Education, Healthy Habits, and Inequality

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

[☆] CEMFI

[◇] EUI

[□] CEMFI, CEPR

SAEe, December 2024

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 facts
  - Health behaviors are an important driver of health outcomes
    McGinnis, Foege (1993); Li et al. (2018); Zaninotto et al. (2020)
  - Individuals with higher levels of education are more likely to adopt health-enhancing behaviors
    Lantz et al. (1998); Cutler, Lleras-Muney (2010); Polvinen et al. (2013)

Objectives

1. Measure the impact of lifestyles on health dynamics and economic outcomes

2. Understand the joint determination of education and lifestyles

   → Why is there an education gradient of lifestyles?

   a) Do higher educated individuals have a higher incentives to invest in 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 (*not today*)

Literature on health and economic inequality

- Models with *exogenous health* dynamics

  De Nardi et al. (2010); Ameriks et al. (2020); Bueren (2023); Nakajima, Telyukova (2023)

  Capatina (2015); Braun et al (2019); Hosseini et al (2021); De Nardi et al. (2023)

  → We model endogenous health dynamics

- Models with *endogenous monetary health investments*

  Fonseca et al. (2023); Ozkan (2023); Hong, Pijoan-Mas, Ríos-Rull (2024)

  → Scarce causal evidence of effects of money on health

  → We focus on health-related behaviour

- Models with *endogenous health behaviour investments*

  Cole et al. (2019); Mahler, Yum (2023); Margaris, Wallenius (2023)

  → Joint decision of education and lifestyle: deal with selection

  → Effect of lifestyle identified by long-run health dynamics

  Hai, Heckman (2024)

  → We model different chanels

→ We study long-run changes of economic and health inequality

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 (a_t)

- Wide array of information on health status and health behavior
  - Health state h_t: self-reported health (good/bad) + death
  - Health behavior $z_{mt} \in \{0, 1\}$:
    1. Preventive cancer tests (mammography / prostate check)
    2. Cholesterol test
    3. Flu shot
    4. Heavy drinking (2+ drinks on the day they drink)
    5. Smoking
    6. Exercise

Latent types

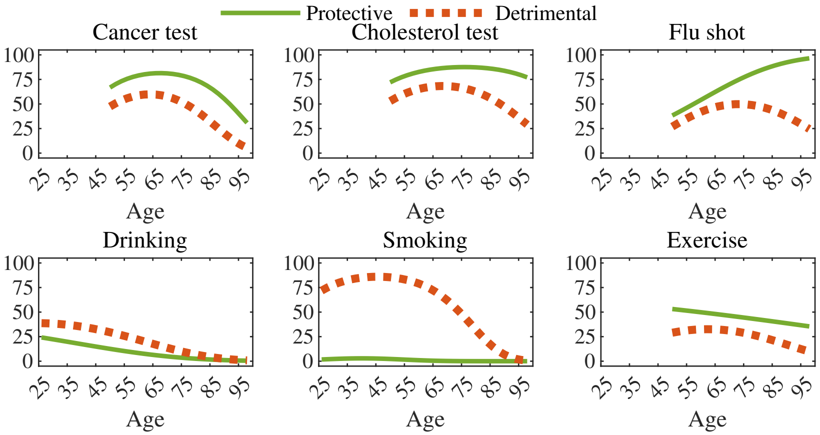
- We assume that observed health behavior is the (noisy) result of a latent time-invariant factor y $\in Y \equiv \{y_1, y_2, ...\}$
  - We interpret y as the individual lifestyle: propensity to engage in healthy behaviors
- We propose a novel econometric model to
  - Allocate individuals to lifestyles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

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

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 8.6 years larger for protective than for detrimental

LE at age 50 across education and lifestyles: males born in 1970s

	All		HSD		HSG		CG		ΔLE (CG-HSD)		
	%	LE	%	LE	%	LE	%	LE	Data	(a)	(b)
All	100	**29.3**	100	24.9	100	28.0	100	32.8	7.9	4.7	3.2
PRO	74.4	**31.5**	44.3	28.5	69.0	30.1	93.6	33.4	4.9		
DET	25.6	**22.9**	55.7	22.0	31.0	23.3	6.4	23.8	1.9		
Δ	48.8	**8.6**	-11.4	6.6	37.9	6.8	87.2	9.6	3.0		

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

Results: Lifestyles, education, and health dynamics

3. Effect of lifestyle on LE larger for the more educated (3 years)

LE at age 50 across education and lifestyles: males born in 1970s

	All		HSD		HSG		CG		ΔLE (CG-HSD)		
	%	LE	%	LE	%	LE	%	LE	Data	(a)	(b)
All	100	29.3	100	24.9	100	28.0	100	32.8	7.9	4.7	3.2
PRO	74.4	31.5	44.3	28.5	69.0	30.1	93.6	33.4	4.9		
DET	25.6	22.9	55.7	22.0	31.0	23.3	6.4	23.8	1.9		
Δ	48.8	**8.6**	-11.4	**6.6**	37.9	**6.8**	87.2	**9.6**	3.0		

