Healthy Habits and Inequality

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

[☆]CEMFI

[◇]EUI

[□]CEMFI, CEPR

University of Bristol, November 2023

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 and Hauser (1973); Pijoan-Mas and Rios-Rull (2014); Chetty *et al* (2016)
  b) Growing educational gradients of health inequality
  Preston and Elo (1995); Meara et al. (2009); Montez *et al* (2011); Case and Deaton (2015)

  → Reasons not well-understood
- We aim to study to which extent differences in lifestyles across education groups and time can account for these facts
  - Differences in lifestyles are an important driver of health outcomes
  Li *et al* (2018); Zaninotto, Head, Steptoe (2020)
  - More educated individuals tend to adopt healthier habits
  Lantz et al. (1998); Martikainen et al. (2013)

Objectives

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

2. Understand the joint determination of education and lifesyle

   → Why is there an education gradient of lifestyles?

3. Quantify the effect of increased economic inequality on the increase in the educational gradient of health outcomes

What we do: Data

- There are many indicators of health behavior in HRS and PSID (*preventive tests, substance abuse, exercise*)
- Ideally, we would like to incorporate all this info into a structural model
- Problems:
  - Observed health behaviors are noisy signals of lifestyles as imperfectly correlated across individuals and over time
  - Curse of dimensionality
- Contribution → Novel methodology to reduce the dimensionality of the data by identifying (permanent) patterns in lifestyle behavior exploiting:
  - Cross-sectional and panel variation on health behavior
  - Relationship between health behavior and health dynamics

What we do: Data

Results

1. Health behavior is parsimoniously well represented by two different lifestyles: protective and detrimental

2. Large LE gradient at age 50: 8 years between protective and detrimental

3. Lifestyles are correlated w/ education
   - Harmful types much more frequent among the less educated
   - Habit gradient of similar size within education groups
   - Lifestyles explain around 40% of the education gradient in LE

4. Individuals holding detrimental health behaviors accumulate less wealth (controlling by education)

5. There is an increasing dispersion in lifestyles across education groups
   - Lifestyles explain 1/3 of the LE edu gradient for individuals born in the 1930s
   - Lifestyles explain 1/2 of the LE edu gradient for individuals born in the 1970s

What we do: Model

Main ingredients

- We build an heterogeneous agents model featuring two different stages

  1. Early life: individuals choose education and lifestyle.
  2. Working/retirement age: individuals solve a standard life-cycle model with idiosyncratic labor income and health risks

     (conditional on given education and lifestyle)

- Complementarities between education and lifestyle investments

  - An extra year of life is more valuable with higher consumption possibilities
  - The benefit in health transitions of investing in protective health behavior differs across education groups
  - Early life determinants of education and lifestyle may be correlated

What we do: Model

Numbers

- We take health dynamics conditional on education and lifestyle as given
- Calibrate the model to match:
  1. The joint distribution of education and lifestyles choices for different cohorts: born in the 30's, 50's and 70's.
  2. Savings decisions by education and lifestyles of individuals born in 1930's.
- Main results:
  - → Education gradient of lifestyles. Both the income and health advantages of college education are equally important
  - → Increase in LE gradient .
    - The increases in income inequality between the cohorts born in 1930s and 1970s explains around 78%
    - Worse behavior of highschool dropouts (college) mostly explained by selection (increases in wages)

Literature on health and economic inequality

- Models with *exogenous health*
  Hosseini et al. (2021); De Nardi et al. (2023)
  → We model endogenous health

- Models with *endogenous monetary health investments*
  Fonseca et al. (2023); Hong et al.(2023)
  → We focus on health-related behaviour

- Models with *endogenous health behaviour investments*
  Cole et al. (2019); Mahler and Yum (2023); Margaris and Wallenius (2023)
  → We focus on once-and-for-all choices of lifestyles and education
  → Identification of lifestyles on health based on differences in health dynamics and mortality.

→ We study how long-run changes in economic inequality shape changes in health inequality

The Data

- The HRS and PSID provide an unbalanced panel of individuals $i = 1, ..., N$ followed for $t = 1, ..., T$ periods
- Standard demographic information: gender (g), 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}):
    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 want to incorporate heterogeneity in health dynamics across lifestyles into a structural model.
- In principle, we could include all the health behavior variables into the state-space but that would imply: 2^6 grid points for describing it.
- Instead, we are going to assume that observed health behavior (z_{mt}) is the result of some latent time-invariant factor (y)
  - The latent factor is represented by a few discrete groups $y \in \{y_1, y_2, ...\}$.
- We interpret the latent factor (y) as the lifestyle
  - Allocate individuals to lifestyles
  - Measure the importance of lifestyles on health dynamics

