Healthy Habits and Inequality*

Dante Amengual
CEMFI

Jesús Bueren
EUI

Josep Pijoan-Mas
CEMFI and CEPR

February 2023

Abstract

There are two important features of inequality in health outcomes across individuals. First, there is an important socio-economic gradient in health outcomes. Second, the gradient has widened over time. In this paper we study the role of differences in lifestyle habits and risky health behavior as a driver of these two facts. We start by jointly estimating latent lifestyle habit types and health dynamics in both the HRS and the PSID. We do so by exploiting data on health behavior (preventive tests, substance abuse, obesity) and health outcomes (survival and health state) through MCMC methods. We find the following results. First, health behavior is well represented by two types: protective and harmful. Second, the types generate large gradients in life expectancy: at age 50 there are 8 years of life expectancy difference between protective and harmful types. Third, healthy habit types are correlated with education (but also carry independent information): the different incidence of healthy habits types across education groups explains 40% of the education gradient in life expectancy (but the healthy habit gradient is still large within education groups). We next build a life cycle heterogeneous agents model with idiosyncratic labor market and health risks. In the model, education and lifestyles are jointly chosen early in life. We first use the model to understand why more educated individuals tend to choose health protective lifestyles. Second, we use the model to quantify how the increase in the college wage premium has led to an increase in the education gradient of life expectancy.

*Emails: amengual@cemfi.es, jesus.bueren@eui.eu, pijoan@cemfi.es

1 Introduction

In recent decades, Western economies have witnessed an increase in both income and health inequalities, a phenomenon that has attracted considerable attention from economists, demographers and health researchers alike. This trend is underpinned by two fundamental observations: a strong correlation between economic and health outcomes (Kitagawa and Hauser, 1973; Pijoan-Mas and Ríos-Rull, 2014; Chetty et al., 2016) and a widening educational gap in health outcomes (Preston and Elo, 1995; Meara et al., 2008; Case and Deaton, 2015). However, the precise mechanisms linking economic status and health are not well understood. Our study seeks to address this gap by examining the role of lifestyle factors and health risk behaviors.

Lifestyle habits are important determinants of health outcomes, playing a significant role in shaping health inequalities. Firstly, differences in lifestyle factors and health risk behaviors such as exercise habits, dietary patterns, or smoking have been consistently identified as key drivers of health disparities (Li et al., 2018; Zaninotto et al., 2020). Individuals who engage in healthier behaviors are more likely to experience positive health outcomes, while those with unhealthy habits are at greater risk of developing chronic diseases and experiencing premature mortality. Secondly, research has consistently shown that individuals with higher levels of education tend to adopt healthier lifestyles (Lantz et al., 1998; Polvinen et al., 2013). Hence, heterogeneous lifestyle choices may contribute to the educational gradient in health outcomes, wherein individuals with lower levels of education are disproportionately affected by poor health outcomes due to less favorable lifestyle choices.

Our study aims to firstly measure the impact of different lifestyle habits on both health dynamics and economic outcomes. Secondly, we seek to understand the early life joint determination of education and lifestyle habits, particularly exploring why there exists an education gradient in lifestyle choices. Finally, given these ingredients, we aim to assess the extent to which increases in labor earnings inequality across education groups can drive increases in health inequalities.

Using data from the Health and Retirement Study (HRS) and the Panel Study of Income Dynamics (PSID), we exploit a rich array of health behavior indicators. These include preventive tests, substance abuse, and exercise habits, among others. Ideally, we would incorporate all this information into a structural model. However, challenges arise due to the noisy nature of observed health behaviors, which are imperfectly correlated across individuals and over time. Additionally, the curse of dimensionality poses a significant obstacle. To address these issues, our study makes a novel contribution by developing a methodology to reduce the dimensionality of the data. This involves identifying permanent patterns in lifestyle behavior, leveraging cross-sectional and panel variation on health behavior and examining the relationship between health behavior and health dynamics.

We find that health behavior can be parsimoniously represented by two distinct lifestyles: protective and detrimental. At age 50, there exists a substantial life expectancy gradient of 8

years between individuals with protective and detrimental lifestyles. Moreover, we observe a strong correlation between lifestyle habits and education, with harmful behaviors being more prevalent among the less educated. Within education groups, we find larger gradients in lifestyle habits for college graduates than for high-school dropouts. Furthermore, differences in lifestyles across education groups explain approximately 40% of the education gradient in life expectancy. Finally, we find an increasing dispersion in lifestyles across education groups for individuals born in the 1930s and the 1970s.

Then, in order to understand the joint determination of education and health behavior choices, we propose an heterogeneous agents model comprising two distinct stages. In the first stage, individuals make a one-time early-life education and health behavior choice. In this stage, individuals exhibit heterogeneity in their costs associated with education and engagement in protective health behaviors.[1] In the second stage, during the working/retirement phase, individuals are modeled to solve a standard life-cycle model with idiosyncratic labor income and health risks, with outcomes conditioned on their specific education and lifestyle choices made in the initial stage.

Our economic model incorporates complementarities between education and lifestyle investments because of two important reasons. First, an extra year of life is more valuable with higher consumption possibilities. This means that the value of a health-protective lifestyle is larger for the mode educated. Second, as implied by our econometric model, the benefit in health transitions of investing in protective health behavior is also larger for higher-educated individuals. Furthermore, these complementarities between health and education investment influence the selection of individuals into different education categories. Specifically, individuals facing lower costs associated with health behavior are more inclined to opt for higher education investments.

We calibrate our model to accurately replicate the joint distribution of education and lifestyle choices for cohorts born in 1930 and 1970, as well as to match the median savings across different education and health behavior categories. In the model, cohorts differ in their education-specific average labor earnings. Notably, our model effectively captures the overall wealth distribution and is able to explain 50% of the observed increase in life-expectancy inequalities across education groups between the 1930 and 1970 cohorts.

Using our model as a laboratory economy, we identify the mechanisms driving the observed disparities in life expectancy and health outcomes across education groups. Our model incorporates three key mechanisms to explain why higher educated individuals tend to adopt protective health behaviors more frequently. Firstly, higher expected income for the more educated incentivizes healthier behavior by increasing the value of life. Secondly, our econometric model on health dynamics reveals that gains in life expectancy due to protective health behavior are more favorable for those with a college education. Lastly, the selection mechanism suggests that early in life,

[1]The heterogeneous costs of education may arise due to differences in cognitive abilities or differences in distance to quality education centers. Regarding the heterogeneous costs of lifestyle choices, there is a recent literature relating brain characteristics of young adolescents with later health risk behavior, see (Xiang et al., 2023).

individuals facing lower costs of protective behavior are more likely to choose higher education, given the complementarities between health and education investment.

We find that both the income and health advantage are equally important and explain most of the differences in health behavior choices across education groups. Selection on the other hand, plays a quantitatively more modest role. With respect to changes across cohorts, we find worse economic conditions of the high-school dropout explain 78% of the fall in life-expectancy while 22% is explained by selection. Better economic outcomes for the college educated explain all of the increases in life-expectancy across cohorts.

Our paper is related to a growing literature that employs life cycle models to quantify how heterogeneous health dynamics impact economic outcomes. Similar to our approach, Hosseini et al. (2021) and De Nardi et al. (2017) utilize dynamic panel models to characterize health dynamics and estimate significant welfare losses associated with adverse health episodes. While the former investigates the effects of social security and disability insurance, the latter demonstrates that the impact of health on economic outcomes largely depends on fixed characteristics predetermined earlier in life. In contrast, our study provides an econometric framework to leverage information on health behaviors, enabling us to interpret these fixed types and model health behavior decisions.

The paper is also related to previous work featuring endogenous monetary health investments. Fonseca et al. (2021) investigate the causes behind the increase in health spending and longevity in the United States from 1965 to 2005, attributing technological progress to half of the rise in life expectancy over the period. In contrast, our study focuses on exploring the role of health behavior choices.

The paper importantly connects with previous studies that model endogenous health behavior investments (Cole et al., 2019; Mahler and Yum, 2022; Margaris and Wallenius, 2023). However, while these models allow agents to adjust their behavior while keeping their education choices as given, our approach differs in that we leverage long-term changes in health behavior across cohorts to identify model parameters. These studies often assume significant costs associated with changing lifestyles to match the lack of switching observed in the data. In contrast, our model capitalizes on the persistence of health behavior, which is often established early in life. This simplifying assumption enables us to explore the joint determination of health behavior and education choices instead.

The remainder of the paper is organized as follows: Section 2 and Section 3 present the econometric model used to identify lifestyles and its results, respectively. Section 4 outlines the economic model, while Section 5 details the calibration strategy. Section 6 presents the quantitative findings, followed by concluding remarks in Section 7.