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

Results: Lifestyles, education, and health dynamics

4. Lifestyles explain around 40% of the education gradient of LE

LE at age 50 across education and lifestyles: males born in 1970s

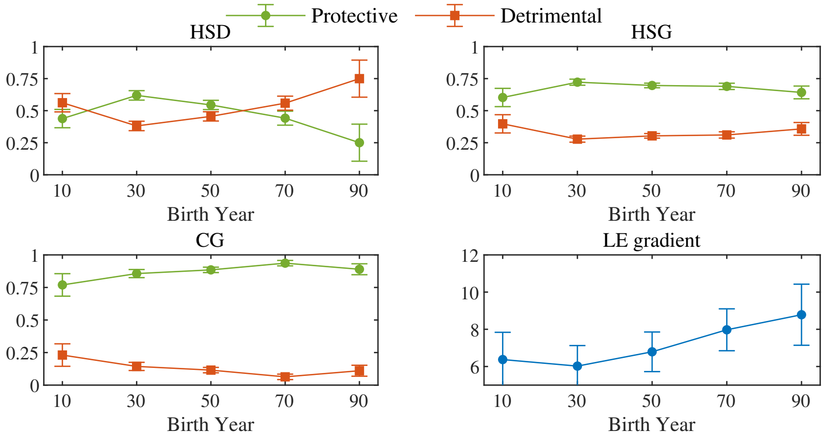
	All		HSD		HSG		CG		ΔLE (CG-HSD)		
	%	LE	%	LE	%	LE	%	LE	Data	(a)	(b)
All	100	29.3	100	24.9	100	28.0	100	32.8	**7.9**	**4.7**	**3.2**
PRO	74.4	31.5	**44.3**	28.5	**69.0**	30.1	**93.6**	33.4	4.9		
DET	25.6	22.9	**55.7**	22.0	**31.0**	23.3	**6.4**	23.8	1.9		
Δ	48.8	8.6	**-11.4**	6.6	**37.9**	6.8	**87.2**	9.6	3.0		

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

Results: Changes across cohorts

5. 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
   a) Working age: standard life-cycle incomplete-markets model of consumption with health and labor market risks

   b) Retirement: as before, but without labor market risks

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:

    $$\tau_{\text{HSD}} = 0 \quad | \quad \tau_{\text{HSG}} \sim N(\mu_{\text{HSG}}, \sigma_{\text{HSG}}) \quad | \quad \tau_{\text{CG}} \sim N(\mu_{\text{CG}}, \sigma_{\text{CG}})$$

  - Cost τ_y of lifestyle y:

    $$\tau_{\text{DET}} = 0 \quad | \quad \tau_{\text{PRO}} \sim N(\mu_{\text{PRO}}, \sigma_{\text{PRO}})$$

  - The choices are independent if $V_0^{\text{CG},\text{PRO}} - V_0^{\text{CG},\text{DET}} = V_0^{\text{HSD},\text{PRO}} - V_0^{\text{HSD},\text{DET}}$
  - There are complementarities between education and lifestyle if:

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

    [See theory]

Stage 2: Life cycle

State variables

- Working agents are heterogeneous with respect to:

1. Types
   - Education e $\in$ {HSD, HSG, CG}
   - Lifestyle y $\in$ {DET, PRO}
   - Cohort c $\in$ {1930, 1970}
2. Exogeneous and deterministic state
   - Age $t \in \{25, 27, 29, ...\}$
3. Exogeneous and stochastic states
   - Health status $h_t \in \{h_g, h_b\}$
   - Employment status $l_t \in \{0, 1\}$
   - Shock to earnings $\zeta_t \in \mathbb{R}$
4. Endogenous state
   - Cash-on-hand $x_t \in [\underline{x}, \infty)$

INGREDIENTS

- Cohort-specific ingredients
  - Perceptions about effect of type y on health dynamics
    - Individuals do not understand well the link between type y and health dynamics
    - $\lambda^{c} \in [0,1]$ indicates how much individuals born in cohort c know about it
  - Employment status $l_t^{ec}(\varepsilon, h)$
  - Labor earnings $w_t^{ec}(\zeta, \epsilon, h)$
- Cohort-independent ingredients
  - Health dynamics $s_t^{ey}(h)$ and $\Gamma_t^{ey}(h'|h)$

    (but cohort effects appear through the joint distribution of e and y)
  - Medical expenses $m_t^{e}(\xi, h)$
  - Progressive tax system $T()$ and minimum income floor $\underline{x}$