Econometric Model

Overview

- We *jointly estimate* health dynamics and lifestyles using a mixture model:

$$
\begin{aligned}
p(\boldsymbol{z},\boldsymbol{h}|c,s,e,a,h_0) &= \sum_{y\in Y} p(\boldsymbol{z},\boldsymbol{h}|c,s,e,a,h_0,y)p(y|c,s,e,a,h_0) \\
&= \sum_{y\in Y} p(\boldsymbol{z}|\boldsymbol{h},a,h_0,y)p(\boldsymbol{h}|s,e,a,h_0,y)p(y|c,s,e,a,h_0)
\end{aligned}
$$

  - By estimating types and transition jointly, we find the types that better represent both the observed behaviour and the health transitions
    (vs. k-means clustering on habits and then transitions)

Econometric Model

1. Healthy Habits

- 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:

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

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,

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

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:

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

Results: Lifestyles

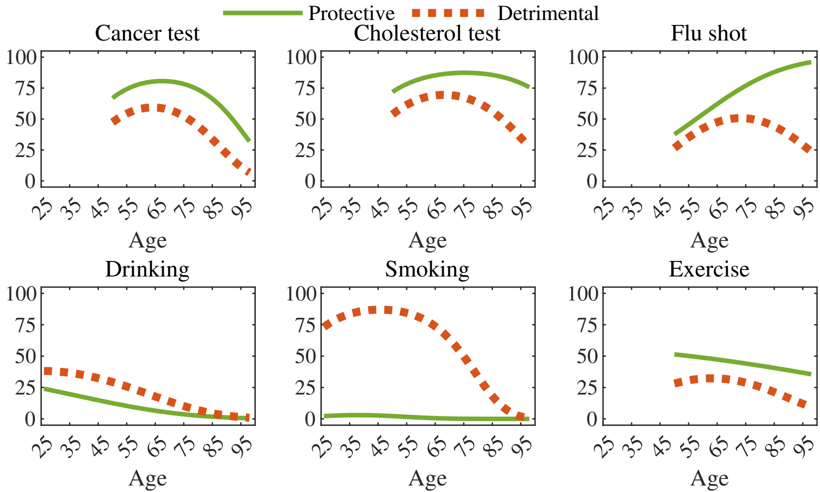


Figure: Probability of reporting health behaviors by lifestyle

Results: Lifestyles, education, and health dynamics

Males

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

	Dropouts		High-school		College		College-Dropout	
	Share	LE	Share	LE	Share	LE	ΔLE	ΔLE
All	100.0	25.5	100.0	28.0	100.0	32.3	6.8	4.3
Protective	54.3	29.0	69.1	30.3	88.3	33.4	4.4	
Detrimental	45.7	21.4	30.9	23.0	11.8	24.5	3.1	
Δ	8.6	7.6	38.2	7.2	76.5	8.9	1.4	

Results: Lifestyles, education, and health dynamics

Females

Table: LE at age 50 across education and lifestyles: females born in 1950s

	Dropouts		High-school		College		College-Dropout	
	Share	LE	Share	LE	Share	LE	ΔLE	ΔLE
All	100.0	28.1	100.0	31.5	100.0	34.3	6.2	4.6
Protective	68.5	30.3	75.9	33.2	90.0	34.9	4.6	
Detrimental	31.5	23.2	24.2	26.1	10.0	28.2	5.0	
Δ	37.0	7.1	51.7	7.2	80.1	6.7	-0.4	

Results: Health Dynamics

- More educated individuals tend to adopt healthier lifestyles.
  - The fraction of males with harmful lifestyle is 4 times bigger among high-school dropouts than among college graduates.
- If dropout males had the same lifestyles as college males, their life expectancy would increase by 2.5 extra years.
  - This corresponds to 37% of the observed difference in life-expectancy.

Results: Changes across cohorts

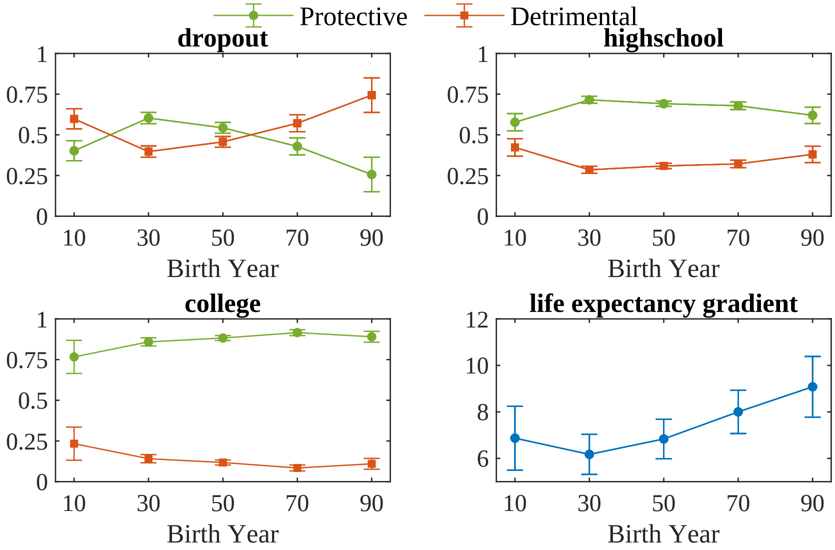


Figure: Probability of lifestyle at age 50 across cohorts. Males.