2 An econometric model of health dynamics with latent types

We combine data from the HRS and the PSID to create an unbalanced panel of individuals $i = 1, \ldots, N$ followed for $t = 0, \ldots, T_i$ periods. For each individual and period we observe standard demographic variables: cohort of birth $c_i \in \{c_{10}, c_{30}, c_{50}, c_{70}, c_{90}\}$ (individuals born around 1910, 1930, 1950, and 1990), gender $s_i \in \{s_m, s_f\}$ (male, female), education $e_i \in \{e_c, e_h, e_d\}$ (college degree, high school degree, high school dropout), and age $a_{it} \in \{25, 26, \ldots, 100\}$, plus a wide array of health-related variables, which we classify into two groups: health outcomes and health behaviour. The health outcome $h_{it} \in H \equiv \{h_g, h_b, h_d\}$ takes three values: good health (h_g), bad health (h_b), or dead (h_d), which is an absorbing state. We build this variable by use of the 5-category self-rated health variable (where the best two categories form the good health state) and the information on survival. The health behavior vector $\mathbf{z}_{it} = \{z_{1,it}, z_{2,it}, \ldots, z_{N_z,it}\}$ contains information on N_z different categorical variables $z_{m,it} \in \{0, 1\}$ describing whether individual i in period t does some particular action. These actions are whether the individual has taken a cancer test (prostrate or mammography), a cholesterol test, or a flu shot in the last year, whether the individual drinks (more than two drinks on the day she drinks), whether the individual smokes, and whether the individual is performs any type of physical activity.[2]

We assume that both the observed health behaviour $\mathbf{z}_{it}$ and health outcomes h_{it} depend on some unobserved time-invariant latent variable $y_i \in Y \equiv \{y_1, y_2, \ldots, y_{N_y}\}$, with $N_y < N_z$. We interpret the latent variable as lifestyle / healthy habit type —determined before the start of working life— that captures the idea that individuals differ in their propensity to undertake actions that are good for their health. This notion of an unobserved latent variable is important because in the data observed health behaviour is imperfectly correlated across individuals and over time. We want the latent types to also affect health outcomes because we want the classification of individuals based on behavior to be meaningful in terms of health dynamics.

We aim to estimate the parameters of our econometric model by maximizing the probability of observing the joint sequence of health behaviours $\mathbf{z}_i^T$ and health outcomes $\mathbf{h}_i^T$ of each individual i conditional on the demographic variables and initial health. We can write the likelihood of the data as a mixture model:

$$p\left(\mathbf{z}_i^T, \mathbf{h}_i^T | c_i, s_i, e_i, a_i^T, h_{i0}\right) = \sum_{y \in Y} p\left(\mathbf{z}_i^T | y, h_i^T, a_i^T\right) p\left(\mathbf{h}_i^T | y, s_i, e_i, a_i^T, h_{i0}\right) p\left(y | c_i, s_i, e_i, a_{i0}, h_{i0}\right) \quad (1)$$

where the elements in the right hand side, the probability of observing a sequence of health behaviours, the probability of observing a sequence of health outcomes, and the initial distribution of types, are explained in Section 2.1, 2.2, and 2.3 respectively. It is worth discussing here the role

[2] Some of the health outcome and health behavior variables for a given individual may be missing for some period t. Indeed, in the PSID we do not observe the variables for health protective behaviour (cancer test, cholesterol test, flu shot). We take missing observations into account under the assumption that they occur completely at random, but we abstract from them in the model description to simplify the exposition.

of cohort c_i.

2.1 Health behaviour

We assume that the probability of individual i in period t reporting the m^{th} behavior ($h_{m,it} = 1$) depends on the health behaviour type y_i but also on age a_{it}, on health status h_{it}, and on an idiosyncratic shock $\varepsilon_{m,it}$. The idea is that the association of observed behavior such as smoking or cancer tests with types may differ over age and across health states. Instead, conditional on these variables, we impose that health behaviour does not depend on cohort c_i, gender g_i, or education e_i. We do so because we want the definition of types to be stable across demographic groups. This makes the types comparable across groups, and it lets the variation in behavior across demographic groups to arise from their different distribution of types.

We assume that, conditional on y_i, a_{it}, and h_{it}, the shock $\varepsilon_{m,it}$ is iid across m, i, and t, and that it follows a standard normal distribution. Then, we can model the probability of individual i in period t reporting the mth behavior ($h_{m,it} = 1$) as a probit model. In particular, let $z^*_{m,it} = z^*_m(y_i, a_{it}, h_{it})$ be a latent variable such that $z^*_{m,it} + \varepsilon_{m,it} > 0 \Rightarrow z_{m,it} = 1$.[3] Then, the probability of observing the mth behavior is given by $\Phi\left(z^*_{m,it}\right)$ where $\Phi()$ is the cdf of the standard normal distribution, and the probability of observing a sequence $\mathbf{z}_i^T$ of health behaviour vectors $\mathbf{z}_{it}$ for individual i is given by,

$$p\left(\mathbf{z}_i^T | y_i, h_i^T, a_i^T\right) = \Pi_{t=1}^T \Pi_{m=1}^M \left[\Phi\left(z^*_{m,it}\right)\right]^{z_{m,it}} \left[1 - \Phi\left(z^*_{m,it}\right)\right]^{1-z_{m,it}} \tag{2}$$

2.2 Health dynamics

We assume that health dynamics for individual i depends on gender s_i, education e_i, health behaviour type y_i, and age a_{it}. The dependence of health dynamics on gender and education is meant to capture differences in health outcomes associated to these variables that are not captured by differences in health behaviour types across these demographic groups. The absence of cohort c_i from the set of conditioning variables is an identification assumption. We do not observe full lifespans for individuals in different cohorts. Rather, we combine information on health dynamics at old ages from individuals born in earlier cohorts with information on health dynamics at young ages from individuals born in later cohorts. This is standard in estimation of models of health dynamics with survey data, see for instance Pijoan-Mas and Ríos-Rull (2014). Hence, our assumption is that health dynamics of individuals of given gender and education are, conditional on type, identical across cohorts. However, it is important to note that this allows gender and education health dynamics to differ across cohorts due to the different composition of types.

We assume that conditional on gender s_i, education e_i, health behaviour type y_i, and age a_{it},

[3] We model $z^*_m(y_i, a_{it}, h_{it})$ as a flexible low order polynomial on age.

the evolution of health outcomes is markovian, that is, only depends on one lag of health outcomes. We model survival and health transition probabilities as a nested probit model. In the first nest, individuals are exposed to a survival shock $\varepsilon_{S,it} \sim N(0,1)$. Let $h^*_{S,it} = h^*_S(y_i, s_i, e_i, a_{it-1}, h_{it-1})$ be a latent variables such that $h^*_{S,it} + \varepsilon_{S,it} > 0 \Rightarrow h_{it} \in \{h_g, h_b\}$. Then, the survival probability between $t-1$ and t is given by $\Phi\left(h^*_{S,it}\right)$. In the second nest, conditional on surviving, individuals receive a health shock $\varepsilon_{G,it} \sim N(0,1)$. Let $h^*_{G,it} = h^*_G(y_i, s_i, e_i, a_{it-1}, h_{it-1})$ be a latent variable such that $h^*_{G,it} + \varepsilon_{G,it} > 0 \Rightarrow h_{it} = h_g$. Then, the probability of good health transition between $t-1$ and t (conditional on survival) is given by $\Phi\left(h^*_{G,it}\right)$.[4] With these elements, one can write the probability $p\left(\mathbf{h}_i^T | y_i, s_i, e_i, a_i^T, h_{i0}\right)$ of observing a sequence $\mathbf{h}_i^T$ of health outcomes h_{it} conditional on the health behavior type y_i, the demographic variables s_i, e_i, and a_{it}, and the health h_{i0} in the first period that the individual is observed.

2.3 Distribution of health types

The final element we need is the fraction of individuals of each type, that is, the prior probability of each individual of being of a given health behaviour type before observing the data on health behavior and health outcomes. In particular, we define $p(y_i | c_i, s_i, e_i, a_{i0}, h_{i0})$ as the probability that individual i born into cohort c_i, of gender s_i, education e_i, first observed with age a_{i0} and health h_{i0} is of type y_i. One could add this term to the likelihood function and estimate it non-parametrically from the data. However, because the panel is not balanced and the first age of observation of many individuals is quite advanced (a_{i0} is large), there would an identification problem in that observed changes of health behavior $\mathbf{z}_{it}$ with age (for instance, the decline with age of the incidence of smoking) could not be separated into (a) changes over age of health behavior conditional on type $p\left(\mathbf{z}_i^T | y_i, h_i^T, a_i^T\right)$ and (b) changes in the distribution of types with age $p(y_i | c_i, s_i, e_i, a_{i0}, h_{i0})$. Therefore, we exploit the model of health dynamics described above to obtain an expression for how the posterior distribution of types conditional on observables evolves with age. This leaves us with the need to only estimate the prior probabilities of each individual of being of each type at the initial age of 25.

In order to do so, note that we can write,

$$p(y_i | c_i, s_i, e_i, a_{it}, h_{it}) = \frac{p(y_i, h_{it} | c_i, s_i, e_i, a_{it})}{\sum_{h \in H} p(y_i, h | c_i, s_i, e_i, a_{it})} \tag{3}$$

The joint probability $p(y_i, h_{it} | c_i, s_i, e_i, a_{it})$ can be decomposed as,

$$p(y_i, h_{it} | c_i, s_i, e_i, a_{it}) = \sum_{h_{it-1} \in H} p(h_{it} | y_i, h_{it-1}, s_i, e_i, a_{it-1})\, p(y, h_{it-1} | c_i, s_i, e_i, a_{it-1}) \tag{4}$$

The first term in the right hand side describes the health dynamics and it has been discussed in Section 2.2. The second term in the right hand side is the same as the left hand side, just one

[4] We model $h^*_S(y_i, s_i, e_i, a_{it-1}, h_{it-1})$ and $h^*_G(y_i, s_i, e_i, a_{it-1}, h_{it-1})$ as flexible low order polynomials on age.

period before. Hence, we can use equation (4) recursively up to age 25, which is our initial age, that is, up to $p\left(y_i, h_{it} | c_i, s_i, e_i, a_{it} = 25\right)$. This term describes the joint probability of an individual in cohort c_i, of gender s_i and education e_i of being of health h_{it} and type y_i at age 25. We decompose this probability in two pieces:

$$p\left(y_i, h_{it} | c_i, s_i, e_i, a_{it} = 25\right) = p(y_i | c_i, s_i, e_i, a_{it} = 25, h_{it}) p(h_{it} | c_i, s_i, e_i, a_{it} = 25)$$

The second term in the right hand side describes the share of individuals of age 25 of given cohort c_i, gender s_i, and education e_i that have health h_{it}. This can be measured directly in the PSID for many but not all cohorts. We thus assume that conditional on education and gender, the probability of good and bad health at age 25 does not change across cohorts and we set it equal to the one that we observe in the PSID.[5] The vast majority of observed individuals of age 25 display $h_{it} = h_g$, so we set this probability to one. The first term in the right hand side describes the fraction of individuals of age 25, cohort c_i, gender s_i, education e_i, and health h_{it} that are of type y_i. We model this probability through a multinomial probit. That is, we define thresholds $y^*_{1,i} = y^*_1\left(c_i, s_i, e_i, h\right)$ and $y^*_{2,i} = y^*_2\left(c_i, s_i, e_i, h\right)$ that, given the realization of a shock $\varepsilon_{y,i} \sim N\left[0, 1\right]$, separates individuals into types.