Worker's problem

- Worker's problem can be written as:

$$V_t^{\text{eyc}}(h,\zeta,x) = \max_{c,k'} \left\{ u(c) + \beta \Big[\lambda^{\text{c}} W_t^{\text{eyc}}(h',\zeta',x') + (1-\lambda^{\text{c}}) \widehat{W}_t^{\text{eyc}}(h',\zeta',x') \Big] \right\}$$

$$W_t^{\text{eyc}}(h',\zeta',x') = s_t^{\text{ey}}(h) \sum_{h'} \Gamma_t^{\text{ey}}(h'|h)\, \mathbb{E}\left[V_{t+1}^{\text{eyc}}(h',\zeta',x')\right] + \left(1 - s_t^{\text{ey}}(h)\right) v_{t+1}(k')$$

$$\widehat{W}_t^{\text{eyc}}(h',\zeta',x') = s_t^{\text{e}}(h) \sum_{h'} \Gamma_t^{\text{e}}(h'|h)\, \mathbb{E}\left[V_{t+1}^{\text{eyc}}(h',\zeta',x')\right] + \left(1 - s_t(h)\right) v_{t+1}(k')$$

$$\begin{aligned} \text{s.t.}\quad k' &= x - c \\ \tilde{x}' &= (1+r)\,k' + T\left[l_{t+1}^{\text{ec}}(\varepsilon',h') w_{t+1}^{\text{ec}}(\zeta',\epsilon',h')\right] - m_{t+1}^{\text{e}}(\xi',h') \\ x' &= \max\left\{\tilde{x}', \underline{x}\right\} \end{aligned}$$

$$\text{Flow utility:} \quad u(c) = \frac{c^{1-\sigma}-1}{1-\sigma} + b$$

$$\text{Bequest motive:} \quad v_{t+1}(k) = \beta^{T-(t+1)}\theta_1 \frac{(k+\theta_2)^{1-\sigma}-1}{1-\sigma}$$

Calibration

Estimation

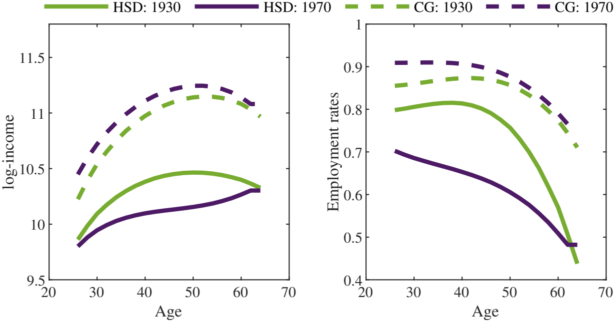
- Life-cycle:
  - External:
    - Parameters related to: demographics, taxes, social security
    - Cohort-specific wages $w_t^{\text{ec}}(\zeta, \epsilon, h)$ and labor force participation $l_t^{\text{ec}}(\varepsilon, h)$
    - Cohort-independent health dynamics $s_t^{\text{ey}}(h)$, $\Gamma_t^{\text{ey}}(h'|h)$, and medical spending $m_t^{\text{e}}(\xi, h)$
  - Internal: SMM to calibrate remaining 4 parameters $(\underline{x}, \theta_1, \theta_2, b)$
    - Median wealth across age (by education and lifestyles) for the 1930s cohort
    - Value of statistical life
- Early life:
  - Estimate cost parameters by matching the joint distribution of education and lifestyles for two different cohorts: 1930 and 1970.
- Outside loop:
  - Perceptions λ^c best matches the joint distribution of education and lifestyles: 1930, 1970

Estimation: Life-cycle

Cohort-specific lifetime earnings

- Lifetime earnings
  - CG: $\Delta_{sc} = +12\%$
  - HSD: $\Delta_{sc} = -18\%$

College premium and labor force participation across cohorts



Source: US Census (1940-2022)

Estimation: Life-cycle

Internally estimated parameters

- The model is able to replicate
  - Higher wealth accumulation for the more educated
  - Higher wealth accumulation for the protective (conditional on education)

- We choose a VSL of 1 Million $

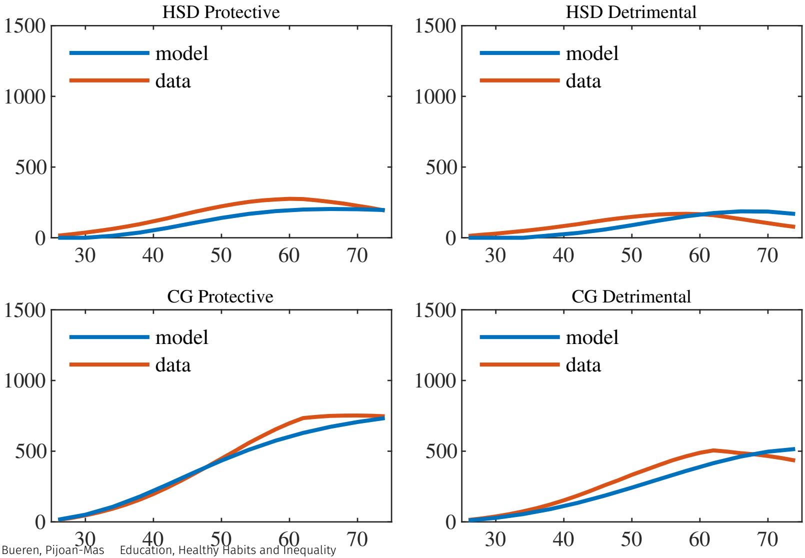
- Parameter values

Parameter	Description	Value
$\underline{x}$	income floor	17.34
θ_1	bequest motive: marginal utility	9.23
θ_2	bequest motive: non-homoteticity	455.35
b	value of life	1.25

