestyles based on:
  - Cross-sectional and panel variation of health behaviors
  - Health dynamics

  → *Types* are permanent, but health behaviors change with age and health status

What we do: Model

- Two different stages

  2) Adulthood — standard life-cycle model with labor income and health risks
     - Education and lifestyle given
     - Choice of consumption and savings
     - → Provides *early life* value of each education and lifestyle choice

  1) Early life — simple static problem
     - Heterogeneous costs of acquiring education and lifestyle
     - Once and for all choice of education and lifestyle
     - → Provides joint distribution of education and lifestyle

- **Complementarities** between education and lifestyle choices
  - + An extra year of life is more valuable with higher consumption possibilities

⇒ **Selection**: individuals from healthier backgrounds tend to pursue higher education.

Health Dynamics and Health Behavior

The Data

- HRS and PSID
  - Unbalanced panels of individuals $i = 1, ..., N$ followed $t = 1, ..., T$ periods
- Demographic information: birth cohort (c), sex (s), education (e), age of entry (a_0)
- Wide array of information on health status and health behavior
  - Health state h_t: self-reported health (good/bad) + death
  - Health behavior $z_{mt} \in \{0, 1\}$:
    - Preventive cancer tests (mammography / prostate check)
    - Cholesterol test
    - Flu shot
    - Heavy drinking (2+ drinks every day)
    - Smoking
    - Exercise

  → Across demographic groups: associated w/ health outcomes

  → At individual level: imperfectly correlated across individuals and over time

[See data]

LATENT TYPES

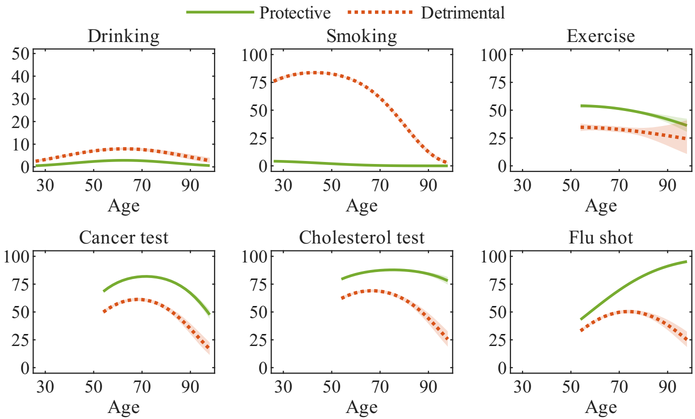
- We assume that observed health behavior is the (noisy) result of a latent time-invariant factor $y \in Y \equiv \{y_1, y_2, \dots\}$
  - We interpret y as the individual lifestyle: propensity to engage in healthy behaviors
- We propose a novel econometric model to
  - Allocate individuals to lifestyles y
  - Measure the importance of lifestyles on health dynamics
- We *jointly estimate* health dynamics and lifestyles using a mixture model:

$$
\begin{aligned}
p(\boldsymbol{z}, \boldsymbol{h} | \mathrm{c}, \mathrm{s}, \mathrm{e}, a_0) &= \sum_{y \in Y} p(\boldsymbol{z}, \boldsymbol{h} | \mathrm{c}, \mathrm{s}, \mathrm{e}, a_0, y)\, p(y | \mathrm{c}, \mathrm{s}, \mathrm{e}, a_0) \\
&= \sum_{y \in Y} p(\boldsymbol{z} | \boldsymbol{h}, \mathrm{c}, \mathrm{s}, \mathrm{e}, a_0, y)\, p(\boldsymbol{h} | \mathrm{c}, \mathrm{s}, \mathrm{e}, a_0, y)\, p(y | \mathrm{c}, \mathrm{s}, \mathrm{e}, a_0) \\
&\simeq \sum_{y \in Y} \underbrace{p(\boldsymbol{z} | \boldsymbol{h}, a_0, y)}_{\text{health behavior}}\; \underbrace{p(\boldsymbol{h} | \mathrm{s}, \mathrm{e}, a_0, y)}_{\text{health dynamics}}\; \underbrace{p(y | \mathrm{c}, \mathrm{s}, \mathrm{e}, a_0)}_{\text{lifestyle distribution}}
\end{aligned}
$$