3 Results from the econometric model

We put together the PSID from 1999 to 2019 and the HRS between 1996 and 2018 to estimate the econometric model by use of bayesian methods as directly maximizing the likelihood function is numerically very complex. In particular, we use flat priors and aim to recover the posterior distribution of all the parameters and the latent variables that classify the health behaviors to which each individual belongs. To do so, we use a Gibbs sampling procedure. In what follows we describe the results.

3.1 Health behaviour

In our main estimation and in favor of parsimony, we choose $N_y = 2$, that is, we classify individuals into two latent groups. Figure 1 reports the probability of displaying each health behaviour z_{it} as a function of health type y_i and age a_{it} and for $h_{it} = h_g$ (the case $h_{it} = h_b$ is not too different). Individuals in one group have higher likelihood of reporting health protective habits (cancer test, cholesterol test, flu shot, and exercise) and lower probability of reporting health detrimental habits (smoking and drinking). We label this group as *protective* (solid green line). Individuals in the other group, the probability of reporting all the protective habits is lower and their smoking probability is very high. We label this group as detrimental (red dashed line).

[5] The probability of being in good health at age 25 varies between 77% for dropout females to 98% for male college graduates.

3.2 Distribution of health types

Our estimation assigns a different fraction of individuals to each type y depending on cohort c, gender s, and education e. The first column in Table 1 reports these fractions for the 1950 cohort. The main finding is that there is a large educational gradient of health behaviour types: the share of *protective* individuals grows from 42.9% to 91.6% as we move from high school dropout to college graduate. The pattern is similar similar among women, with the difference that among women there are less *harmful* types among the high-school dropouts. These figures are large and reflect a strong correlation between education and lifestyle. In a sense, these numbers are not too surprising: it is well known that the incidence of smoking, drinking, and obesity declines with education, and it is easy to check that —in the HRS— the share of individuals taking cancer tests, cholesterol tests, and flu shots increases with education. It is therefore natural that our classification of individuals into types according to the observed health-related behaviour retains the education gradient of these latter variables. However, we highlight that our classification of individuals into types also uses longitudinal information on health dynamics *conditional on education*. That is, it also uses the fact that within education *protective* types have better health dynamics than *harmful* types.

A very interesting question is how has the distribution of health types evolved across education groups over time. In Panels (a) to (c) of Figure 2 we report the distribution of types by education group from the 1910 to 1990 cohorts. Within the high school dropouts, the *harmful* types increased monotonically from 40% in the 1930 cohort to 75% in the 1990 cohort. This implies a severe deterioration in the lifestyle of individuals in the least educated group, which reverses a slight improvement in the type distribution between the 1910 and the 1930 cohorts. In contrast, among college educated individuals, there is little change in the share of *protective*, a slight decline of the *harmful* and a slight increase in the *detrimental*. All in all, this implies that the educational gradient in lifestyles has widened remarkably between the 1930 and the 1990 cohort. As we will see in the next Section, this will generate an increasing life expectancy gap across education groups.

3.3 Putting all together

Combining all previous results, we can compute life expectancies by gender, education, and health behaviour type and quantify the role of each covariate. In particular, in Table 1 we report total life expectancy, healthy life expectancy, and unhealthy life expectancy at age 50. We highlight three main results. First, as it is well known, the education gradient in life expectancy is large: males with a college degree live around 7.8 more years than males without a high school degree, while the gradient is 6.7 years among females. Second, there is a large gradient in life expectancy across health behaviour types *within* each education category: *protective* types live 7.6 more years than *harmful* types among males without a high school degree, 7.2 more years among males with a high school degree, and 8.9 more years among males with a college degree. For females, these numbers are 7.1, 7.2, and 6.7 respectively. And third, the different type composition across education groups

explains a big part of the educational gradient of life expectancy but there is still much left. In particular, if males without a high school degree had a distribution of health behaviour types as the males with a college degree, their average life expectancy would rise to 28.3 years (up from 24.6) and the life expectancy differential with college graduates would fall from 6.8 to 4.3. That is, the different distribution of health behaviour types across education groups among males accounts for 40% of the life expectancy differential between college graduates and high school dropouts. The numbers are of the same order of magnitude among females: the different distribution of health behaviour types across education groups accounts for 33% of the life expectancy differential between college graduates and high school dropouts.

Finally, in Panel (d) of Figure 2 we report the age-50 predicted life expectancy gap between college educated males and high school dropouts over different cohorts. In our estimation, different cohort have different life expectancy because of different composition of latent types as described in Panels (a) to (c) of Figure 2, but health dynamics conditional on type are identical across cohorts. This allows us to infer health dynamics at old ages of younger cohorts, for which most individuals are still alive today. Our finding show a growing education gap in life expectancy: from 6.8 years in the 1930 cohort to 9.1 years in the 1990 cohort, which follows a 0.7 year decline in the gradient between the 1910 and the 1930 cohorts. The education gradient of life expectancy that grows across cohorts is consistent with the estimates of a growing education gradient of life expectancy over time, but it is a different concept.

3.4 Health inequality and economic inequality

Our estimation results classify individuals within three different health behaviour categories based on observed health behaviour and health transitions. We have seen that these types correlate with education. The next question is whether these types also correlate with some economic outcomes. In this Section we show that wealth accumulation is positively linked to health behavior types. To do so, we need to recover the wealth distribution across the unobserved health behavior types. For this purpose, we model the observed wealth distribution as a mixture model. In order to separate the mass point at zero wealth and the distribution of wealth conditional on positive wealth we proceed in two stages. First, we write the distribution of (positive) wealth conditional on observables as:

$$p(w_{i,t}|e_i, a_{it}, z_i^T, h_i^T, w_{i,t} > 0) = \sum_{y \in Y} p(w_{i,t}|y, e_i, a_{it}, z_i^T, h_i^T, w_{i,t} > 0) p(y|e_i, z_i^T, h_i^T),$$

The first term in the right hand side is the conditional probability of observing wealth $w_{i,t}$. We assume that wealth conditional on age a, education e, and type y is lognormally distributed, that is, $\log p(w|y, e, a, w > 0) \sim N(\mu^1(y, e, a), \sigma^1(y, e))$. This implies that we are imposing that given a, e, and y, wealth is independent from h_i^T and z_i^T.

The second term in the right hand side gives the conditional distribution of types, which we

have estimated above, see Section 2.3. Hence, we only need to estimate $\mu^1(y, e, a)$ and $\sigma^1(y, e)$ for the sample of male individuals with positive asset holdings.[6]

Second, we can similarly write the probability of reporting zero (or negative) assets conditional on observables as

$$p(w_{it} = 0|e_i, a_{it}, z_i^T, h_i^T) = \sum_{y \in Y} p(w_{it} = 0|y, e_i, a_{it}, z_i^T, h_i^T) p(y|e_i, z_i^T, h_i^T),$$

where the first term on the right hand side is modelled as a probit, that is, it is given by $\Phi\left(w^*\left(y, e_i, a_{it}\right)\right)$, where the threshold $w^*\left(y, e_i, a_{it}\right)$ is modelled as a flexible low order polynomial on age. As above, we assume that this probability does not depends on h_i^T or z_i^T.

We present selected moments of the estimated wealth distribution conditional on age, education and type in scatter plots in Figure 3. As it is well known, wealth accumulation is positively correlated with education. What is interesting of our results is that, within each education category, wealth accumulation is also stronger for better health types. We would like to highlight two results. First, wealth accumulation is stronger for the *protective* type. Second, the difference in wealth accumulation across types is especially apparent within college educated individuals, and much smaller (or null) within the other two education categories.

These results could happen for two different reasons. On the one hand, better health types lead to better health trajectories and this could cause more wealth accumulation through higher income, lower medical expenditures, or anticipation of longer lifespan. On the other hand, the stronger wealth accumulation of better health types could be the result of selection. For instance, it could be that the unobserved health types are correlated with discount rates, which determine wealth accumulation. In the next Section we write a life cycle economic model to understand the relative importance of each explanation.

4 An economic model

Our economic model considers two distinct life stages. The *early life* stage is a static problem where young individuals belonging to a cohort c choose their education $e \in \{\text{HSD}, \text{HSG}, \text{CG}\}$ and their lifestyle or health-related behaviour $y \in \{\text{DET}, \text{PRO}\}$. This problem is a stand-in for choices and investments made by either parents in early childhood or young adults before entering the labor market. The objective is to maximize the expected value of starting working life with a given type (education and lifestyle) minus the (idiosyncratic) costs of choosing each type. This stage serves to account for the observed correlation between education and health types. The *adult life* stage is a dynamic life-cycle consumption-saving problem under uncertainty in both labor market and health outcomes where individuals differ in type (education and lifestyle). This stage serves

[6] We model $\mu^1(y, e, a)$ as a flexible low-order polynomial on age and $\sigma^1(y, e)$ non-parametrically, see Section XXX for details.

to link inequality in health and economic outcomes, and provides the value of starting life in each type, which is used in the *early life* stage.