Estimation: Life-cycle

Model Fit

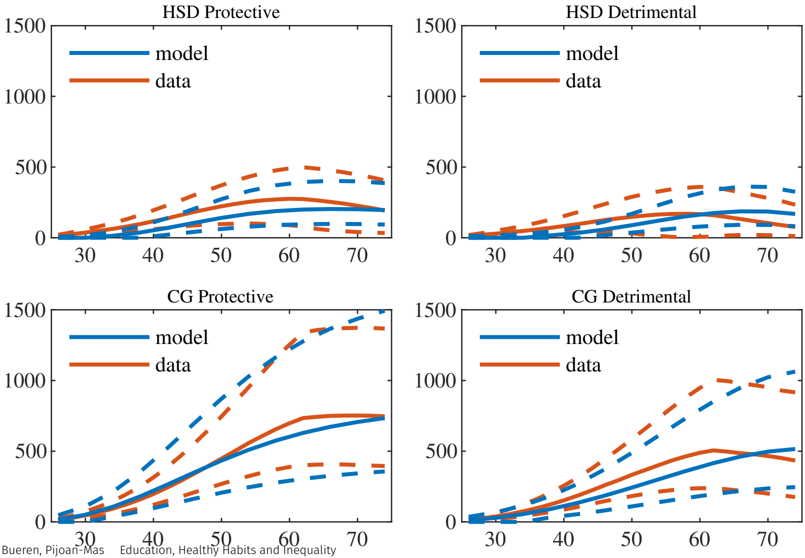
Wealth profiles: median



Estimation: Life-cycle

Model Fit

Wealth profiles: median, 25th, 75th



Estimation: Early-life

- Match the joint distribution of e and y in each cohort c $\in \{1930, 1970\}$
  - → A total of 6 parameters and 10 statistics

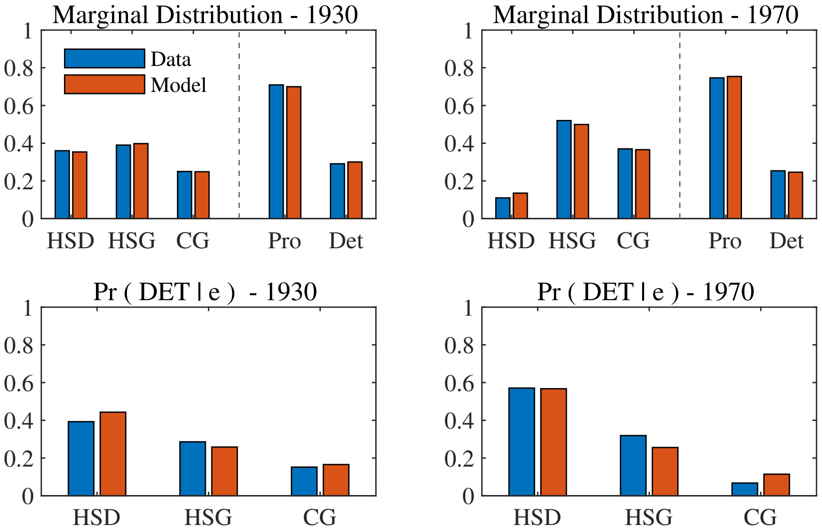
- Identification
  - – μ_{hsg}, μ_{CG}, μ_{PRO} drive the average share of e and y for one cohort
  - – σ_{hsd}, σ_{cg}, σ_{PRO} drive changes in e and y over cohorts c as V_0^{eyc} changes (due to changes in wages and labor force participation)

- Calibrate $\lambda^{1970} = 1$ and λ^{1930} to match the increase in the life-expectancy gradient between the 1930 and the 1970 cohort.

Parameter	Value	Parameter	Value
μ_{HSD}	6.53	σ_{HSD}	2.02
μ_{CG}	26.55	σ_{CG}	18.92
μ_{PRO}	-1.27	σ_{PRO}	10.38
λ^{1930}	0.85		