Results: Lifestyles and health behavior

1. Lifestyles "well" approximated by 2 types: protective and detrimental

Probability of reporting health behaviors by lifestyle



Results: Lifestyles and health dynamics

2. LE at age 50 is 9.0 years larger for protective than for detrimental

LE at age 50 by education and lifestyles: males born in 1950s

	All		HSD		HSG		CG		Δ_eLE (CG-HSD)		
	%	LE	%	LE	%	LE	%	LE	Data	(a)	(b)
All	100.0	**27.7**	100.0	24.1	100.0	26.6	100.0	31.2	7.1	4.5	2.6
PRO	72.1	**30.2**	55.6	27.8	67.7	29.2	86.7	32.4	4.6		
DET	27.9	**21.2**	44.4	19.4	32.3	21.3	13.3	23.5	4.1		
Δ_y	44.2	**9.0**	11.2	8.4	35.4	7.9	73.4	8.9	0.5		

(a) Gradient explained by difference in health dynamics across education groups for given lifestyle, $\Delta_e p(\boldsymbol{h}|\mathrm{e}, y)$
(b) Gradient explained by difference in lifestyles across education groups for given health dynamics, $\Delta_e p(y|\mathrm{e})$

Results: Lifestyles, education, and health dynamics

3. Lifestyles explain almost 40% of the education gradient of LE

LE at age 50 by education and lifestyles: males born in 1950s

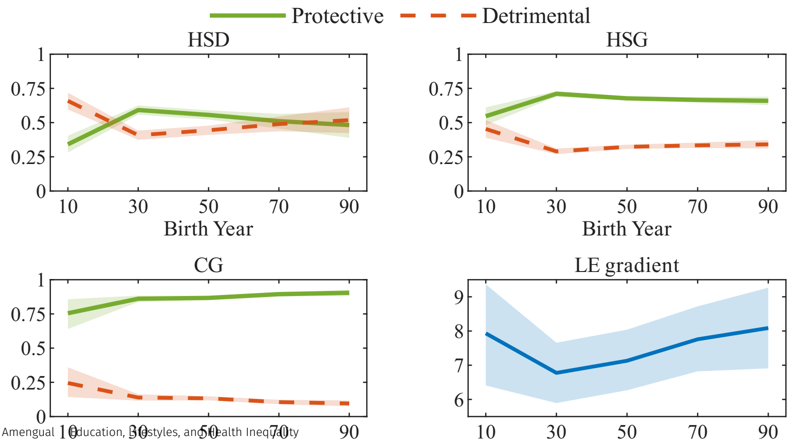
	All		HSD		HSG		CG		Δ_eLE (CG-HSD)		
	%	LE	%	LE	%	LE	%	LE	Data	(a)	(b)
All	100.0	27.7	100.0	24.1	100.0	26.6	100.0	31.2	**7.1**	**4.5**	**2.6**
PRO	72.1	30.2	**55.6**	27.8	**67.7**	29.2	**86.7**	32.4	4.6		
DET	27.9	21.2	**44.4**	19.4	**32.3**	21.3	**13.3**	23.5	4.1		
Δ_y	44.2	9.0	**11.2**	8.4	**35.4**	7.9	**73.4**	8.9	0.5		

(a) Gradient explained by difference in health dynamics across education groups for given lifestyle, $\Delta_e p(\boldsymbol{h}|\mathrm{e}, y)$
(b) Gradient explained by difference in lifestyles across education groups for given health dynamics, $\Delta_e p(y|\mathrm{e})$