Results: Changes across cohorts

- Differences in lifestyles across education groups have increased.
  - The share of dropouts holding a detrimental lifestyle has increased from 40% for those born in 1930 to 57% for those born in 1970.
  - The share of college graduates holding a detrimental lifestyle has decreased from 14% for those born in 1930 to 8.4% for those born in 1990.
- This divergence in lifestyles across education groups has led to an increase in the life expectancy gradient.
  - From 6.2 years in 1950 to 8.0 years in 1970.
  - The importance of the differences in lifestyles to explain the life expectancy gradient has increased from 31% to 46% for those born in 1930 and 1970, respectively.

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

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 $\in \{1930, 1950, 1970\}$
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 $\zeta_t \in \mathbb{R}$
   - Shock to medical expenses $\xi_t \in \mathbb{R}$
4. Endogenous state
   - Cash-on-hand $x_t \in [\underline{x}, \infty)$

Important ingredients

- Health dynamics $s_t^{\text{ey}}(h)$ and $\Gamma_t^{\text{ey}}(h'|h)$ as estimated before
  - Survival and health transition are *not* cohort-specific
- Medical expenses $m_t^{\text{e}}(\xi, h)$
- Employment status $l_t^{\text{e}}(l_{-1}, \varepsilon, h)$ depends on
  - Education e and age t
  - Previous period employment l_{-1} and shock ε
  - Health h
- Labor earnings $w_t^{\text{ec}}(\zeta, \epsilon, h)$ depend on
  - Education e, cohort c, and age t
  - Persistent and transitory stochastic component ζ and ϵ
  - Health h
- Education costs $\bar{\tau}_t^{\text{ec}}$ paid over 8 first years of working life
- Progressive tax system $T()$ and minimum income floor $\underline{x}$

Worker's problem

- Worker's problem can be written as:

$$V_t^{\text{eyc}}(h,l,\xi,\zeta,x) = \max_{c,k'} \Bigg\{ u(c) + \beta s_t^{\text{ey}}(h) \sum_{h'} \Gamma_t^{\text{ey}}(h'|h)\, \mathbb{E}\left[V_{t+1}^{\text{eyc}}(h',l',\xi',\zeta',x')\right] + \beta^{T-t}\left(1 - s_t^{\text{ey}}(h)\right) v(k') \Bigg\}$$

s.t.

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

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

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

Estimation

A two-step estimation strategy

1. External: parameters related to demographics, taxes, social security, and the stochastic processes for earnings and for health dynamics
   - This includes cohort-specific wages $w_t^{ec}(\zeta, h, \epsilon)$ and education costs $\bar{\tau}_t^{ec}$
2. Internal: SMM to calibrate remaining parameters
   - Median wealth across age (by education and lifestyles) for the 1930s cohort
   - Value of statistical life

Estimation

Externally estimated parameters

- Labor income (PSID):
  - Extensive margin:

$$l_t^*(\mathsf{e}, l_{t-1}, \varepsilon_t, h_t) = f_t(\mathsf{e}, l_{t-1}, h_t) + \varepsilon_t, \quad \varepsilon_t \sim N(0,1)$$
$$l_t^{\mathsf{e}}(l_{t-1}, \varepsilon_t, h_t) = 1 \text{ if } l_t^*() > 0$$

  - Intensive margin:

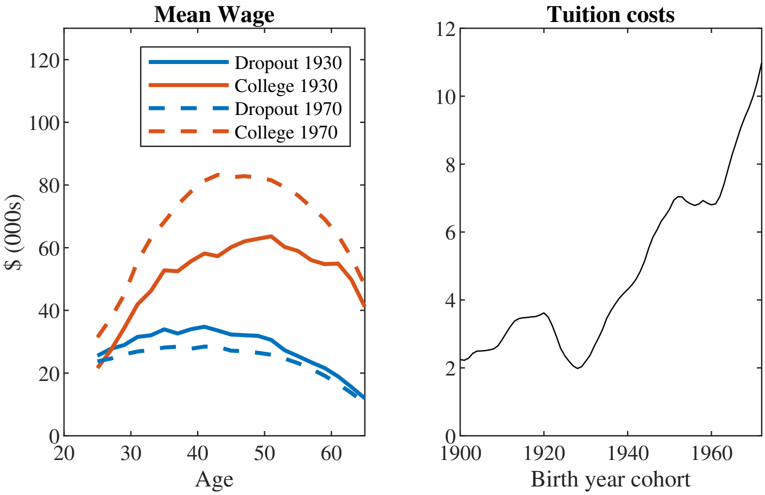
$$w_t^{\mathsf{ec}}(\zeta_t, \epsilon_t, h_t) = \omega_t^{\mathsf{ec}}(h_t) + \zeta_t + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2)$$
$$\zeta_t = \rho_\zeta \zeta_{t-1} + \nu_t, \quad \nu_t \sim N(0, \sigma_\nu^2)$$

