

Healthy Habits and Inequality

Dante Amengual^{*}

Jesús Bueren[◇]

Josep Pijoan-Mas[□]

^{*}CEMFI

[◇]EUI

[□]CEMFI, CEPR

Barcelona Summer Forum, June 2024

INTRODUCTION

- In recent decades, there has been an **increase in inequalities** (*wealth, consumption, income, health outcomes*)
- 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 lifestyles** across education groups can account for these facts
 - Lifestyles are an important driver of health outcomes
McGinnis, Foegen (1993); Li et al. (2018); Zaninotto et al. (2020)
 - More educated individuals tend to adopt healthier lifestyles
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?
- 3 Understand the increase in the education gradient of life expectancy
 - Quantify the role played by the increase in the education wage premium
 - a) Income effect on lifestyle choices
 - b) Selection: different composition of education groups

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); Margaritis, 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 channels

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

Health Dynamics and Health Behavior

THE DATA

- HRS and PSID
 - Unbalanced panels of individuals $i = 1, \dots, N$ followed $t = 1, \dots, T$ 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

HEALTH BEHAVIOR AND HEALTH OUTCOMES

- Across demographic groups: **health behaviors** associated w/ health outcomes
 - But, at individual level:
 - Different **health behaviors** imperfectly correlated across individuals and over time
[\[See data\]](#)
 - Hard to identify their long-run effect on health outcomes
 - Curse of dimensionality if you want to put them in a model
 - Objective: identify latent types in lifestyle behavior based on:
 - Cross-sectional and panel variation of health behavior
 - Health dynamics
- Types are permanent, but health behaviors change with age and health status

LATENT TYPES

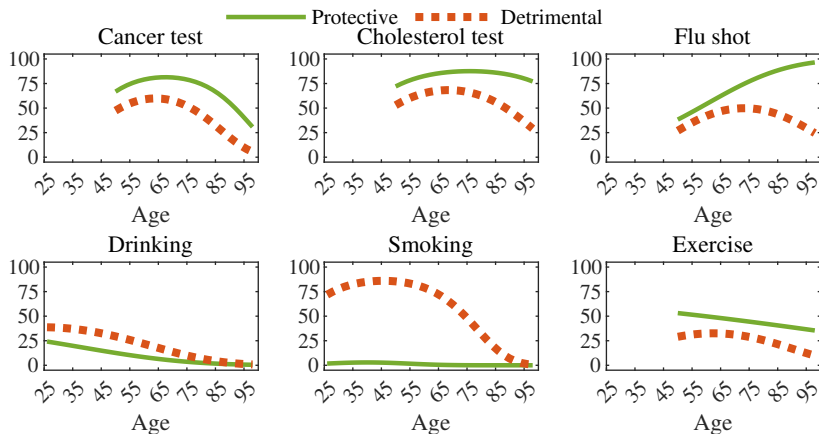
- We assume that observed **health behavior** is the (noisy) result of a **latent time-invariant** factor $y \in Y \equiv \{y_1, y_2, \dots\}$
 - We interpret y as the individual lifestyle: propensity to engage in healthy behaviors
- We propose a novel econometric model to
 - Allocate individuals to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

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

RESULTS: LIFESTYLES AND HEALTH BEHAVIOR

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

Figure 1: **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**

Table 1: **LE at age 50 across education and lifestyles: males born in 1970s**

	All		HSD		HSG		CG		Δ LE (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, $\Delta_e p(\mathbf{h}|e, y)$

(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)

Table 1: LE at age 50 across education and lifestyles: males born in 1970s

	All		HSD		HSG		CG		Δ LE (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, $\Delta_e p(\mathbf{h}|e, y)$