4.1 Stage 1: early life

Let $V_0^{c,e,y}$ be the value of starting working life with a type (e, y) for individuals in cohort c. Before entering the labor market, young individuals choose their type by solving

$$max_{e,y}\Big\{V_0^{c,e,y} - \tau_e(\epsilon_e) - \tau_y(\epsilon_y)\Big\}$$

where τ_e and τ_y represents the cost of choosing a education and lifestyle, respectively. The education and health behavior cost is heterogeneous in the population. Agents differ in the cost that they pay depending on ϵ_y and ϵ_e.

The cost of being a high-school dropout is normalized to 0, while the cost of graduating from high-school is $\tau_e = \mu_e + \epsilon_e$ and the cost of graduating from college is $\tau_e = \mu_{\text{CG}}(\mu_e + \epsilon_e)$. ϵ_e is assumed to follow a normal distribution with mean 0 and variance σ_e. The cost of having a detrimental health behavior is normalized to 0, while the cost of protective behavior is $\tau_{\text{PRO}} = \mu_{\text{PRO}} + \epsilon_{\text{PRO}}$. ϵ_{PRO} is assumed to follow a log-normal distribution with mean $\sigma^2_{\text{PRO}}/2$ and variance σ^2_{PRO}.[7]

How can this framework generate correlation between choices such that better lifestyle types are more frequent among the high educated? The correlation will appear due to the complementarity of investments. For instance, let's consider the choice of lifestyle PRO over DET across education categories HSD and CG. If $V_0^{\text{CG,PRO}} - V_0^{\text{CG,DET}} > V_0^{\text{HSD,PRO}} - V_0^{\text{HSD,DET}}$ then more people will choose PRO over DET within the education group CG than within the education group CG. This difference in values is generated by the model of the *adult life* stage. In addition to this, non-linearities of τ_e and τ_y may increase or decrease the correlation between lifestyle and education.

4.2 Stage 2: adult life

4.2.1 Demographics, preferences, and shocks

The model period corresponds to two years. Individuals live for at most T periods but survival is stochastic every period. During the first $R-1$ periods of life people are exposed to health shocks, medical expenditure shocks, and labor income shocks. Individuals retire at age R, when they start receiving a retirement pension instead of stochastic labor income.

[7] The static nature of this problem overlooks the fact that more educated individuals start their working life later than less-educated individuals. The opportunity cost of labor market earnings while studying will hence be part of τ_e.

Preferences over consumption flows c_t are described by a standard CRRA period utility:

$$u(c_t) = \frac{(c_t/\bar{n})^{1-\sigma}}{1-\sigma} + \bar{b},$$

where σ is the risk-aversion, $\bar{n}$ is an age specific household size and $\bar{b}$ is a positive term to ensure that individuals in our model value their life. In the period when they die, individuals also derive utility from leaving a bequest of size k_t:

$$v(k_t) = b_0 \frac{(k_t + b_1)^{1-\sigma}}{1-\sigma}$$

where b_0 drives the strength of the bequest motive and b_1 the extent to which preferences for bequest increase with wealth.

Following the empirical model in Section 2, health h_t can be either good (h_g) or bad (h_b) and it evolves according to the age-dependent Markov chain $\Gamma_t^{e,y}(h)$, which depends on education e and health type behavior y. The survival probability $s_t^{e,y}(h)$ depends on health h but also (possibly) on education e and health-behavior type y.

Every period of their working life, individuals receive an exogenous income that we model in two components. First, there is an employment shock $\ell_t \in \{0, 1\}$ that determines if the individual has the chance of working in the labor market ($\ell_t = 1$) or not ($\ell_t = 0$). We model $\text{Prob}(\ell_t = 1 | t, e_t, h_t, \ell_{t-1})$ as a Probit model. This component aims to capture the prolonged non-working spells of some individuals in bad health that are at the core of health gradient of labor income. Second, conditional on working individuals receive labor income is given by,

$$\log w_t^{c,e,y}(h_t, \xi_t, \epsilon_t) = \omega_t^{c,e,y}(h_t) + \xi_t + \epsilon_t,$$

where $\omega_t^{c,e,y}(h)$ is a deterministic component depending on cohort, age, education, lifestyle, and health, while ξ_t and ϵ_t are persistent and transitory shocks. The initial value of the persistent component ξ_0 is drawn from a normal distribution with mean zero and variance $\sigma_{\xi_0}^2$. Whenever the worker is attached to the labor force ($\ell_t = \ell_{t-1} = 1$), the stochastic persistent component is assumed to follow a Gaussian AR(1) process with persistence ρ_ξ and variance of the innovations σ_ν^2. If the worker enters the labor force after a non-working spell ($\ell_t = 1$ and $\ell_{t-1} = 0$) then ξ_t is sampled from the unconditional distribution of ξ_t. The transitory component is i.i.d. and distributed with a Gaussian distribution of variance σ_ϵ^2.

Finally, medical expenses are given by,

$$\log m_t^e(h_t, \zeta_t) = \mu_t^e(h_t) + \zeta_t$$

where $\mu_t^e(h_t)$ is a deterministic component, as a function of age t, education e and health h_t, while ζ_t is an i.i.d gaussian white noise with variance σ_ζ^2.

4.2.2 Taxation and social transfers

We model the tax system as follows. Working households pay payroll taxes, which include the Medicare tax (τ_{MCR}) and the Social Security tax (τ_{ss}). The latter only affects earnings below w_{ss}. There is a progressive labor income tax $T(w)$ which we specify as Heathcole et al. (2020)

$$T(w) = w - a_{\tau 0} w^{1-a_{\tau 1}}$$

We represent several existing mean-tested programs in a stylized way through a public safety-net program. This program guarantees every household a minimum income floor $\underline{x}$. Retirees receive Social Security benefits. In practice, these payments depend on an individual's history of earnings. To capture the existing variation in pension benefits without increasing computational costs, we approximate the benefits using the following approach. First, we divide individuals into groups based on their labor force participation just before retirement, their last draw of the persistent productivity shock and on their education and health behavior type. Then, for each group, we compute average earnings over the 17 model periods (34 years) with the highest earnings. Then we apply the Social Security benefits formula to these average earnings.

4.2.3 The optimization problem

At the beginning of the period, working-age individuals of type (c, e, y) and age t learn their cash in hand x_t, labor force status ℓ_t, persistent component of productivity ξ_t (conditional on participating in the labor force), health state h_t, and medical expenditure shock ζ_t. All these variables form the state of the individual: x_t is payoff relevant in the current period and the other variables serve to predict next period outcomes. Based on this information, individuals decide on consumption c_t and savings k_{t+1}. At the end of the period, there are new realizations of the shocks for survival, health, labor force participation, productivity (persistent and transitory), and medical expenses. The timing for retired individuals is similar, with the difference that there are no employment or labor earnings shocks.

The optimization problem for working age individuals ($t < R$) can be written in recursive form as:

$$V_t^{c,e,y}(x,h,\ell,\xi) = \max_{c,k'} \left\{ u(c) + \beta s_t^{e,y}(h) \sum_{h'} \Gamma_t^{e,y}(h)\, \mathbb{E}_{\ell,\xi,\zeta}[V_{t+1}^{c,e,y}(x',h',\ell',\xi')] + \beta\left(1 - s_t^{e,y}(h)\right) v(k') \right\}$$

s.t.

$$c + k' = x$$

$$x' = \min\left\{ w_{t+1}^{c,e,y}(h',\xi',\epsilon')\ell' - Tax + (1+r)k' - m_t^e(h,\zeta),\ \underline{x} \right\}$$

$$Tax = T\left(w_{t+1}^{c,e,y}(h',\xi',\epsilon')\ell'\right) + \tau_{MCR} w_{t+1}^{c,e,y}(h',\xi',\epsilon')\ell' + \tau_{ss} \min\{w_{t+1}^{e,y}(h',\xi',\epsilon')\ell', w_{ss}\}$$

The optimization problem for retired individuals ($t \geq R$) can be written as:

$$V_t^{c,e,y}(x,h,\xi_{R-1}) = \max_{c,k'} \left\{ u(c) + \beta s_t^{e,y}(h) \sum_{h'} \Gamma_t^{e,y}(h)\, \mathbb{E}_\zeta[V_{t+1}^{c,e,y}(x',h',\xi_{R-1})] + \beta\,(1 - s_t^{e,y}(h))\, v(k') \right\}$$

s.t.

$$c + k' = x$$

$$x' = \min\left\{ p^{c,y,e}(\xi_{R-1}) + (1+r)k' - m_t^e(h,\zeta),\, \underline{x} \right\}$$

$$Tax = T\left(p^{c,y,e}(\xi_{R-1})\right)$$

where retirement income is constant and determined by the last productivity/labor force shock before retirement ξ_{R-1}.

5 Calibration

5.1 Working/retirement phase

Parameters related to demographics, taxes, social security, and the stochastic process for earnings are calibrated outside the model (See online appendix for details). We calibrate the discount factor β to 0.98 a standard value in the literature and calibrate the remaining parameters internally by requiring the model to match the median wealth across age and lifestyle for individuals born in the 1930s cohort. Figure 3 shows that the model is able to generate larger wealth accumulation for higher education and for protective health behavior. Even if the model is not required to target the overall distribution it fits the 25th percentile and the 75th percentile well.