- Medical expenses (HRS):

$$m_t^{\mathsf{e}}(\xi_t, h_t) = \lambda_t^e(h_t) + \xi_t$$
$$\xi_t = \rho_\xi \xi_{t-1} + \psi_t, \quad \psi_t \sim N(0, \sigma_\psi^2)$$

Estimation

Externally estimated parameters



Estimation

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)

- Parameter values

Parameter	Description	Value
$\underline{x}$	income floor	16.05
θ_1	bequest motive: marginal utility	9.57
θ_2	bequest motive: non-homoteticity	130.36
b	value of life	0.66

Estimation

Model Fit

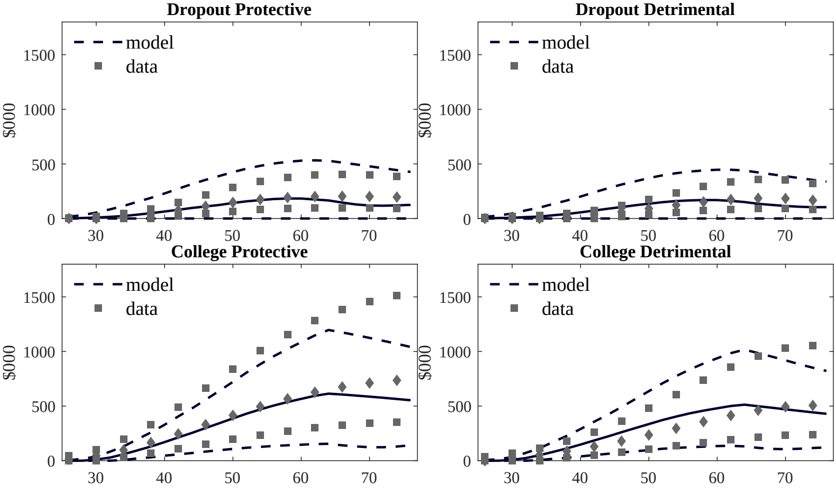


Figure: Wealth distribution: model vs data

Set up

- Teenager/parents make once-and-for-all simultaneous choices of
  - education $\mathsf{e} \in \{\mathrm{HSD}, \mathrm{HSG}, \mathrm{CG}\}$
  - lifesytle $\mathsf{y} \in \{\mathrm{DET}, \mathrm{PRO}\}$

- They solve $\max\limits_{\mathsf{e},\mathsf{y}} \left\{ V_0^{\mathsf{eyc}} - \tau_{\mathsf{ec}} - \tau_{\mathsf{yc}} \right\}$
  - Value V_0^{eyc} of starting stage 2 with type $(\mathsf{e}, \mathsf{y}, \mathsf{c})$
  - Cost τ_{ec} of education e for cohort c:

$$\tau_{\mathrm{HSD},\mathsf{c}} = 0 \quad \Big| \quad \tau_{\mathrm{HSG},\mathsf{c}} = \mu_{\mathrm{HSG}} + \epsilon_{\mathrm{HSG},\mathsf{c}} \quad \Big| \quad \tau_{\mathrm{CG},\mathsf{c}} = \mu_{\mathrm{CG}} + \epsilon_{\mathrm{CG},\mathsf{c}}$$

  - Cost τ_{yc} of lifestyle y for cohort c:

$$\tau_{\mathrm{DET},\mathsf{c}} = 0 \quad \Big| \quad \tau_{\mathrm{PRO},\mathsf{c}} = \mu_{\mathrm{PRO}} + \epsilon_{\mathrm{PRO},\mathsf{c}}$$

  - Where μ_{e}, μ_{y} are average costs of actions e and y
  - Where $\epsilon_{\mathsf{e},\mathsf{c}}$, $\epsilon_{\mathsf{y},\mathsf{c}}$ are (jointly distributed) idiosyncratic costs of actions e and y

Shocks

- Let $(\epsilon_{\mathrm{PRO},\mathsf{c}}, \epsilon_{\mathrm{HSG},\mathsf{c}}, \epsilon_{\mathrm{CG},\mathsf{c}})$ be joint normally distributed:

$$\begin{bmatrix} \epsilon_{\mathrm{PRO},\mathsf{c}} \\ \epsilon_{\mathrm{HSG},\mathsf{c}} \\ \epsilon_{\mathrm{CG},\mathsf{c}} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{\mathrm{PRO}} & \sigma_{\mathrm{PRO},\mathrm{HSG},\mathsf{c}} & \sigma_{\mathrm{PRO},\mathrm{CG},\mathsf{c}} \\ & \sigma^2_{\mathrm{HSG}} & \sigma_{\mathrm{CG},\mathrm{HSG}} \\ & & \sigma^2_{\mathrm{CG}} \end{bmatrix} \right)$$

where

$$\begin{aligned} \sigma_{\mathrm{PRO},\mathsf{e},\mathsf{c}} &= \rho^{\mathsf{c}}_{\mathrm{PRO},\mathsf{e}} \sigma_{\mathrm{PRO}} \sigma_{\mathsf{e}} \\ \sigma_{\mathrm{CG},\mathrm{HSG}} &= \rho_{\mathrm{CG},\mathrm{HSG}} \sigma_{\mathrm{CG}} \sigma_{\mathrm{HSG}} \end{aligned}$$