(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

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

Table 1: LE at age 50 across education and lifestyles: males born in 1970s

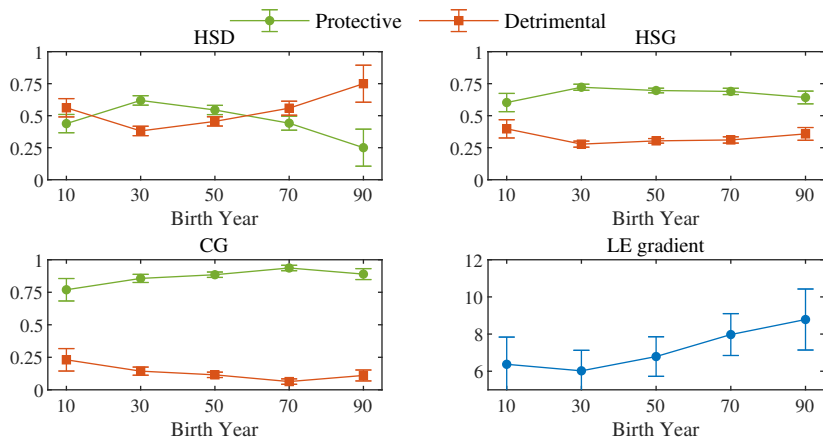
	All		HSD		HSG		CG		Δ LE (CG-HSD)		
	%	LE	%	LE	%	LE	%	LE	Data	(a)	(b)
All	100	29.3	100	24.9	100	28.0	100	32.8	7.9	4.7	3.2
PRO	74.4	31.5	44.3	28.5	69.0	30.1	93.6	33.4	4.9		
DET	25.6	22.9	55.7	22.0	31.0	23.3	6.4	23.8	1.9		
Δ	48.8	8.6	-11.4	6.6	37.9	6.8	87.2	9.6	3.0		

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

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

RESULTS: CHANGES ACROSS COHORTS

5. Education gradient of lifestyles widens over time

Figure 2: **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 $e \in \{HSD, HSG, CG\}$
 - lifestyle $y \in \{DET, PRO\}$

- They solve $\max_{e,y} \{V_0^{eyc} - \tau_e - \tau_y\}$

- Value V_0^{eyc} of starting stage 2 with type (e, y, c)
- Cost τ_e of education e :

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

- Cost τ_y of lifestyle y :

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

- The choices are independent if $V_0^{CG,PRO} - V_0^{CG,DET} = V_0^{HSD,PRO} - V_0^{HSD,DET}$
- There are [complementarities](#) between education and lifestyle if:

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

Stage 2: Life cycle

STATE VARIABLES

- Working agents are heterogeneous with respect to:

1 Types

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

2 Exogenous and deterministic state

- Age $t \in \{25, 27, 29, \dots\}$

3 Exogenous 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 $w_t^{ec}(\zeta, \varepsilon, h)$

- 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 $T(\cdot)$ and minimum income floor \underline{x}

WORKER'S PROBLEM

- Worker's problem can be written as:

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

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

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

$$\text{s.t. } k' = x - c$$

$$\tilde{x}' = (1 + r)k' + T[l_{t+1}^{\text{ec}}(\varepsilon', h')w_{t+1}^{\text{ec}}(\zeta', \epsilon', h')] - m_{t+1}^{\text{e}}(\xi', h')$$

$$x' = \max\{\tilde{x}', \underline{x}\}$$

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

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

Calibration

ESTIMATION

- Life-cycle:
 - External:
 - Parameters related to: demographics, taxes, social security
 - Cohort-specific wages $w_t^{ec}(\zeta, \epsilon, h)$ and labor force participation $l_t^{ec}(\epsilon, h)$
 - Cohort-independent health dynamics $s_t^{ey}(h)$ and $\Gamma_t^{ey}(h'|h)$
 - Cohort-independent medical spending $m_t^e(\xi, h)$
 - 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 marginal distributions of education and lifestyles for two different cohorts: 1930 and 1970.
- Outside loop:
 - Perceptions λ^c to match joint distribution of education and lifestyles: 1930, 1970

ESTIMATION: LIFE-CYCLE

Externally estimated parameters

- Labor income (PSID and Census):

- Extensive margin (probit):

$$\hat{l}_t^{\text{ec}}(\varepsilon_t, h_t) = f_t^{\text{ec}} + f_t^e(h_t) + \varepsilon_t, \quad \varepsilon_t \sim N(0, 1)$$

$$l_t^{\text{ec}}(\varepsilon_t, h_t) = 1 \text{ if } \hat{l}_t^{\text{ec}}(\varepsilon_t, h_t) > 0$$

- Intensive margin:

$$\log w_t^{\text{ec}}(\zeta_t, \varepsilon_t, h_t) = \omega_t^{\text{ec}} + \omega_t^e(h_t) + \zeta_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

$$\zeta_t = \rho_\zeta \zeta_{t-1} + \nu_t, \quad \nu_t \sim N(0, \sigma_\nu^2)$$

- Medical expenses (HRS):