Results: Changes across cohorts

4. Education gradient of lifestyles widens over time

Probability of lifestyle at age 50 across cohorts. Males



The Model

The Model

Two different stages

1. Early life
   - Choice of education and lifestyle
2. Life cycle
   - Standard life-cycle incomplete-markets model of consumption and wealth accumulation
   - Health and labor market risks that are related to education and lifestyle

Stage 1: Early life

Set up

- Teenager/parents in cohort c make once-and-for-all simultaneous choices of
  - education $e \in \{\text{HSD}, \text{HSG}, \text{CG}\}$
  - lifestyle $y \in \{\text{DET}, \text{PRO}\}$
- They solve $\max_{e,y} \left\{ V_0^{eyc} - \tau_e - \tau_y \right\}$
  - Value V_0^{eyc} of starting stage 2 with type (e, y, c)
  - Cost τ_e of education e: $\tau_{\text{HSD}} = 0$ | $\tau_{\text{HSG}} = \mu_{\text{HSG}} + \sigma_{\text{HSG}} \epsilon_e$ | $\tau_{\text{CG}} = \mu_{\text{CG}} + \sigma_{\text{CG}} \epsilon_e$
  - Cost τ_y of lifestyle y: $\tau_{\text{DET}} = 0$ | $\tau_{\text{PRO}} = \mu_{\text{PRO}} + \sigma_{\text{PRO}} \epsilon_{\text{PRO}}$
  - *Correlated* bivariate normal: $(\epsilon_e, \epsilon_{\text{PRO}})' \sim N\left(\mathbf{0}, \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)\right)$
- There are complementarities between education and lifestyle choices if:

$$V_0^{\text{CG,PRO}} - V_0^{\text{CG,DET}} > V_0^{\text{HSD,PRO}} - V_0^{\text{HSD,DET}} \Leftrightarrow V_0^{\text{CG,PRO}} - V_0^{\text{HSD,PRO}} > V_0^{\text{CG,DET}} - V_0^{\text{HSD,DET}}$$

[theory]

Stage 2: Life cycle

Model: states and ingredients

State variables

Agent types

- Education: $\mathrm{e} \in \{\text{HSD}, \text{HSG}, \text{CG}\}$
- Lifestyle: $\mathrm{y} \in \{\text{DET}, \text{PRO}\}$
- Cohort: $\mathrm{c} \in \{1930, 1970\}$

Exogenous states

- Age: $t \in \{25, 27, 29, \ldots\}$
- Health: $h_t \in \{h_g, h_b\}$
- Earnings shock: $\zeta_t \in \mathbb{R}$

Endogenous state

- Cash-on-hand: $x_t \in [\underline{x}, \infty)$

Model ingredients

Cohort-specific

- Employment probability: $l_t^{\mathrm{ec}}(\varepsilon, h)$
- Labor earnings: $w_t^{\mathrm{ec}}(\zeta, \epsilon, h)$

Cohort-independent

- Health and survival:$s_t^{\mathrm{ey}}(h), \Gamma_t^{\mathrm{ey}}(h'|h)$
- Medical expenses and insurance: $m_t^{\mathrm{e}}(\xi, h)$
- Taxes and income floor: $T(\cdot)$ and $\underline{x}$
- Utility from consumption and survival

$\Rightarrow$ The solution provides the initial value of each type $(\mathrm{e}, \mathrm{y}, \mathrm{c})$: V_0^{eyc}

Education gradient of lifestyle

Education gradient of lifestyle

Mechanisms

- Why higher educated individuals are more likely to be protective?