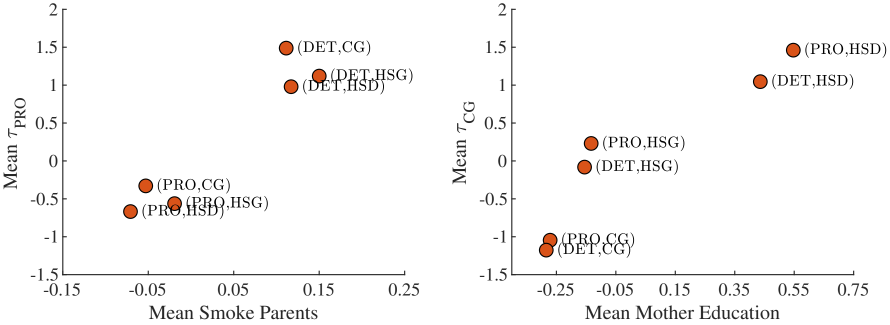
Estimation: Early-life

Model Fit: Early-life



Estimation: Early-life

Model Validation: Early-life



Counterfactuals

Two questions

1. Why is there an education gradient of lifestyles?

2. What has been the effect of the rise in the education wage premium on the increase in education gradient of LE? *(not today)*

Question 1: Education gradient of lifestyle

Education gradient of lifestyle

Mechanisms

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

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

   Higher expected income for the more educated motivates healthier behavior because the value of life increases w/ consumption possibilities

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

2. Complementarities in health dynamics: $s_t^{\text{ey}}(h)$ and $\Gamma_t^{\text{ey}}(h'|h)$

   Gains in life expectancy due to protective health behavior are larger for those with a college education

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

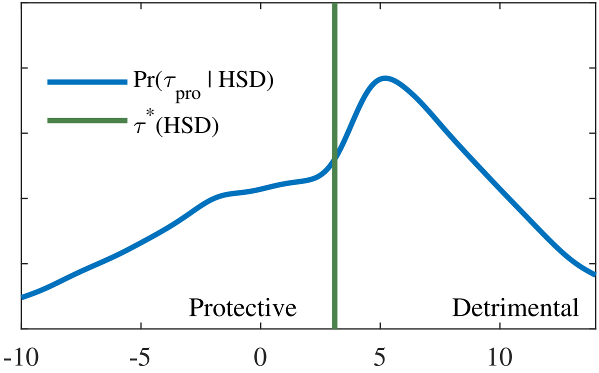
3. Selection:

   Given the complementarities between lifestyle and education, individuals facing lower cost of protective behavior (τ_{PRO}) are more likely to choose higher education.

Education gradient of lifestyle

Lifestyle Choice for HSD

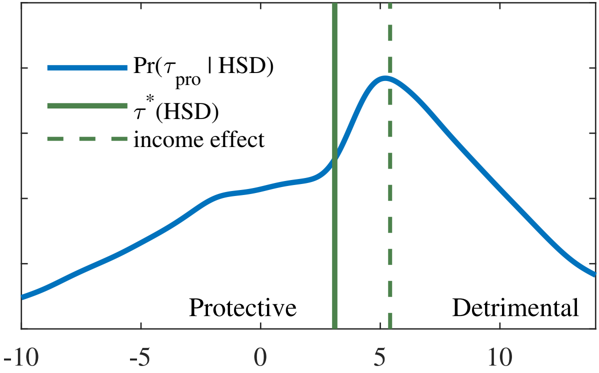
- HSD choose y = PRO iff $\tau_{\text{PRO}} < V_0^{\text{HSD,PRO}} - V_0^{\text{HSD,DET}} \equiv \tau^*(\text{HSD})$



Education gradient of lifestyle

1. Income Effect

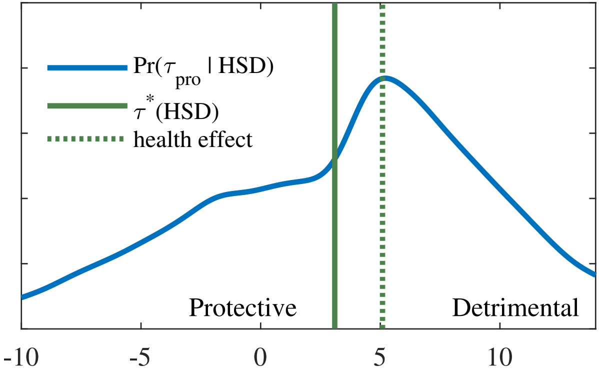
→ If HSD had same income as CG: 16pp more of PRO (out of 43pp gap)



Education gradient of lifestyle

2. Health Effect

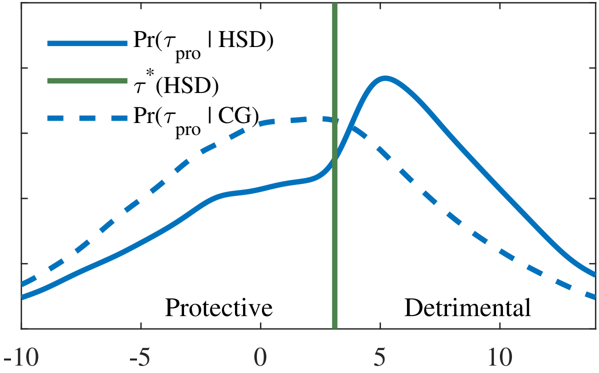
→ If HSD had same health gain of PRO as CG: 13pp more of PRO (out of 43pp gap)