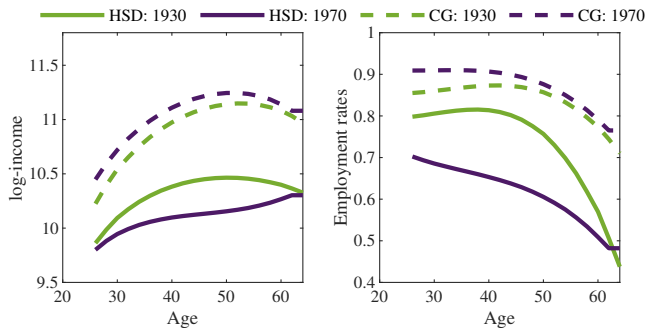
$$\log m_t^e(\xi_t, h_t) = \lambda_t^e(h_t) + \xi_t, \quad \xi_t \sim N(0, \sigma_\xi^2)$$

ESTIMATION: LIFE-CYCLE

Cohort-specific lifetime earnings

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

Figure 3: College premium and labor force participation across cohorts



Source: US Census (1940-2022)

ESTIMATION: LIFE-CYCLE

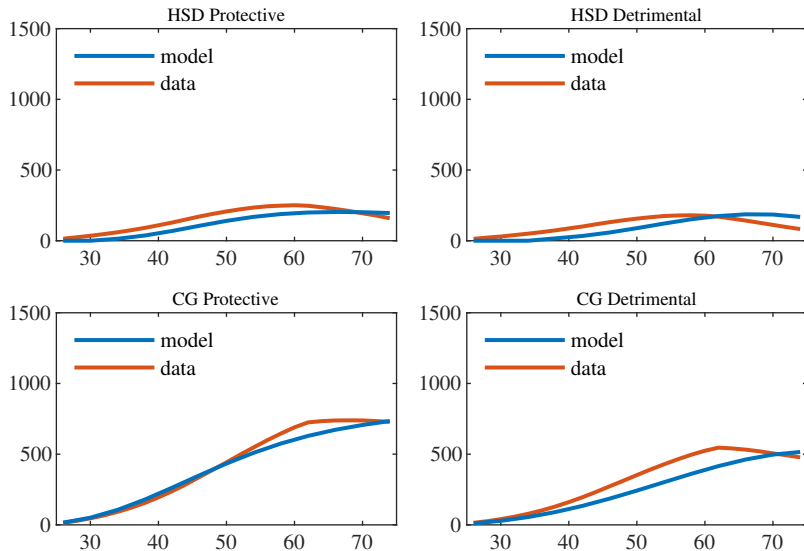
Internally estimated parameters

- 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

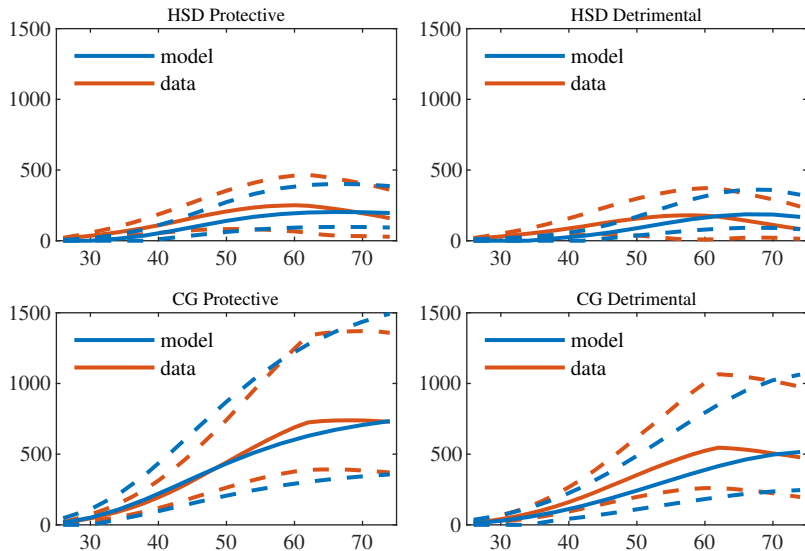
ESTIMATION: LIFE-CYCLE

Model Fit

Figure 4: **Wealth profiles: median**

ESTIMATION: LIFE-CYCLE

Model Fit

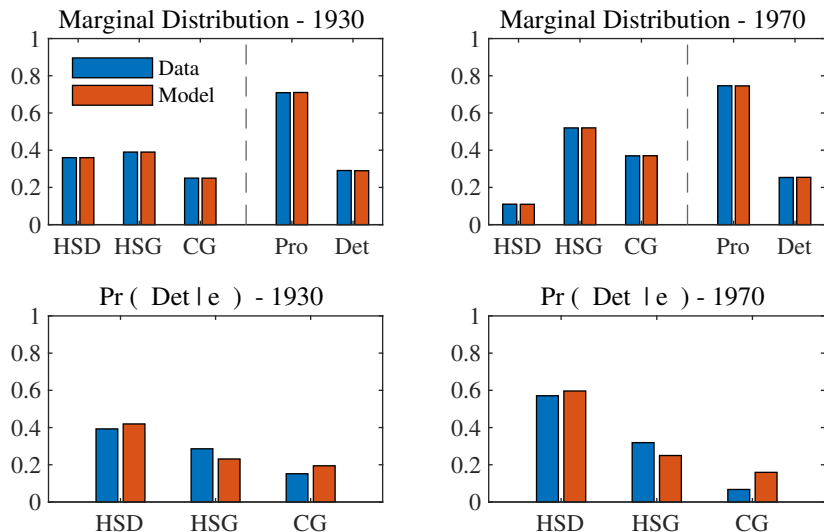
Figure 6: **Wealth profiles: median, 25th, 75th**

ESTIMATION: EARLY-LIFE

- Match the marginal distribution of e and y in each cohort $c \in \{1930, 1970\}$
 - A total of 6 parameters and 6 statistics
- Identification
 - $\mu_{\text{HSG}}, \mu_{\text{CG}}, \mu_{\text{PRO}}$ drive the average share of e and y for one cohort