5.2 Value of a Statistical Life (VSL)

The VSL comes from the estimated wage premium for a given probability of a fatal accident in risky jobs. This literature delivers numbers in the range of \$1 to \$7 million to save one life. We want the model to deliver this same marginal rate of substitution between income and survival probability. Using the value function expressed in Section 4.2.3 we can obtain the total differential:

$$\frac{\partial V_t^{c,e,y}(x,h,\ell,\xi)}{\partial x} dx + \frac{\partial V_t^{c,e,y}(x,h,\ell,\xi)}{\partial s_t^{e,y}(h)} ds_t^{e,y}(h) = 0$$

relating changes in cash-on-hand x and survival probabilities $s_t^{e,y}(h)$ that leave individuals indifferent. Rearranging we obtain,

$$-\frac{dx}{ds_t^{e,y}(h)} = \frac{\partial V_t^{c,e,y}(x,h,\ell,\xi,\zeta)}{\partial s_t^{e,y}(h)} \left[\frac{\partial V_t^{c,e,y}(x,h,\ell,\xi)}{\partial x} \right]^{-1}$$

Hence, for an individual of type (c, e, y) with state variables (x, h, ℓ, ξ) at age t to accept an increase in his survival probability in say 1%, he would require $0.01 \times dx/ds_t^{e,y}(h)$ units of income. Putting 100 identical agents together we would have one death on average in exchange of $dx/ds_t^{e,y}(h)$ units of income. Hence, this expression gives the model equivalent of the VSL. Because the empirical estimates of a VSL typically come from blue-collar jobs, we want the model to deliver a VSL for the average high school dropout of 35 years of age. We target a VSL of 2,000,000. This identifies the parameter $\bar{b}$.

5.3 Early life

In order to calibrate the parameters of the early in life model, we take the value function as given and we require the model to match the joint distribution of education and lifestyles for two different cohort: 1930 and 1970. This part of the model has 5 parameters to estimate: $\mu_e, \mu_{\text{PRO}}, \mu_{\text{CG}}, \sigma_e$, and σ_{PRO}. Given that we require the model to match the joint distribution of e, y for two different cohorts we have 10 moments. μ_e, μ_{PRO}, and μ_{CG} are identified by the average share individuals across education and life-style choices. σ_e and σ_{PRO} are identified as they drive the changes across education choices and lifestyles across cohorts.

Figure 4 and 5 show that the model is able to match well the marginal distributions of education and health behavior and their changes across cohorts. Moreover it is also able to match well the fact that lower educated individuals tend to invest less in their health. While the life-expectancy gradient for the 1970 it is perfectly matched, the model slightly overestimates the gradient for the 1930 cohort. Overall, the model is able to explain 50% of the increase in the life-expectancy gradient between college and high-school graduates between the 1930 and 1950 cohorts.

6 Results

The model incorporates different mechanisms that can explain why individuals with higher education invest more in their health. Firstly, higher expected income among the more educated encourages healthier behavior as life becomes more valuable. Secondly, as detailed in section 2, the benefits of protective behavior for life expectancy are greater for those with a college education. Lastly, considering the first two points, individuals facing lower costs of protective behavior (τ_{PRO}) are more likely to pursue higher education as the returns on health investments are higher for those who opt for protective measures.

To gauge the importance of each mechanism, we conduct a series of counterfactual experiments using the model. In these experiments, we keep individuals' education choices fixed and observe how their health investments would differ under various scenarios. In the first scenario, we simulate the behavior of high-school dropouts if they were to have the income prospects of college graduates. The increase in expected wages leads to higher values for both $V^{c,\text{HSD},\text{PRO}}$ and $V^{c,\text{HSD},\text{DET}}$, denoted

as $\tilde{V}^{c,\text{HSD,PRO}}$ and $\tilde{V}^{c,\text{HSD,DET}}$ respectively. Since higher income enhances the flow utility individuals in this counterfacual economy value more being alive. Thus, $\tilde{V}^{c,\text{HSD,PRO}} - \tilde{V}^{c,\text{HSD,DET}} > V^{c,\text{HSD,PRO}} - V^{c,\text{HSD,DET}}$, leading dropout individuals to be more inclined to choose protective behavior despite the associated costs.

The solid line in the upper panel of Figure 6 illustrates the distribution of protective behavior costs (τ_{PRO}) for high-school dropouts in the benchmark model. The vertical line represents $V^{c,\text{HSD,PRO}} - V^{c,\text{HSD,DET}}$. Individuals with $\tau_{\text{PRO}} < V^{c,\text{HSD,PRO}} - V^{c,\text{HSD,DET}}$ opt for protective behavior, while the rest choose detrimental behavior. The integral of the distribution of τ_{PRO} between zero and $V^{c,\text{HSD,PRO}} - V^{c,\text{HSD,DET}}$ represents the fraction of dropouts adopting protective behavior.

As income rises, the threshold value that prompts individuals to adopt protective health behavior shifts to the right. The dashed vertical line in Figure 6 marks the value at $\tilde{V}^{c,\text{HSD,PRO}} - \tilde{V}^{c,\text{HSD,DET}}$. This indicates that income largely influences why high-school dropouts tend to adopt more detrimental behavior. If faced with the same expected income as college graduates, the proportion of high-school dropouts choosing detrimental behavior would decrease from 45.7% to 22.9%, reducing the life-expectancy gap from 6.9 years to 5.2 (a 25% reduction).

Another reason why college graduates invest more in health behavior in the model is because the gains in life expectancy due to protective behavior are larger. To analyze this effect, we solve for $V^{c,\text{HSD},y}$ assuming high-school dropouts experience the same health transitions as college graduates. As in the previous counterfactual, the improved health transitions result in higher values for both $V^{c,\text{HSD,PRO}}$ and $V^{c,\text{HSD,DET}}$, denoted as $\hat{V}^{c,\text{HSD,PRO}}$ and $\hat{V}^{c,\text{HSD,DET}}$ respectively. Given that in this counterfactual, the gains in life expectancy are now larger for protective behavior, $\hat{V}^{c,\text{HSD,PRO}} - \hat{V}^{c,\text{HSD,DET}} > V^{c,\text{HSD,PRO}} - V^{c,\text{HSD,DET}}$, leading dropout individuals to be more inclined to choose protective behavior despite the associated costs.

The vertical dotted line in the middle panel of Figure 6 represents the value at $\hat{V}^{c,\text{HSD,PRO}} - \hat{V}^{c,\text{HSD,DET}}$ in this counterfactual scenario. This illustrates that higher returns on health behavior for college-educated individuals are as influential as the income effect.

Lastly, given the complementary nature of health and education investments, individuals facing lower costs of protective behavior (τ_{PRO}) are more likely to pursue higher education. This leads to a sorting effect where the distribution of protective behavior costs for high-school dropouts first-order stochastically dominates the distribution for college graduates. To quantify this effect, we examine the fraction of detrimental types if high-school dropouts had the same cost distribution as college graduates.

The lower panel in Figure 6 displays the distribution of the cost of protective behavior (represented by the dashed line). It shows that the mass of individuals below the $V^{c,\text{HSD,PRO}} - V^{c,\text{HSD,DET}}$ threshold (vertical line) is larger. If faced with the distribution of costs τ_{PRO} of college-educated individuals, the proportion of detrimental behavior would decrease from 45.7% to 40.7%. Therefore, selection would account for approximately 5.4% of the gradient in life expectancy.

6.1 Changes over time

The econometric model presented in section 2 showed an increase in the life-expectancy gap between college graduates and high-school dropouts of 1.9 years across the 1930 and 1970 cohorts. This increase is driven by lower (higher) health investments among the high-school dropouts (college graduates).

The model successfully reproduces 50% of this observed increase. To further elucidate the mechanisms driving the expansion of the life-expectancy gap in the model, we delve into identifying these factors. As explained in the previous section, changes in individuals' income prospects influence their educational and health investment decisions. Between the 1930 and 1970 cohorts, there has been a significant rise in the education premium, driven by increases in college wages and slight declines in dropout wages.

Given the complementarity between health and education behaviors in the model, income increases (decreases) for higher (lower) educated individuals lead to greater (lesser) health investments. Moreover, changes in the wage distribution across educational choices also impact the distribution of individuals in terms of the cost of protective behavior across education alternatives. Increases in the wage premium incentivize individuals to pursue educational investments, particularly among those facing lower costs of protective behavior.

To quantify the income effect while controlling for selection, we fix the distribution of health behavior costs faced by individuals in different education categories to the one in 1930. We then analyze how their health investments would have changed solely due to changes in income. In the upper panel of Figure 7, the solid line represents the distribution of the cost of protective behavior faced by individuals in 1930. The vertical solid line indicates the threshold value at which individuals of that cohort decided to switch from a protective to a detrimental lifestyle. The dashed vertical line illustrates that for high-school dropouts, the threshold value decreased across cohorts due to declines in income and consequently, a lower willingness to prioritize health. Declines in income account for 78% of the reduction in life expectancy for high-school dropouts between individuals born in the 1930 and 1970.

The dashed line in the upper panel of Figure 7 illustrates the distribution of behavior costs for high-school dropouts in the 1970s. Driven by changes in income across cohorts and education categories, compared to the distribution in 1930, the cost distribution worsened for high school dropouts born in the 1970s. Nevertheless, the effect of selection is modest. We find that if high-school dropouts in 1930 had the same cost distribution than high-school dropouts in 1970, the life-expectancy of the high-school dropout would only have felt 22% of what the original model predicts.

Finally the lower panel of Figure 7 plots the equivalent decompositions for the college graduates. It shows that all increases in life-expectancy for college graduates across cohorts is driven by the better expected income. Selection, on the other hand plays a quantitatively negligible role.

7 Conclusions

In this paper, we propose a latent variable model to characterize how health behavior influences health dynamics across different education groups. Our findings indicate that health behavior can be parsimoniously summarized into two lifestyles: protective and detrimental. We observe that individuals with higher levels of education tend to more frequently choose protective behavior, and differences in lifestyles explain 40% of the variations in life expectancy across education groups. Additionally, conditional on behavior, we find that a protective lifestyle has a greater impact on extending life expectancy for college graduates than for dropouts. Finally, we identify an increasing life-expectancy gradient across education groups between the 1930s and 1970s, driven by worsening lifestyles among the less educated and improved lifestyles among the more educated.