- Direct Effect

  1. Income: $w_t^{\text{ec}}(\zeta, \epsilon, h)$ and $l_t^{\text{ec}}(\varepsilon, h)$

  $\Rightarrow$ Generates complementarities: $V_0^{\text{CG,PRO}} - V_0^{\text{CG,DET}} > V_0^{\text{HSD,PRO}} - V_0^{\text{HSD,DET}}$

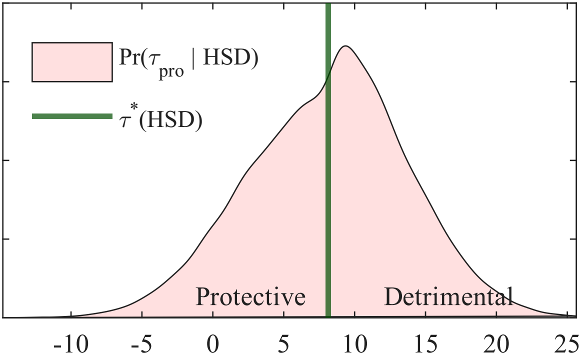
- Induced by complementarities

  2. Selection: $F(\tau_{\text{PRO}} | \text{HSD}) \leq F(\tau_{\text{PRO}} | \text{CG}) \quad \forall \tau_{\text{PRO}}$

Education gradient of lifestyle

Lifestyle Choice for HSD

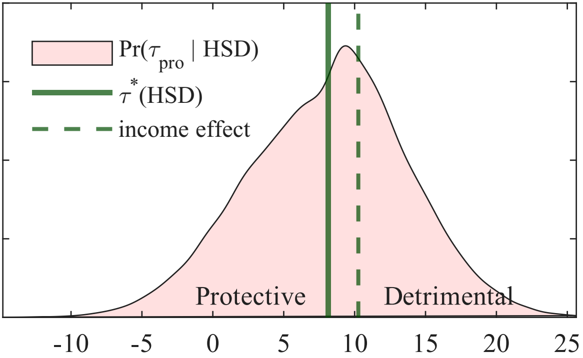
- HSD choose y = PRO iff $\tau_{\mathrm{PRO}} < V_0^{\mathrm{HSD,PRO}} - V_0^{\mathrm{HSD,DET}} \equiv \tau^*(\mathrm{HSD})$



Education gradient of lifestyle

1. Income Effect

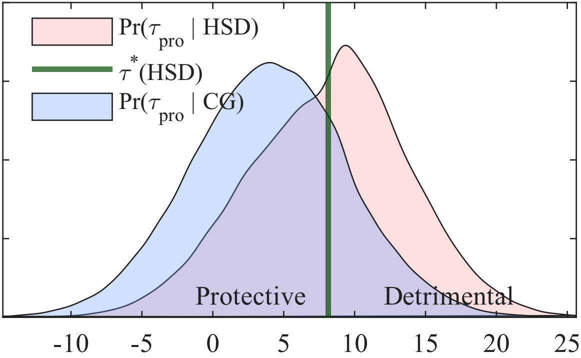
→ If HSD had same income prospects as CG: 16pp more of PRO (1.3 more years of LE)



Education gradient of lifestyle

2. Selection

→ If HSD had same distribution of τ_{PRO} as CG: 31pp more of PRO (2.6 more years of LE)



Education gradient of lifestyle

Summary

Decomposition: 1970s cohort

	Pr(e)		Pr(PRO|e)			LE(e)		
	CG	HSD	CG	HSD	CG-HSD	CG	HSD	CG-HSD
Data	0.36	0.08	0.89	0.51	0.38	31.5	23.7	7.8
Model	0.36	0.10	0.87	0.47	0.39	31.2	23.4	7.8
Same lifestyle			0.87	0.87	0.00	31.2	26.7	4.5
Income			0.87	0.63	0.23	31.2	24.7	6.5
Selection			0.87	0.78	0.09	31.2	25.9	5.2
No correlated costs	0.35	0.10	0.81	0.61	0.20	30.7	24.5	6.1

Policy implications

Policy implications

Feedback between education and lifestyle choices

- A policy directed at either early-life investment can affect the other.
- Given that education and lifestyle choices are complementary:
  - Lowering the cost of one investment increases the incentive to undertake the other:
    1. Lower cost of protective behavior $\Longrightarrow$ healthier lifestyles $\Longrightarrow$ higher educational attainment.
    2. Lower education costs $\Longrightarrow$ more education $\Longrightarrow$ greater adoption of protective behavior.
- This feedback amplifies the reach of each intervention beyond its direct target.