 - $\sigma_{\text{HSD}}, \sigma_{\text{CG}}, \sigma_{\text{PRO}}$ 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 $p(y|e)$ for 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

Figure 8: **Model Fit: 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?

Question 1: Education gradient of lifestyle

EDUCATION GRADIENT OF LIFESTYLE

Mechanisms

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

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

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

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

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

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

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

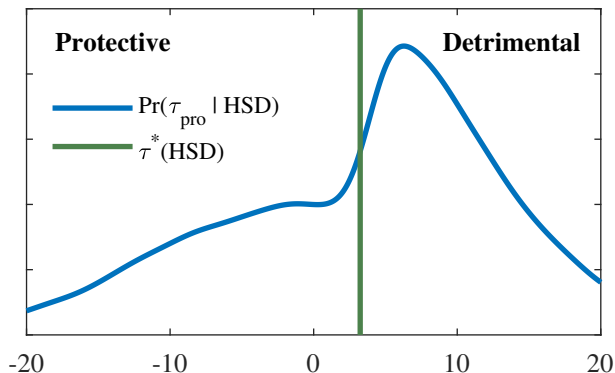
3 Selection:

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

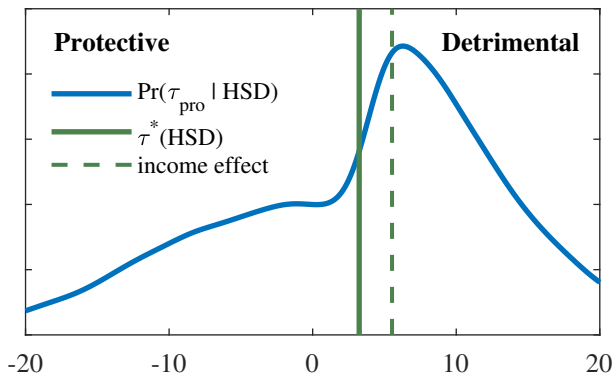
- HSD choose $y = \text{PRO}$ iff $\tau_{\text{PRO}} < V_0^{\text{HSD,PRO}} - V_0^{\text{HSD,DET}} \equiv \tau^*(\text{HSD})$



EDUCATION GRADIENT OF LIFESTYLE

1. Income Effect

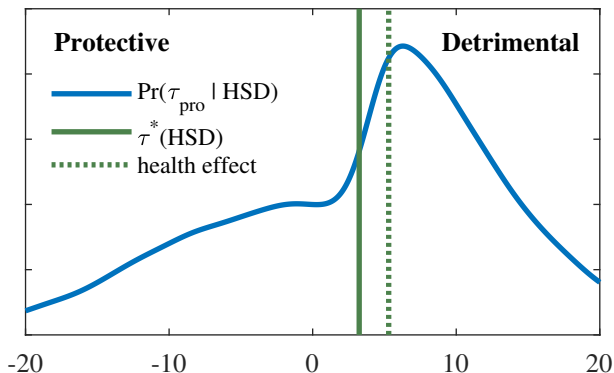
→ If HSD had **same income** as CG: 11pp more of PRO (out of 44pp gap)