  - $\rho^{\mathsf{c}}_{\mathrm{PRO},\mathsf{e}}$ captures complementarities in education and health investments beyond the ones incorporated in the 2nd stage of the model
(genes, parents, friends, neighborhood, etc.)
  - $\rho_{\mathrm{CG},\mathrm{HSG}}$ captures complementarities in the different education choices

- Note that only $\rho^{\mathsf{c}}_{\mathrm{PRO},\mathsf{e}}$ is cohort-specific

Calibration

- Match the joint distribution of e and y in each cohort c
  - → A total of 13 parameters and 15 statistics

- Identification
  - μ_{e}, μ_{y} drive the <u>average</u> share of e and y over time
  - σ^2_{e}, σ^2_{y}, $\rho_{\mathrm{CG,HSG}}$ drive <u>changes</u> in e and y over cohorts c as V_0^{eyc} changes (due to changes in wages and tuition fees)
  - $\rho^{\mathsf{c}}_{\mathrm{PRO},\mathsf{e}}$ residually matches the <u>joint distribution</u> of e and y in each cohort c

Calibration

Parameters

Parameter	Value	Parameter	Value	Parameter	Value
μ_{PRO}	11.3	$\rho^{1930}_{\text{PRO,HS}}$	0.02	$\rho^{1930}_{\text{PRO,COL}}$	0.01
μ_{HSG}	8.9	$\rho^{1950}_{\text{PRO,HS}}$	0.02	$\rho^{1950}_{\text{PRO,COL}}$	0.01
μ_{CG}	35.7	$\rho^{1970}_{\text{PRO,HS}}$	0.04	$\rho^{1970}_{\text{PRO,COL}}$	0.13
σ_{PRO}	10.6				
σ_{HSG}	1.6				
σ_{CG}	14.6				
$\rho_{\text{HS,COL}}$	0.0				

Calibration

Fit: marginal distributions

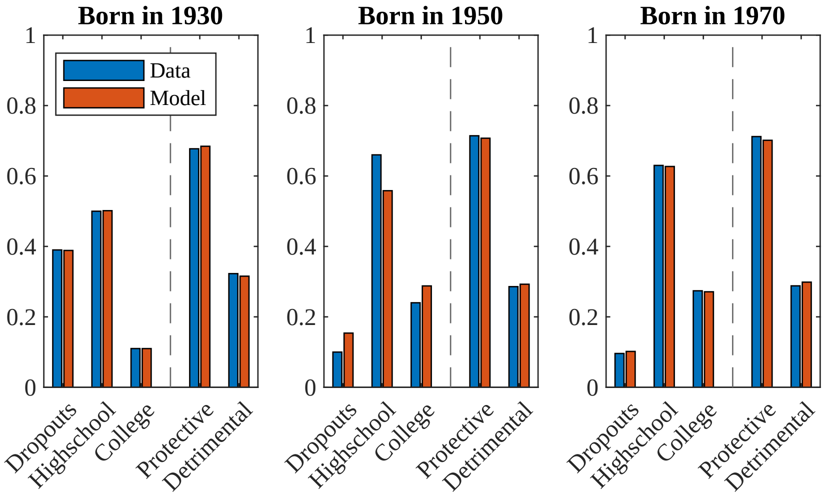


Figure: Marginal distributions: Education and Health Behavior

Calibration

Fit: conditional distributions

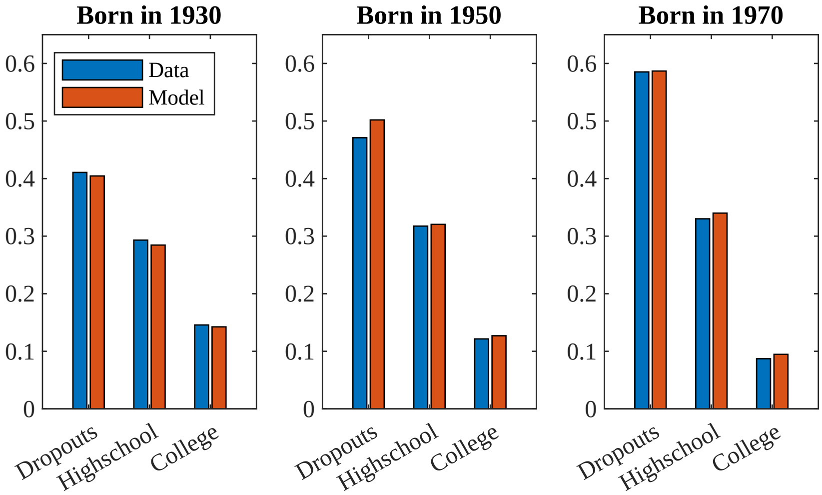


Figure: Conditional distribution of Detrimental Behavior by Education

Two questions

1. Why do individuals with higher education opt for more protective health behaviors?
2. To what extent the rise in earnings inequality has led to increased health inequalities?