Education gradient of lifestyle

3. Selection

→ If HSD had same distribution of τ_{PRO} as CG: 19pp more of PRO (out of 43pp gap)



Education gradient of lifestyle

Summary

Decomposition: 1970s cohort

	Pr(PRO|HSD)	Pr(PRO|CG)	ΔPr(PRO)	LE(HSD)	LE(CG)	ΔLE
Model	0.42	0.85	0.43	24.7	32.1	7.3
Same lifestyle	0.85	0.85	0.00	27.5	32.1	4.7
Health	0.55	0.85	0.30	25.6	32.1	6.5
Income	0.58	0.85	0.27	25.8	32.1	6.3
Selection	0.61	0.85	0.24	26.0	32.1	6.1

Conclusions

Conclusions

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

  ⇒ Differences in lifestyles across education groups account for 40% of the LE gradient (3.1 out of 7.8 years)

- HA model w/ complementarities in edu and health investments

  1. Education gradient of lifestyles
     - Income advantage explains 25% (⇒ 8 months of LE gradient)
     - Health advantage of protective behavior explains 23% (⇒ 7 months)
     - Selection explains 66% (⇒ 1 year 10 months)
  2. 1.9 years increase in education gradient of LE between 1930s and 1970s cohorts
     - Changes in wages and perceptions explains 2/3 of the increase
     - 44% explained by changes in wages
     - 5% explained by changes in perceptions
     - 60% explained by changes in selection

Econometric Model

1. Health Behavior

- We model the probability of individual i of reporting the m'th behaviour ($z_{mt} = 1$) at time t as a probit model.
  - There is a latent variable (z^*_{mt}) that depends on type (y), age (a_t), health (h_t), and an idiosyncratic shock (ϵ_t)

$$z^*_{mt} = \gamma_{0,m,y} + \gamma_{1,m,y} a_t + \gamma_{2,m,y} a_t^2 + \gamma_{3,m,y} h_t + \epsilon_t, \quad \epsilon_t \sim N(0,1)$$

  - Then,

$$\text{Prob}\left(z_{mt} = 1\right) = \text{Prob}\left(z^*_{mt} > 0\right) = \underbrace{\alpha_m(y, a_t, h_t)}_{\alpha_{mt}}$$

  - Considering independence of health behaviour given type, the probability of observing a sequence of health behaviours $\boldsymbol{z}$ for an individual across time, is assumed to be given by:

$$p(\boldsymbol{z}|\boldsymbol{h}, y) = \prod_{t=1}^{T} \prod_{m=1}^{M} \alpha_{mt}^{z_{mt}} \left(1 - \alpha_{mt}\right)^{1 - z_{mt}}$$

Back to Mixture Model

Econometric Model

2. Health Dynamics

- We model the probability of reporting some health $h' \in \{Good, Bad, Dead\}$ next period as a multinomial probit model
  - There are latent variables ($h^*_{h,h'}$) that depend on gender (g), education (e), type (y), health (h), age (a), and an idiosyncratic shock ($\epsilon_{h'}$)

$$h^*_{h,h'} = f(a, s, e, y; \boldsymbol{\beta}_{h,h'}) + \epsilon_{h'}$$

with,

$$f(a, g, e, y; \boldsymbol{\beta}_{h'}) = \beta_{0,y,e,g,h,h'} + \beta_{1,y,e,g,h,h'} a$$

Back to Mixture Model

Econometric Model

3. Weights

- The mixture weights at the initial age (age 25 are modeled as a multinomial probit model:

$$
\begin{aligned}
y_1^* =& \lambda^1_{0,s,e,c} + \lambda^1_{1,s,e} bh + \epsilon_1 \\
&\vdots \\
y_Y^* =& \lambda^Y_{0,s,e,c} + \lambda^Y_{1,s,e} bh + \epsilon_Y \, ,
\end{aligned}
$$

- We compute weights for future ages using the health transition model:

$$
p(y, h_t | s, e, c) = \sum_{h_{t-1}} p(h_t | h_{t-1}, y, s, e) p(y, h_{t-1} | s, e, c)
$$

Back to Mixture Model

Health Behaviour Data

Mean health behavior and 4-year auto-correlation

	Mean					AC	
	HSD	HSG	CG	50-60	70-80	50-60	70-80
Drinking	0.09	0.10	0.07	0.13	0.05	0.53	0.48
Smoking	0.23	0.16	0.07	0.21	0.08	0.81	0.78
Cancer test	0.66	0.76	0.85	0.71	0.77	0.42	0.41
Cholesterol	0.77	0.84	0.89	0.79	0.89	0.37	0.30
Flu shot	0.58	0.62	0.69	0.49	0.77	0.55	0.62
Exercise	0.26	0.39	0.55	0.42	0.38	0.40	0.39