EDUCATION GRADIENT OF LIFESTYLE

2. Health Effect

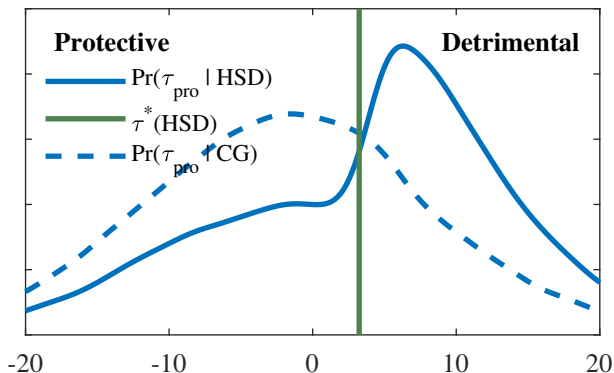
→ If HSD had **same health gain of PRO** as CG: 10pp more of PRO (out of 44pp gap)



EDUCATION GRADIENT OF LIFESTYLE

3. Selection

→ If HSD had same distribution of τ_{PRO} as CG: 29pp more of PRO (out of 44pp gap)



EDUCATION GRADIENT OF LIFESTYLE

Summary

Table 2: **Decomposition: 1970s cohort**

	$\Pr(\text{PRO} \text{HSD})$	$\Pr(\text{PRO} \text{CG})$	$\Delta\Pr(\text{PRO})$	LE(HSD)	LE(CG)	ΔLE
Model	0.40	0.84	0.44	24.6	31.9	7.3
Same lifestyle	0.84	0.84	0.00	27.5	32.2	4.7
Health	0.50	0.84	0.34	25.2	31.9	6.7
Income	0.51	0.84	0.33	25.3	31.9	6.6
Selection	0.69	0.84	0.15	26.5	31.9	5.4

Question 2: Changes over time

CHANGES OVER TIME

Mechanisms

- Income effect: Increases in the education wage premium strengthen the complementarity between education and lifestyles,

$$\Delta_c (V_0^{\text{CG,PRO,C}} - V_0^{\text{CG,DET,C}}) > \Delta_c (V_0^{\text{HSD,PRO,C}} - V_0^{\text{HSD,DET,C}})$$

→ This increases the education gradient in $\Pr[y = \text{PRO} | e]$

- Perceptions: As individuals become aware of the importance of types on health outcomes, they are more willing to invest in **protective** behavior.

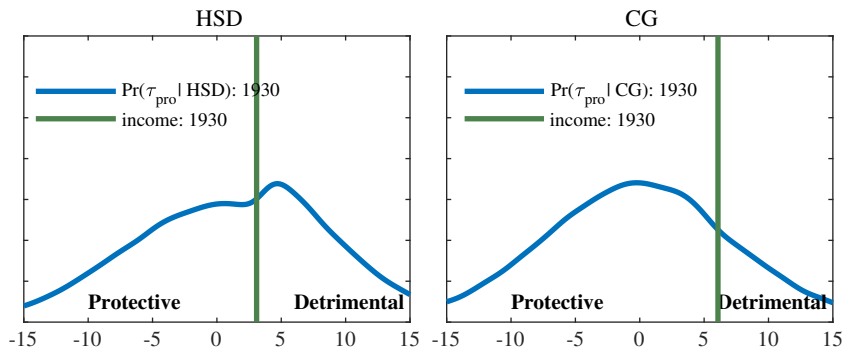
→ Unclear effect on the education gradient

- Selection effect: worse pool of individuals in terms τ_{PRO} for both HSD and CG.

→ Unclear effect on the education gradient

CHANGES OVER TIME: 1930 VS 1970 COHORTS

1930 cohort

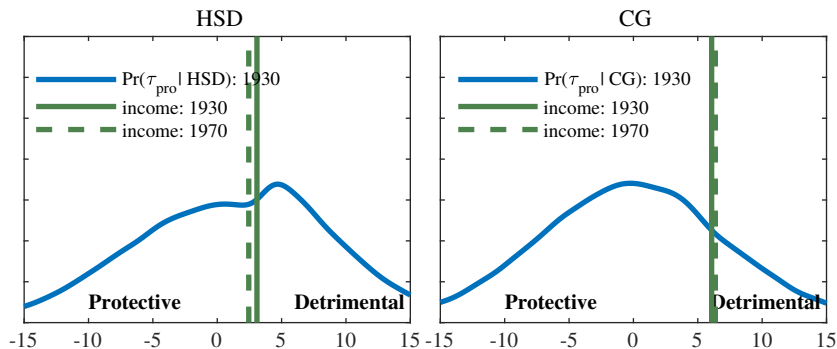


CHANGES OVER TIME: 1930 VS 1970 COHORTS

Income effect

$\Pr(\tau_{\text{PRO}} | e)$ as in 1930, education premium widens

- $\Pr(\text{PRO}|\text{HSD})$ declines 2pp
- $\Pr(\text{PRO}|\text{CG})$ increases 1pp
- LE gradient widens 3 months