Furthermore, we introduce a heterogeneous agents model comprising two distinct stages. Initially, individuals make a one-time decision regarding education and lifestyle during an early-life health stage. Subsequently, in a working/retirement phase, agents address a consumption savings problem subject to income and health risks, as modeled in the econometric framework.

This model enables us to explain the connection between income and health inequality. Health and education decisions are shown to be complementary due to two key factors. Firstly, higher income increases the value of life, leading to greater returns from investing in health. Secondly, as reflected in our calibration process where we integrate the health dynamics from the econometric model, we observe greater returns to health investment for college-educated individuals. Driven by these complementarities, individuals facing higher costs of maintaining protective health behaviors are more likely to select lower education categories.

We calibrate the model to match savings, education, and lifestyle choices across cohorts, and then we use it to understand why lower-educated individuals tend to choose unhealthier lives. Our analysis reveals that lower income and diminished returns in health outcomes from protective behavior largely account for the disparities observed across education groups.

Finally, the model is able to explain 50% of the increase in health inequality across the 1930s and the 1970s. 80% of the deterioration in life expectancy among the less educated is driven by worsening economic conditions, and 20% is attributed to selection effects. All improvements in lifestyle among college graduates are explained by improvements in economic conditions.

The model cannot fully account for the increase in the life-expectancy gradient observed in the data. Factors such as peer influence, segregation, genetic predispositions, and variations in intergenerational mobility across cohorts are likely significant drivers of health behavior choices made by individuals, which we abstract from in our current analysis. These avenues offer promising directions for future research.

References

Case, A. and A. Deaton (2015): "Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century," *Proceedings of the National Academy of Sciences*, 112, 15078–15083. (Cited on page 1.)

Chetty, R., M. Stepner, S. Abraham, S. Lin, B. Scuderi, N. Turner, A. Bergeron, and D. Cutler (2016): "The association between income and life expectancy in the United States, 2001-2014," *JAMA*, 315, 1750–1766. (Cited on page 1.)

Cole, H. L., S. Kim, and D. Krueger (2019): "Analysing the effects of insuring health risks: On the trade-off between short-run insurance benefits versus long-run incentive costs," *The Review of Economic Studies*, 86, 1123–1169. (Cited on page 3.)

De Nardi, M., S. Pashchenko, and P. Porapakkarm (2017): "The lifetime costs of bad health," Tech. rep., National Bureau of Economic Research. (Cited on page 3.)

Fonseca, R., P.-C. Michaud, T. Galama, and A. Kapteyn (2021): "Accounting for the rise of health spending and longevity," *Journal of the European Economic Association*, 19, 536–579. (Cited on page 3.)

Hosseini, R., K. A. Kopecky, and K. Zhao (2021): "How important is health inequality for lifetime earnings inequality?" . (Cited on page 3.)

Kitagawa, E. M. and P. M. Hauser (1973): *Differential mortality in the United States: A study in socioeconomic epidemiology*, Harvard University Press. (Cited on page 1.)

Lantz, P. M., J. S. House, J. M. Lepkowski, D. R. Williams, R. P. Mero, and J. Chen (1998): "Socioeconomic factors, health behaviors, and mortality: results from a nationally representative prospective study of US adults," *JAMA*, 279, 1703–1708. (Cited on page 1.)

Li, Y., A. Pan, D. D. Wang, X. Liu, K. Dhana, O. H. Franco, S. Kaptoge, E. Di Angelantonio, M. Stampfer, W. C. Willett, et al. (2018): "Impact of healthy lifestyle factors on life expectancies in the US population," *Circulation*, 138, 345–355. (Cited on page 1.)

Mahler, L. and M. Yum (2022): "Lifestyle behaviors and wealth-health gaps in Germany," *Available at SSRN 4034661*. (Cited on page 3.)

Margaris, P. and J. Wallenius (2023): "Can Wealth Buy Health? A Model of Pecuniary and Non-Pecuniary Investments in Health," *Journal of the European Economic Association*, jvad044. (Cited on page 3.)

Meara, E. R., S. Richards, and D. M. Cutler (2008): "The gap gets bigger: changes in mortality and life expectancy, by education, 1981–2000," *Health affairs*, 27, 350–360. (Cited on page 1.)

Pijoan-Mas, J. and J.-V. Ríos-Rull (2014): "Heterogeneity in expected longevities," *Demography*, 51, 2075–2102. (Cited on pages 1 and 5.)

Polvinen, A., R. Gould, E. Lahelma, and P. Martikainen (2013): "Socioeconomic differences in disability retirement in Finland: the contribution of ill-health, health behaviours and working conditions," *Scandinavian journal of public health*, 41, 470–478. (Cited on page 1.)

Preston, S. H. and I. T. Elo (1995): "Are educational differentials in adult mortality increasing in the United States?" *Journal of aging and health*, 7, 476–496. (Cited on page 1.)

Xiang, S., T. Jia, C. Xie, W. Cheng, B. Chaarani, T. Banaschewski, G. J. Barker, A. L. W. Bokde, C. Büchel, S. Desrivières, H. Flor, A. Grigis, P. A. Gowland, R. Brühl, J.-L. Martinot, M.-L. P. Martinot, F. Nees, D. P. Orfanos, L. Poustka, S. Hohmann, J. H. Fröhner, M. N. Smolka, N. Vaidya, H. Walter, R. Whelan, H. Garavan, G. Schumann, B. J. Sahakian, T. W. Robbins, and J. Feng (2023): "Association between vmPFC Gray Matter Volume and Smoking Initiation in Adolescents," *Nature Communications*, 14, 4684. (Cited on page 2.)

Zaninotto, P., J. Head, and A. Steptoe (2020): "Behavioural risk factors and healthy life expectancy: evidence from two longitudinal studies of ageing in England and the US," *Scientific Reports*, 10, 6955. (Cited on page 1.)

List of Figures

FIGURE 1: Probability of having a health habit by health behavior type as individuals age

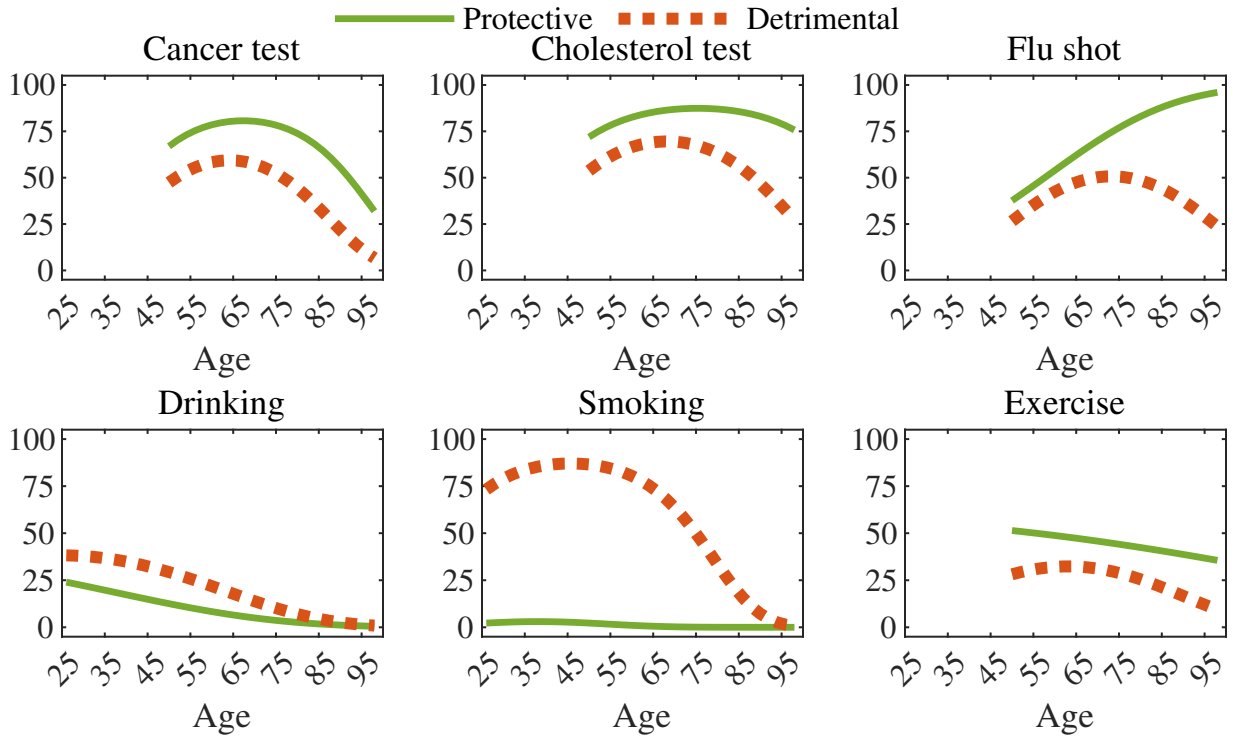


FIGURE 2: Distribution of types by education and cohort (males)