Notes: HRS. HSD: high-school dropout; HSG: high-school graduate; CG: college graduate; 50-60: sub-sample of individuals aged 50 to 60; 70-80: sub-sample of individuals aged 70 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	Cholesterol	Flu shot	Exercise
Drinking	1.00	0.08	-0.02	-0.03	-0.03	0.02
Smoking	0.18	1.00	-0.10	-0.07	-0.06	-0.08
Cancer test	-0.04	-0.13	1.00	0.26	0.19	0.11
Cholesterol	-0.04	-0.11	0.39	1.00	0.24	0.07
Flu shot	-0.05	-0.05	0.23	0.24	1.00	0.02
Exercise	-0.01	-0.14	0.08	0.04	0.02	1.00

Notes: HRS. Upper diagonal: individuals aged 70 to 80. Lower diagonal: individuals aged 50 to 60.

Back to Data

Set up: two education choices

- An individual decides to hold a protective lifestyle if:

$$\tau_{\text{PRO}} < \max\{V^{\text{CG,PRO}} - \tau_{\text{CG}}, V^{\text{HSD,PRO}}\} - \max\{V^{\text{CG,DET}} - \tau_{\text{CG}}, V^{\text{HSD,DET}}\}$$

Go back

Set up: two education choices

- An individual decides to hold a protective lifestyle if:

$$\tau_{\text{PRO}} < \max\{V^{\text{CG,PRO}} - \tau_{\text{CG}}, V^{\text{HSD,PRO}}\} - \max\{V^{\text{CG,DET}} - \tau_{\text{CG}}, V^{\text{HSD,DET}}\}$$

- $V^{\text{CG,PRO}} - V^{\text{CG,DET}} = V^{\text{HSD,PRO}} - V^{\text{HSD,DET}}$:
  - An individual decides to be protective if:

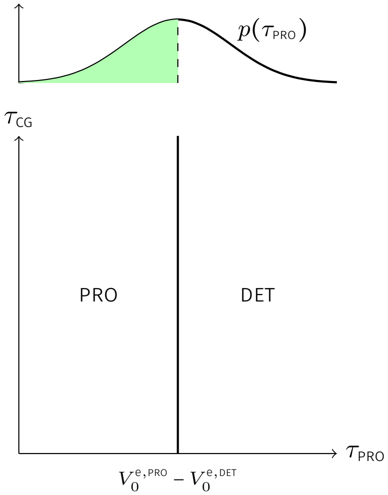
$$\tau_{\text{PRO}} < V^{e,\text{PRO}} - V^{e,\text{DET}}$$

  - The decision of being protective is independent on the value of τ_{CG}

Go back

Choices: two education levels

Independent choices: $V_0^{\text{CG},\text{PRO}} - V_0^{\text{CG},\text{DET}} = V_0^{\text{HSD},\text{PRO}} - V_0^{\text{HSD},\text{DET}}$



Go back

Set up: two education choices

- An individual decides to hold a protective lifestyle if:

$$\tau_{\text{PRO}} < \max\{V^{\text{CG,PRO}} - \tau_{\text{CG}}, V^{\text{HSD,PRO}}\} - \max\{V^{\text{CG,DET}} - \tau_{\text{CG}}, V^{\text{HSD,DET}}\}$$

- With $V^{\text{CG,PRO}} - V^{\text{HSD,PRO}} > V^{\text{CG,DET}} - V^{\text{HSD,DET}}$:
  - An individual decides to be protective if:

    A. $\tau_{\text{CG}} < V^{\text{CG,DET}} - V^{\text{HSD,DET}} : \tau_{\text{PRO}} < V^{\text{CG,PRO}} - V^{\text{CG,DET}}$

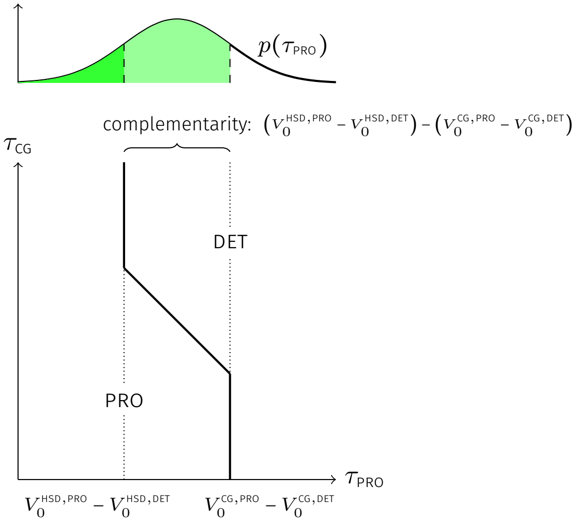
    B. $\tau_{\text{CG}} > V^{\text{CG,PRO}} - V^{\text{HSD,PRO}} : \tau_{\text{PRO}} < V^{\text{HSD,PRO}} - V^{\text{HSD,DET}}$