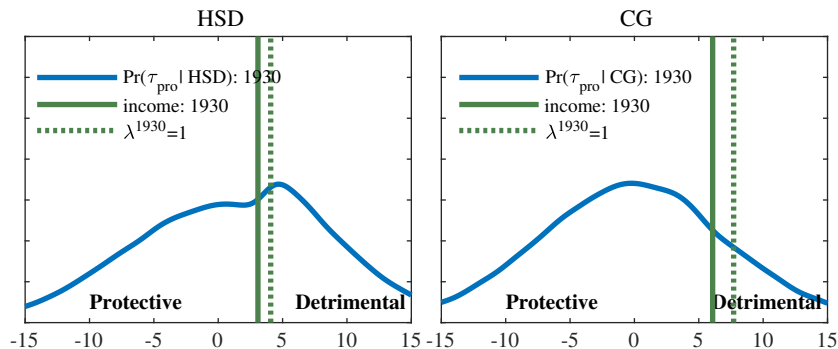


CHANGES OVER TIME: 1930 VS 1970 COHORTS

Changes in Perceptions

$\Pr(\tau_{\text{PRO}} | e)$ and income as in 1930, perceptions as in 1970 (fully aware)

- $\Pr(\text{PRO}|\text{HSD})$ increases 2pp
- $\Pr(\text{PRO}|\text{CG})$ increases 2pp
- LE gradient widens 1 month

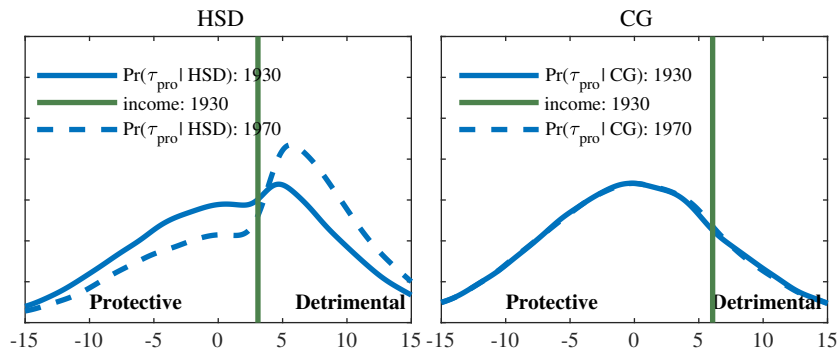


CHANGES OVER TIME: 1930 VS 1970 COHORTS

Selection effect

$\Pr(\tau_{\text{PRO}} | e)$ changes, education premium and perceptions as in 1930

- $\Pr(\text{PRO}|\text{HSD})$ declines 17pp
- $\Pr(\text{PRO}|\text{CG})$ increases 0pp
- LE gradient widens 1 year and 2 months



CHANGES OVER TIME: 1930 VS 1970 COHORTS

Summary

Table 3: **Decomposition**

	$\Delta_c \text{Pr}(\text{CG})$	$\Delta_c \text{Pr}(\text{HSD})$	$\Delta_c \text{Pr}(\text{PRO})$	$\Delta_c \text{Pr}(\text{PRO} \text{CG})$	$\Delta_c \text{Pr}(\text{PRO} \text{HSD})$	$\Delta_c \Delta_e \text{LE}$
Data	0.12	-0.25	0.04	0.08	-0.18	1.92
Model	0.12	-0.25	0.04	0.03	-0.18	1.50
Income				0.01	-0.02	0.22
Perception				0.02	0.02	0.07
Selection				0.00	-0.17	1.16

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 80% of the increase
 - 15% explained by changes in wages
 - 5% explained by changes in perceptions
 - 77% 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,

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

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

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

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 ($\epsilon_{h'}$)

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

with,

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

[Back to Mixture Model](#)

ECONOMETRIC MODEL

3. Weights

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

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

- 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

Table 4: 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

Table 5: **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](#)