Mechanisms

- Our model incorporates various mechanisms through which the incentive to adopt a given health behavior varies across educational choices:

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

   Higher expected income for the more educated motivates healthier behavior as the value of life increases

2. Complementarity of health investments: $\Gamma_t^{\mathsf{ey}}(h'|h)$

   Gains in life expectancy due to health behavior are more favorable for those with college education

3. Early life complementarities: $\rho^{c}_{\mathrm{PRO},\mathsf{e}}$

   Costs of better lifestyles may be related to costs of education (genes, parents, friends, neighborhood)

Results cohort 1930s: summary

- The better income and health transitions of the more educated
  - Are key for education choices: make CG
  - Do not change much the marginal distribution of health behaviour
  - Narrow down the education gradients of health behavior
- Early life complementarities
  - Play no quantitatively relevant role

Results cohort 1930s: figures

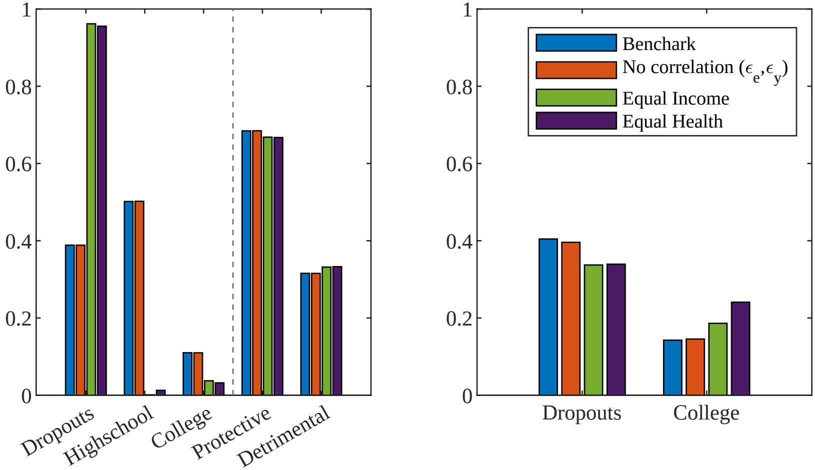


Figure: Education and Health Behavior

Results cohort 1930s: table

	$\Pr(\mathsf{y} = \text{PRO} \mid \mathsf{e})$			
Born in 1930	$\mathsf{e} = \text{HSD}$	$\mathsf{e} = \text{CG}$	$\Delta_{\text{CG,HAD}}$	ΔLE
Benchmark	40.4	14.2	26.2	6.2
$\rho^{c}_{\text{PRO},\mathsf{e}} = 0$	39.6	14.5	25.1	6.1
$w_t^{\text{CG}} = w_t^{\text{HSD}}$	33.7	18.6	15.1	5.3
$\Gamma_t^{\text{CG}}(h' \mid h) = \Gamma_t^{\text{HSD}}(h' \mid h)$	33.9	24.1	9.8	0.7

Mechanisms

1. Increase in the education wage premium

   a) Increases (decreases) the fraction of $e = \text{CG}$ ($e = \text{HSD}$) individuals

   b) Effect on $\Pr[y = \text{PRO} \mid e]$ and $LE(e)$

      - Direct: return on health investments increases more for the more educated,

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

        → This increases the education gradient in $\Pr[y = \text{PRO} \mid e]$ and $LE(e)$

      - Selection: worse pool of individuals (in terms of ϵ_{PRO}) within CG and HSD

        → *Ex ante* ambiguous effect on education gradient in $\Pr[y = \text{PRO} \mid e]$ and $LE(e)$

2. Increase in college enrollment fees

   – Reversed patterns

3. Increase in the correlation of initial conditions

   – Changes patterns of selection

Results: summary

- Increase in the education wage premium accounts for most of the action
  - Direct effect: key for the increase in PRO individuals and LE among CG
  - Selection effect: key for the fall in PRO individuals and LE among HSD

    (life prospects of HSD *did not fall as much to justify fall in* PRO*)*

- Absent financial frictions, the increase in tuition has very small traction

- 1/4th of the increase in the gradients accounted for the increase in the initial correlation between education and lifestyle shocks