    C. $V^{\text{CG,DET}} - V^{\text{HSD,DET}} < \tau_{\text{CG}} < V^{\text{CG,PRO}} - V^{\text{HSD,PRO}} : \tau_{\text{PRO}} < V^{\text{CG,PRO}} - \tau_{\text{CG}} - V^{\text{HSD,DET}}$

Go back

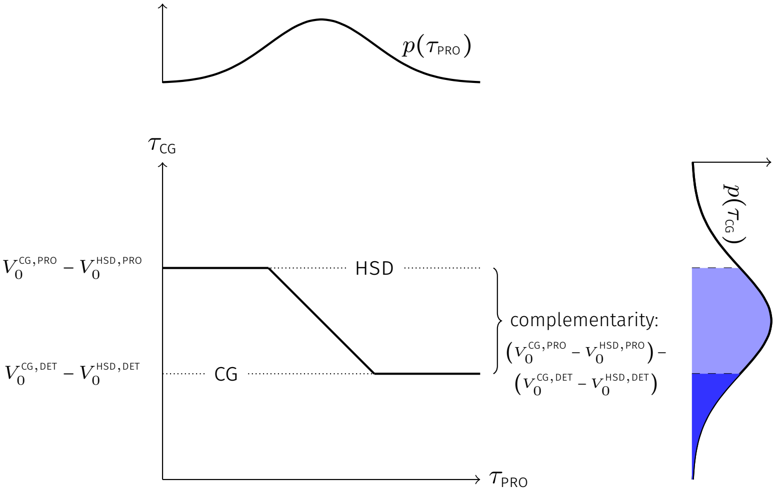
Choices: two education levels

Complementarities: $V_0^{\mathrm{CG,PRO}} - V_0^{\mathrm{CG,DET}} > V_0^{\mathrm{HSD,PRO}} - V_0^{\mathrm{HSD,DET}}$



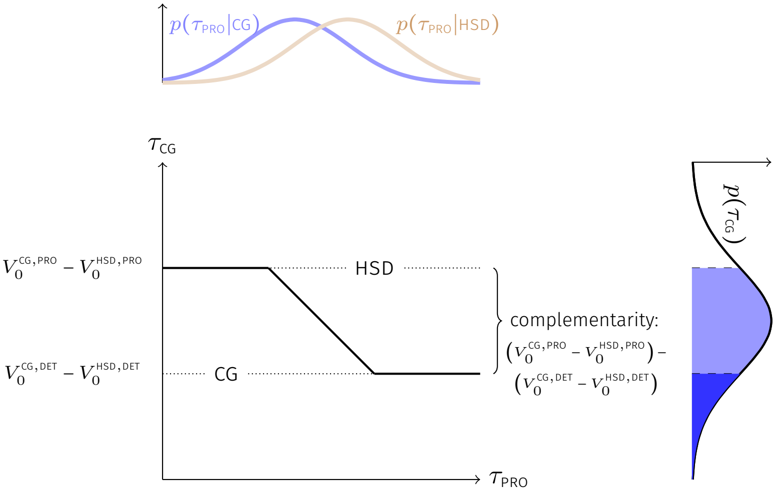
Choices: two education levels

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



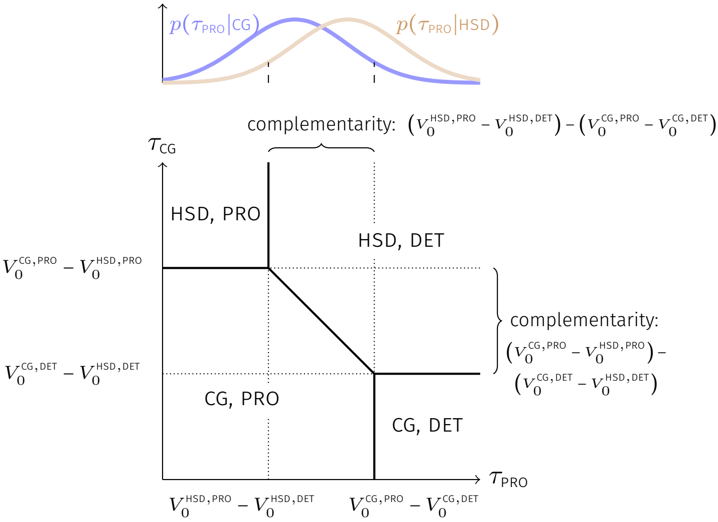
Choices: two education levels

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



Choices: two education levels

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



Choices: two education levels

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

- Complementarities imply:
  - Higher educated individuals are more likely to invest in protective lifestyle as the returns are larger: *direct effect*.
  - Individuals with lower cost cost of adopting health behavior (healthy parents/ peers) are more likely to go to college: *selection effect*.
- Selection drives that the distribution of τ_{PRO} across education groups is different:

$$F(\tau_{\text{PRO}}|\text{CG}) \geq F(\tau_{\text{PRO}}|\text{HSD})$$

$\Rightarrow$ High-school dropouts are negatively selected in terms of τ_{PRO}

Go back