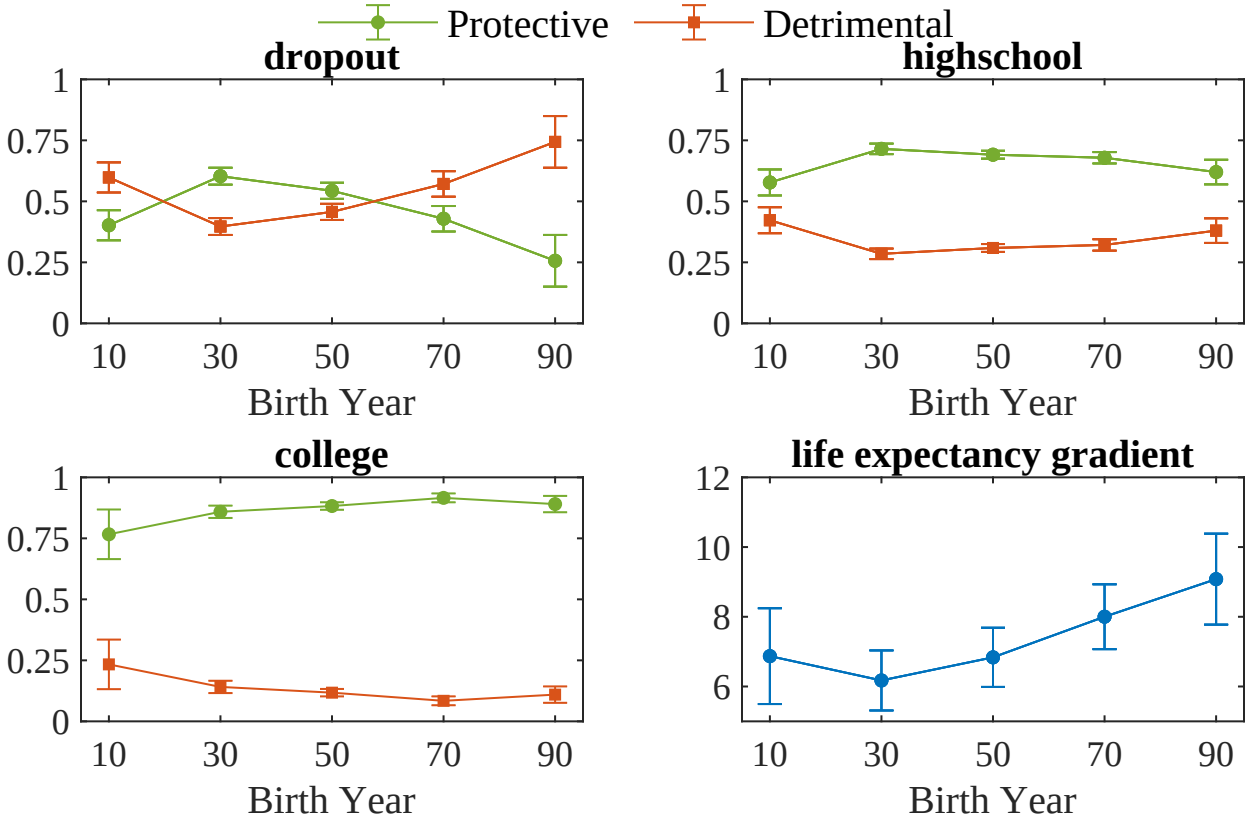


FIGURE 3: Model fit: wealth distribution model (lines) versus data (scatter)