Three mechanisms

	CG	HSD	$\Delta_{\text{CG,HSD}}$	explained
$\Delta\mathsf{Pr}(\mathsf{e})$	16.1	-28.7	44.8	
$\Delta\mathsf{Pr}(y = \text{PRO} \mid \mathsf{e})$	4.8	-18.3	23.0	
$\Delta\mathsf{LE}(\mathsf{e})$	0.4	-1.4	1.8	

Three mechanisms

	CG	HSD	$\Delta_{\text{CG},\text{HSD}}$	explained
$\Delta\text{Pr}(\text{e})$	16.1	-28.7	44.8	
Wage increase	16.1	-29.1	45.2	101%
$\Delta\text{Pr}(y = \text{PRO} \mid \text{e})$	4.8	-18.3	23.0	
Wage increase	1.6	-15.3	17.0	74%
$\Delta\text{LE}(\text{e})$	0.4	-1.4	1.8	
Wage increase	0.1	-1.2	1.3	74%

Three mechanisms

	CG	HSD	$\Delta_{\text{CG,HSD}}$	explained
ΔPr (e)	16.1	-28.7	44.8	
Wage increase	16.1	-29.1	45.2	101%
Tuition increase	-0.2	0.1	-0.3	-1%
ΔPr ($y = \text{PRO} \mid$ e)	4.8	-18.3	23.0	
Wage increase	1.6	-15.3	17.0	74%
Tuition increase	0.2	-0.3	0.4	2%
ΔLE(e)	0.4	-1.4	1.8	
Wage increase	0.1	-1.2	1.3	74%
Tuition increase	0.0	-0.0	0.0	2%

Three mechanisms

	CG	HSD	$\Delta_{\text{CG,HSD}}$	explained
ΔPr (e)	16.1	-28.7	44.8	
Wage increase	16.1	-29.1	45.2	101%
Tuition increase	-0.2	0.1	-0.3	-1%
Change initial conditions	0.2	0.0	0.2	1%
ΔPr $(y = \text{PRO} \mid \text{e})$	4.8	-18.3	23.0	
Wage increase	1.6	-15.3	17.0	74%
Tuition increase	0.2	-0.3	0.4	2%
Change initial conditions	4.4	-1.6	6.0	26%
ΔLE(e)	0.4	-1.4	1.8	
Wage increase	0.1	-1.2	1.3	74%
Tuition increase	0.0	-0.0	0.0	2%
Change initial conditions	0.4	-0.1	0.5	26%

Wage changes

	CG	HSD	$\Delta_{\text{CG,HSD}}$	explained
$\Delta\text{Pr}(\text{e})$	16.1	-28.7	44.8	
$\Delta\text{Pr}(y = \text{PRO} \mid \text{e})$	4.8	-18.3	23.0	
$\Delta\text{LE}(\text{e})$	0.4	-1.4	1.8	

Wage changes

	CG	HSD	$\Delta_{\text{CG,HSD}}$	explained
$\Delta\text{Pr}(\text{e})$	16.1	-28.7	44.8	
Wage increase	16.1	-29.1	45.2	101%
$\Delta\text{Pr}(y = \text{PRO} \mid \text{e})$	4.8	-18.3	23.0	
Wage increase	1.6	-15.3	17.0	74%
$\Delta\text{LE}(\text{e})$	0.4	-1.4	1.8	
Wage increase	0.1	-1.2	1.3	74%

Wage changes

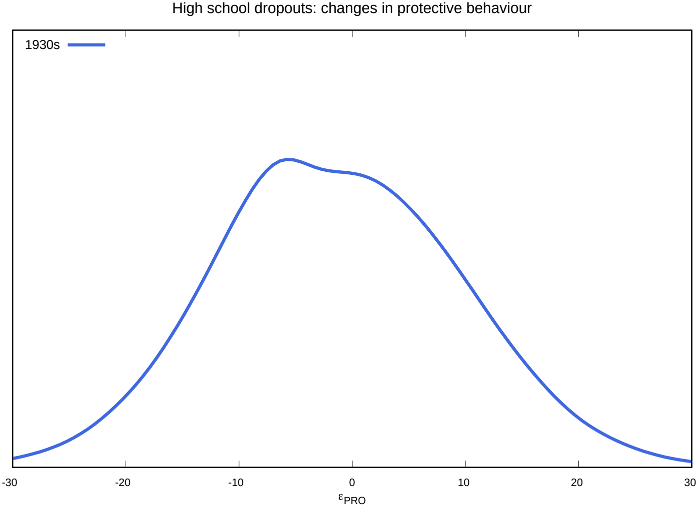
	CG	HSD	$\Delta_{\text{CG},\text{HSD}}$	explained
ΔPr (e)	16.1	-28.7	44.8	
Wage increase	16.1	-29.1	45.2	101%
ΔPr (y = PRO | e)	4.8	-18.3	23.0	
Wage increase	1.6	-15.3	17.0	74%
Direct effect	2.0	-1.5	3.4	26%
ΔLE(e)	0.4	-1.4	1.8	
Wage increase	0.1	-1.2	1.3	74%
Direct effect	0.2	-0.1	0.3	26%