Policy implications

Policy counterfactuals

	Pr(e)		Pr(PRO)	Pr(PRO|e)			LE(e)			
	CG	HSD		CG	HSD	CG-HSD	All	CG	HSD	CG-HSD
Benchmark	0.36	0.10	0.73	0.87	0.47	0.39	28.0	31.2	23.4	7.8
Cost pro - 0.5 sd	0.36	0.09	0.86	0.95	0.67	0.27	29.1	31.9	25.1	6.8
Diff pro	0.01	-0.01	0.14	0.08	0.20	-0.12	1.1	0.7	1.7	-1.0
Cost educ - 0.5 sd	0.55	0.04	0.74	0.84	0.42	0.42	28.8	31.0	22.9	8.1
Diff educ	0.19	-0.06	0.02	-0.02	-0.05	0.03	0.8	-0.2	-0.4	0.2

Conclusions

Conclusions

- We develop an econometric model to identify latent types in lifestyles

  $\Rightarrow$ Differences in lifestyles across education groups account for 40% of the LE gradient (2.6 out of 7.1 years)

- HA model w/ complementarities in education and lifestyle investments

  1. Education gradient of lifestyles
     - Income advantage advantage explain 40% ($\Rightarrow$ 1.3 years)
     - Endogenous + Primitive selection explains 75% ($\Rightarrow$ 2.6 year)
     - Both similarly important
     - Not explained by lifestyles ($\Rightarrow$ 4.5 years)
  2. 2.0 years increase in education gradient of LE between 1930s and 1970s cohorts
     - Direct effect of wage changes explains 4.8 months
     - Indirect effect through selection explains 9.6 months
     - Both together explain 1.2 year

Econometric Model

1. Health Behavior

- How to model $p(\boldsymbol{z}|\boldsymbol{h}, a_0, y)$?
- We assume that, conditional on $(\mathrm{y}_i, h_{it}, a_{it})$ health behaviors $z_{m,it}$ are
  - independent from $h_{it'}$ $\forall t' \neq t$
  - independent between them
  - independent over time
- Hence, we can write the probability of a given history $\boldsymbol{z}_i$ of health behaviors as

$$p(\boldsymbol{z}_i|\boldsymbol{h}_i, \mathrm{y}_i, a_{it}) = \prod_t \prod_m p\left(z_{m,it} \,\middle|\, \mathrm{y}_i, h_{it}, a_{it}\right)$$

- We model each element $p\left(z_{m,it}|\mathrm{y}_i, h_{it}, a_{it}\right)$ as a probit

Back to Mixture Model

Econometric Model

2. Health Dynamics

- How to model $p(\boldsymbol{h}|\mathsf{s},\mathsf{e},a_0,y)$?
- We assume that, conditional on $(\mathsf{s},\mathsf{e},a_0,y)$, the evolution of health outcomes is markovian
- Hence, we can write the probability of a given history $\boldsymbol{h}_i$ of health outcomes as

$$p(\boldsymbol{h}_i|\mathsf{s}_i,\mathsf{e}_i,a_{it},\mathsf{y}_i) = \prod_t p\left(h_{it}|\,\mathsf{s}_i,\mathsf{e}_i,a_{it},\mathsf{y}_i,h_{i,t-1}\right)$$

- We model each element $p\left(h_{it}|\,\mathsf{s}_i,\mathsf{e}_i,a_{it},\mathsf{y}_i,h_{i,t-1}\right)$ as a nested probit
  - First we model survival
  - Next, we model health changes conditional on survival