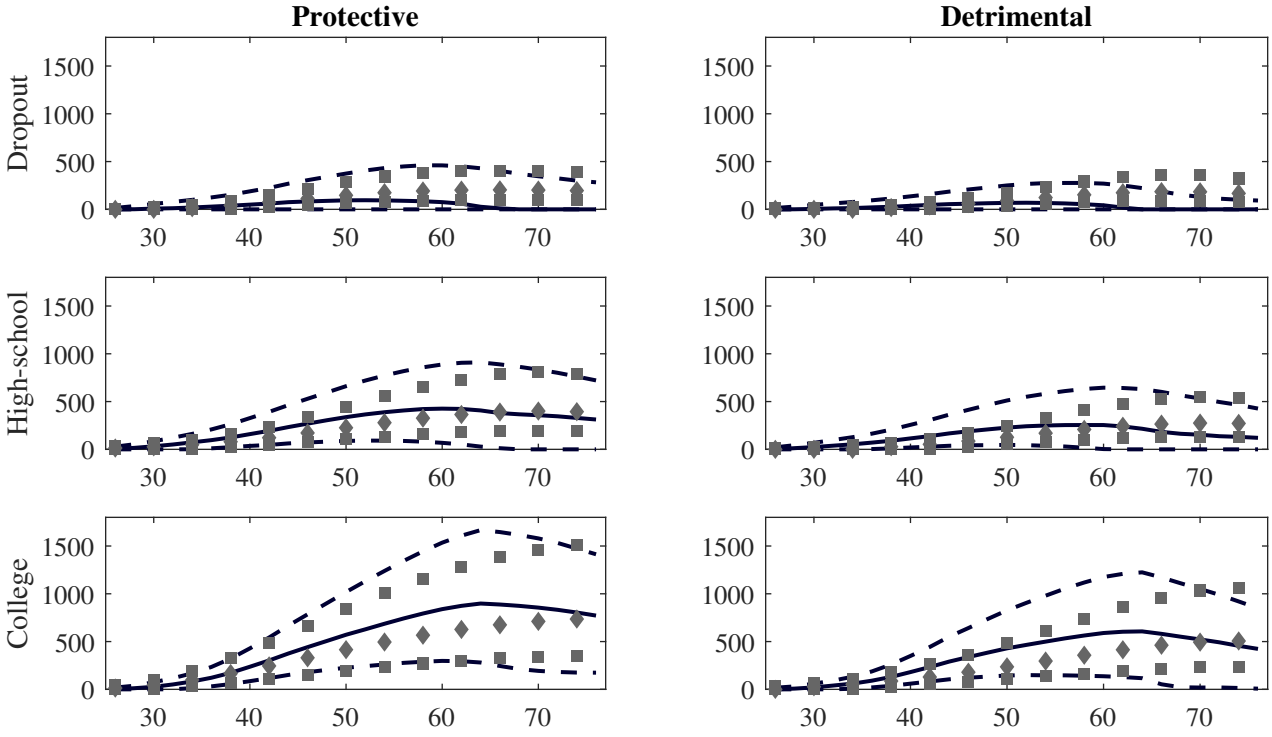


FIGURE 4: Marginal distributions: Education and Health Behavior

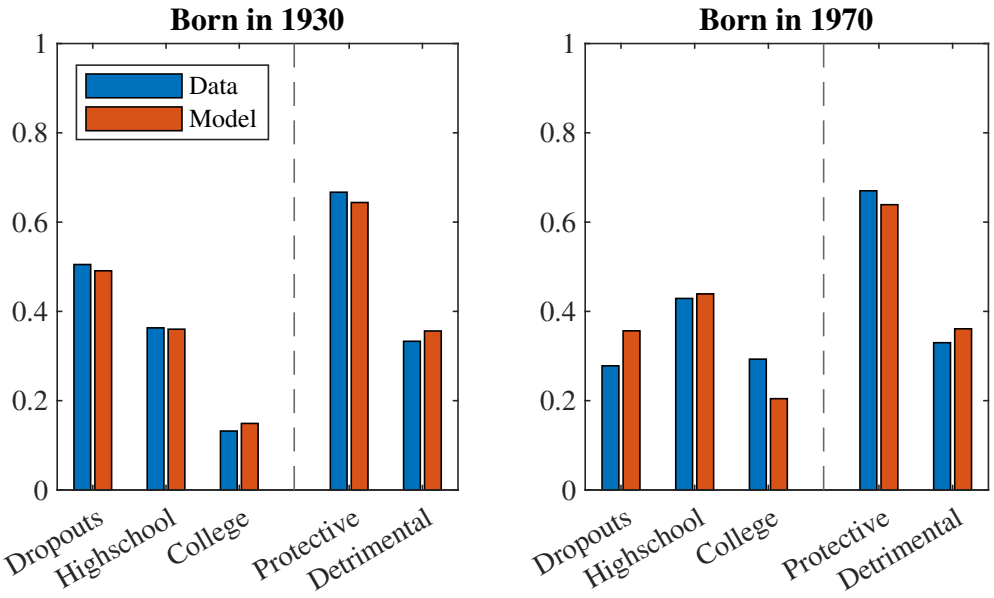


FIGURE 5: Conditional distribution of Detrimental Behavior by Education

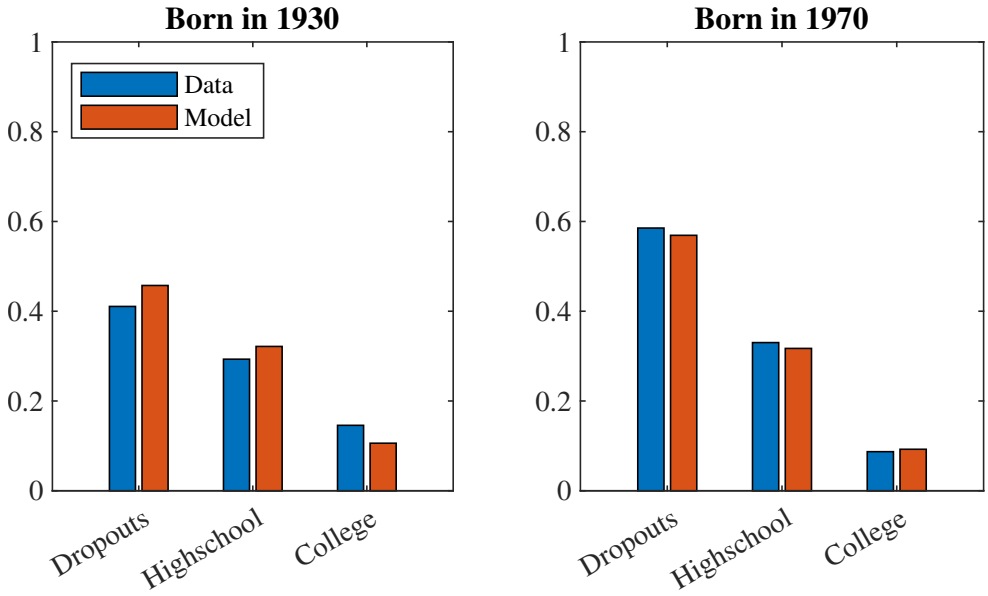


FIGURE 6: Cost of protective lifestyle

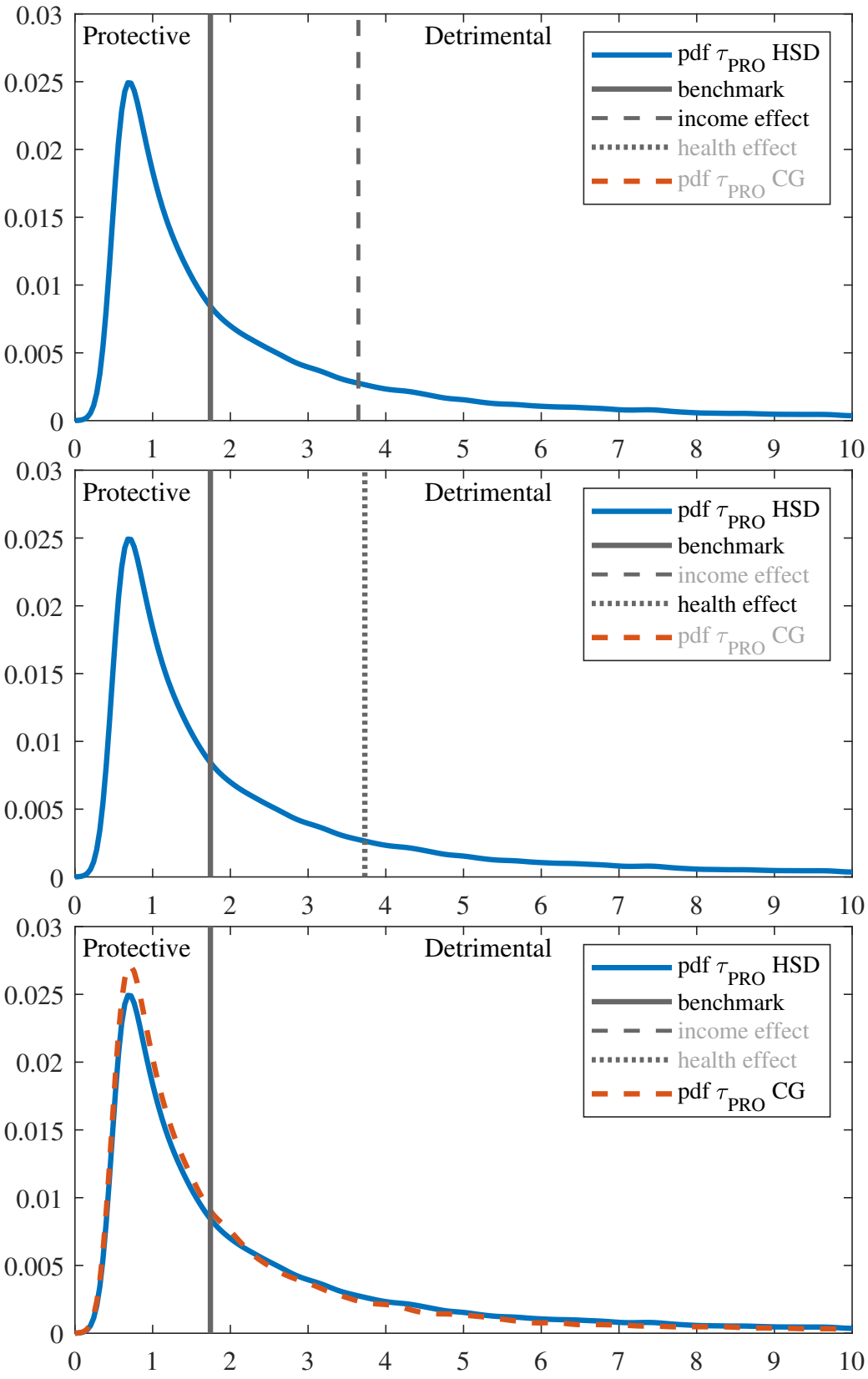
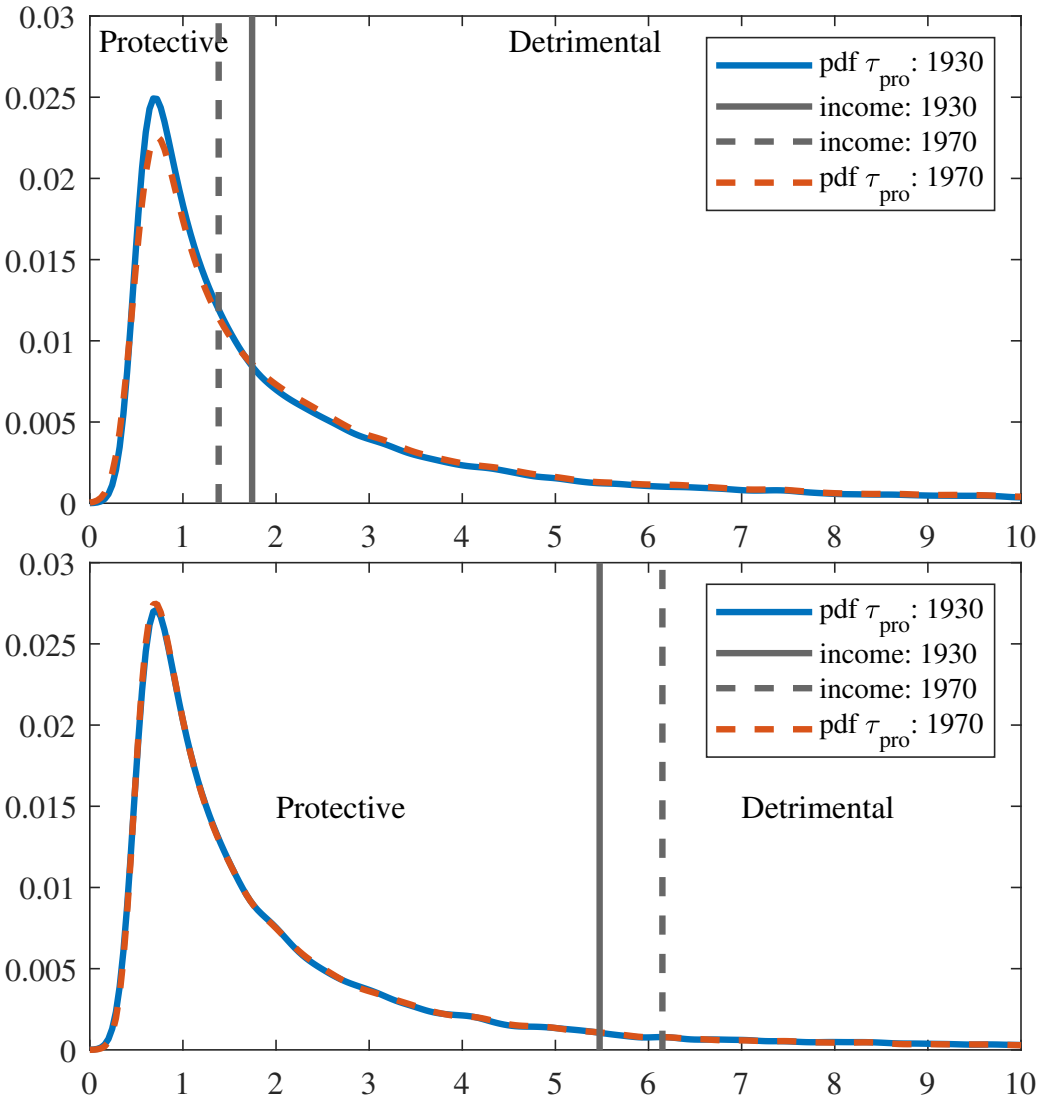


FIGURE 7: Cost of protective lifestyle across cohorts



List of Tables

TABLE 1: Expected duration of each health state at age 50 across behavior types and sex

Health Behavior	Fraction behavior	Life Expectancy	=	Good Health	+	Bad Health
Men: Dropouts						
Protective	42.9	29.0		17.5		11.4
Detrimental	57.1	21.4		11.8		9.6
Δ	-14.3	7.6		5.7		1.8
Men: High-school						
Protective	67.9	30.3		23.9		6.4
Detrimental	32.1	23.0		16.2		6.8
Δ	35.7	7.2		7.7		-0.4
Men: College						
Protective	91.6	33.4		29.6		3.8
Detrimental	8.4	24.5		19.1		5.4
Δ	83.2	8.9		10.5		-1.6
Women: Dropouts						
Protective	55.4	30.3		15.1		15.2
Detrimental	44.6	23.2		9.9		13.3
Δ	10.9	7.1		5.1		2.0
Women: High-school						
Protective	71.8	33.2		25.8		7.4
Detrimental	28.2	26.1		17.6		8.4
Δ	43.6	7.2		8.2		-1.0
Women: College						
Protective	92.7	34.9		30.4		4.5
Detrimental	7.3	28.2		20.8		7.4
Δ	85.4	6.7		9.6		-2.8

Notes: The third column sums the expected duration in all possible health states which is equal to the life expectancy at age 50.

TABLE 2: Internally calibrated parameters

Parameter	Description	Value
$\underline{x}$	income floor	17.60
b_0	bequest motive: marginal utility	3.90
b_1	bequest motive: non-homoteticity	103.71
b	value of life	-0.63

TABLE 3: Early-life parameters

Parameter	Value	Parameter	Value
μ_e	6.58	σ_e	4.02
μ_{PRO}	8.52	σ_{PRO}	1.45
μ_{CG}	5.26		