Wage changes

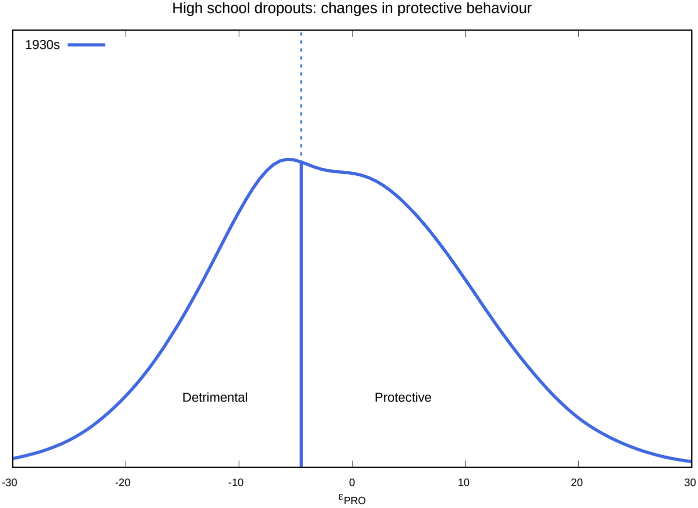
	CG	HSD	$\Delta_{\text{CG,HSD}}$	explained
$\Delta \Pr(e)$	16.1	-28.7	44.8	
Wage increase	16.1	-29.1	45.2	101%
$\Delta \Pr(y = \text{PRO} \mid e)$	4.8	-18.3	23.0	
Wage increase	1.6	-15.3	17.0	74%
Direct effect	2.0	-1.5	3.4	26%
Selection effect	-0.4	-13.8	13.6	59%
$\Delta LE(e)$	0.4	-1.4	1.8	
Wage increase	0.1	-1.2	1.3	74%
Direct effect	0.2	-0.1	0.3	26%
Selection effect	-0.1	-1.1	1.0	59%

Wage changes

1930s



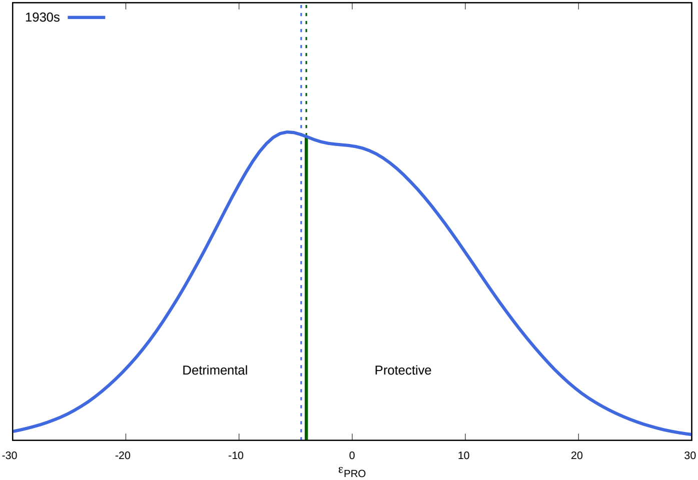
Wage changes
1930s



Wage changes

1930s to 1970s: direct effect

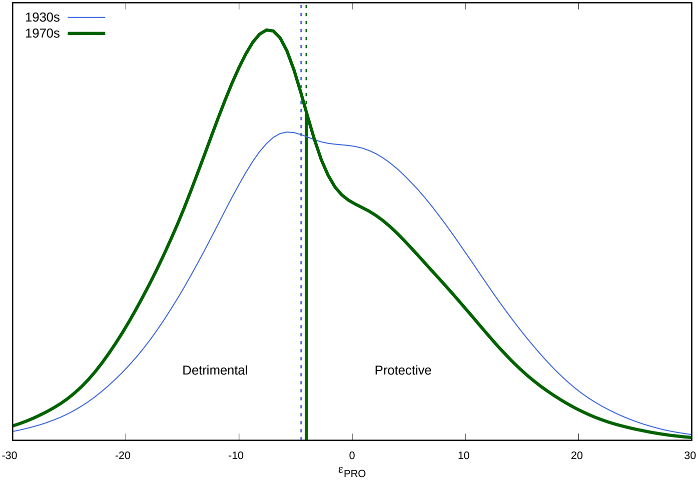
High school dropouts: changes in protective behaviour



Wage changes

1930s to 1970s: selection effect

High school dropouts: changes in protective behaviour



Conclusions

- We develop a econometric model that allows us to incorporate lifestyles into an HA model to analyze complementarities in education and health investments
- Differences in lifestyles across education groups can account for around $1/3$ of the LE gradient
- Both income differences and larger health advantages for college individuals drive differences in lifestyles across education groups.
- Our model implies that increases in wage inequality are tightly linked to increases in health inequality.
- Changes in selection account for most of the decline in lifestyle choices made by the high-school dropouts.