Back to Mixture Model

Econometric Model

3. Weights

- How to model the mixture weights $p(y|\mathsf{c}, \mathsf{s}, \mathsf{e}, a_0)$?
- The conditional mixture weights are different at each age of entry a_0 because of differential health dynamics by type y
  - Detrimental lifestyles become less frequent at older ages because they die.
- What we do
  - Use the health dynamics to express $p\left(y|\ \mathsf{c}_i, \mathsf{s}_i, \mathsf{e}_i, a_{it}, h_{it}\right)$ as a function of $p\left(y|\ \mathsf{c}_i, \mathsf{s}_i, \mathsf{e}_i, a_{it-1}, h_{it-1}\right)$
  - Estimate the mixture weights at the initial age of 25, $p\left(y|\ \mathsf{c}_i, \mathsf{s}_i, \mathsf{e}_i, 25, h_{i,25}\right)$
- We model each element $p\left(y|\ \mathsf{c}_i, \mathsf{s}_i, \mathsf{e}_i, 25, h_{i,25}\right)$ as a multinomial probit

Back to Mixture Model

Health Behaviour Data

Mean health behavior and 4-year auto-correlation

	Mean					AC	
	HSD	HSG	CG	65-70	75-80	65-70	75-80
Drinking	0.07	0.08	0.07	0.08	0.04	0.46	0.39
Smoking	0.18	0.15	0.07	0.14	0.08	0.70	0.64
Cancer test	0.60	0.71	0.79	0.77	0.72	0.40	0.41
Cholesterol	0.78	0.85	0.89	0.84	0.85	0.30	0.27
Flu shot	0.67	0.74	0.81	0.66	0.76	0.59	0.59
Exercise	0.27	0.38	0.54	0.41	0.38	0.41	0.37

Data from the HRS. HSD: high school dropout; HSG: high school graduate; CG: college graduate; 65-70: sub-sample of individuals aged 65 to 70; 75-80: sub-sample of individuals aged 75 to 80. The last two columns show the autocorrelation (AC) of each health behavior with a 4-year lag.

Back to Data

Health Behaviour Data

Cross correlation health behaviors

	Drinking	Smoking	Cancer test	Cholesterol	Flu shot	Exercise
Drinking	1.00	0.07	0.00	-0.00	-0.03	0.02
Smoking	0.12	1.00	-0.08	-0.07	-0.08	-0.06
Cancer test	-0.02	-0.15	1.00	0.28	0.19	0.10
Cholesterol	-0.02	-0.12	0.36	1.00	0.22	0.07
Flu shot	-0.05	-0.09	0.21	0.24	1.00	0.03
Exercise	0.03	-0.09	0.11	0.03	0.01	1.00

Data from the HRS. Lower diagonal: individuals aged between 65 and 70. Upper diagonal: individuals aged between 75 and 80.

Back to Data

Proposition

- Let the distributions of ϵ_e and ϵ_{PRO} be independent
- Let $F(\tau_{\text{PRO}}|\text{e})$ be the CDF of τ_{PRO} conditional on education choice e
- Let's consider only two education choices $\text{e} \in \{\text{HSD}, \text{CG}\}$
- Proposition:
  - If the choices of education and lifestyle are complementary, the distribution $F(\tau_{\text{PRO}}|\text{HSD})$ first-order stochastically dominates $F(\tau_{\text{PRO}}|\text{CG})$. That is,

$$F(\tau_{\text{PRO}}|\text{HSD}) \leq F(\tau_{\text{PRO}}|\text{CG}) \quad \forall \tau_{\text{PRO}}$$

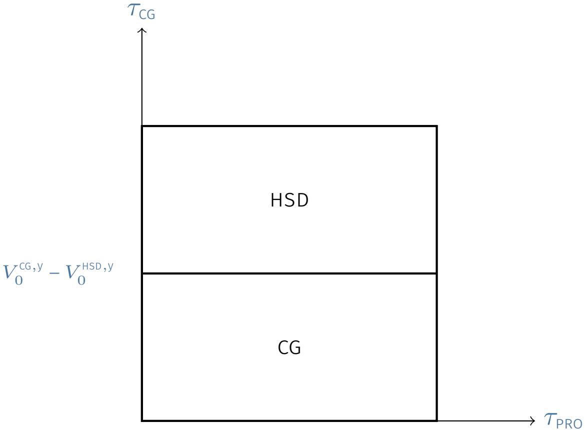
    with strict inequality for some τ_{PRO}.
  - Instead, if the choices of education and lifestyle are independent we have that,

$$F(\tau_{\text{PRO}}|\text{HSD}) = F(\tau_{\text{PRO}}|\text{CG}) \quad \forall \tau_{\text{PRO}}$$

Back to Early Life

Choices: two education levels

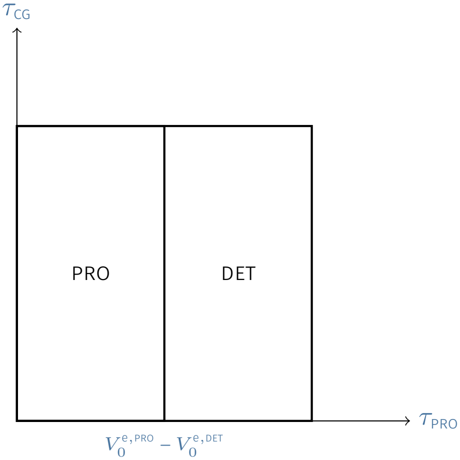
Independent choices



Back to Early Life

Choices: two education levels

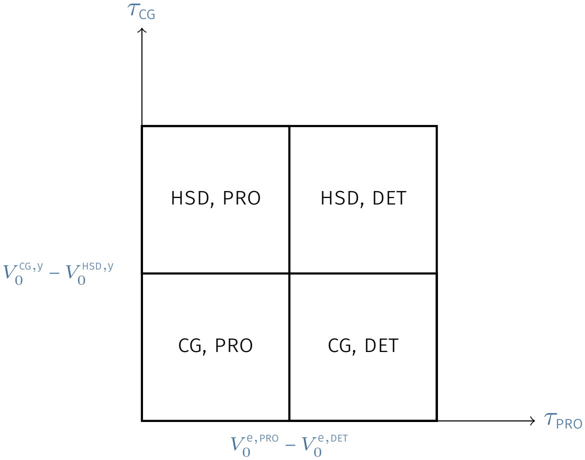
Independent choices



Back to Early Life

Choices: two education levels

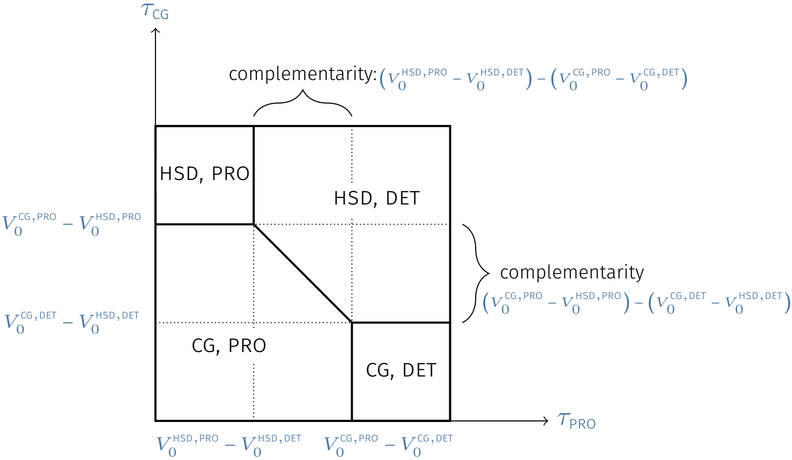
Independent choices



Back to Early Life

Choices: two education levels

Complementarities



Back to Early Life