Pathways to Disability and Death in Middle Ages: Estimates from the Health and Retirement Study Data*

Lakshmi K. Raut†

2024-09-26

Abstract

I review the microbiology literature on aging, finding that much of the process of health development over the lifespan is determined by early childhood epigenetic factors consisting of internal and external environments, health care use, and health related behavior, together with genetic factors. I introduce a statistical 'disablement model', specifically, a multi-state time-to-event life history model of health developments over the lifespan, incorporating childhood factors and health behaviors as regressors. I correct for unobserved heterogeneity biases of parameter estimates, extending a statistical technique from the literature of two-state models. Using the HRS (Health and Retirement Studies) dataset and taking the definition of disability as health states that qualify for the Social Security's DI (Disability Insurance) program or the SSI (Supplemental Security Income) program, I study how childhood factors and health behaviors are associated with probability of following various transition paths through the health states of normal health, chronic diseases, disability, and death. The childhood factors include childhood socioeconomic status, health status and education level, and the health behaviors include smoking and exercising. The regressors also include biomarkers—body mass index (BMI), a measure of stress (CES-D) and a measure of cognitive health. I also carry out quantitative analysis of the effect of social policies improving childhood factors and health behaviors of various social groups on their risks of chronic illness, disability, or death in mid-ages.

*This paper is dedicated in loving memory of my younger brother, Bishnu Pada Raut, who passed away in New Delhi on February 1, 2019, from lung cancer. He never smoked, never drank, and had normal BMI, CES-D and other standard biomarkers (personally observed) throughout his life. Why are there incidence of diseases and death at premature ages? Scientific community is actively exploring the answers to these questions and looking for the ways to improve life. This paper is an inquiry in this vein. An earlier draft was presented at the 2019 Annual Conference of the Society for Government Economists, April 5, 2019, Washington, DC. I got many useful comments from the discussant, Elizabeth Bass at Congressional Budget Office, and from the audience. I had many insightful comments from Han Altae-Tran at MIT, John Phillips at NIH, and Michael V. Leonesio, Javier Meseguer, David Pattison, Mark Sarney, Alexander Strand and Robert Weathers at SSA. Thanks.

Disclaimer: The views, thoughts, and opinions expressed in the paper belong solely to the author, and do not necessarily represent the views of the Social Security Administration or the United States Government, other group or individual.

Contents

List of Figures

List of Tables

1 Introduction

Identification of factors that determine disability and mortality incident rates is important for disability programs such as the Social Security Disability Insurance (DI) program and the Supplemental Security Income (SSI) program. According to the biology of living organisms, all individuals succumb to aging, and experience diseases and disabilities of various kinds as they

age. Diseases and disabilities can also be caused by injuries, genetic abnormalities and epigenetic reprogramming (epigenetic includes environmental factors and health-related individual behaviors). Some individuals stay in normal health for a long period of time and then become disabled or die; some develop early on one or more chronic diseases such as diabetes, cancer, vascular, musculoskeletal, cognitive, and mental disorders that expedite their incidence of disability and death. I use the Health and Retirement Study (HRS) data on individuals in their early 50's to estimate a dynamic multi-state time-to-event econometric model of pathways to disability or to death before disability through various health states — specifically, through normal health and one-or-more chronic diseases — before reaching age 65. Genetic and environmental factors, healthcare use, health related behaviors throughout life and cognitive factors determine the progression of unobserved state of internal health (also known as health-capital in economics, and frailty in gerontology). The state of internal health determines the risks of transitions to other health states and the transit times. I estimate the effect of these factors on the probabilities of transitions and the transit times along the pathways that individuals in their 50's follow before reaching age 65.

Before exploring pathways to disability, I must clarify the definition of disability that I use in this paper. The definition of disability depends on the purpose of its use. Disability is a multidimensional concept and is defined in the literature using simple descriptions, conceptual models, classification schemes, and measurement methods (for details, see Hahn (1985); Marks (1997);Altman (2001); Albrecht and Verbrugge (2003); Snyder et al. (2008)). I use the following statutory definition of disability that the Social Security Administration uses for the DI and SSI programs (specified in the Social Security Act, Title II, § 223(d), paragraph (1)A):

> "inability to engage in any substantial gainful activity by reason of any medically determinable physical or mental impairment which can be expected to result in death or which has lasted or can be expected to last for a continuous period of not less than 12 months"

and with a vocational grid addendum stated in paragraph (2)A,

> "An individual shall be determined to be under a disability only if his physical or mental impairment or impairments are of such severity that he is not only unable to do his previous work but cannot, considering his age, education, and work experience, engage in any other kind of substantial gainful work which exists in the national economy, regardless of whether such work exists in the immediate area

> in which he lives, or whether a specific job vacancy exists for him, or whether he would be hired if he applied for work. For purposes of the preceding sentence (with respect to any individual),"work which exists in the national economy" means work which exists in significant numbers either in the region where such individual lives or in several regions of the country."

The definition of disability for the SSI program is almost identical.[1]

While a lot has been said about the definition of disability, very few papers provide a biological or a behavioral mechanism of the disablement process from which policy implications for clinical practice and healthcare policy can be derived. The first disablement model was introduced by the sociologist Nagi (1965), which he further refined in (Nagi, 1976; Nagi, 1991). This model was extended by (Verbrugge and Jette, 1994; Verbrugge, Latham, et al., 2017) who added biological, environmental and behavioral risk factors affecting various stages of the disablement process.[2] These disablement models are conceptual schemes that describe four distinct but related stages to arrive at a disability: The first stage begins with a pathology[3], progressing to the second stage of impairments of body systems[4], to the third stage of functional limitations[5] and finally to the fourth stage of disability.[6]

The above disablement models are useful for conceptualization, diagnosis and record keeping of disabilities but limited for the study of the causes of disability in epidemiological and policy research. The starting point of the above disablement models is onset of an injury or a chronic disease causing a disability. An injury as the starting point of disablement process serves well

[1]For details, see https://www.ssa.gov/OP_Home/ssact/title02/0223.htm section 223(d)(2) for the OASDI program, and https://www.ssa.gov/OP_Home/ssact/title16b/1614.htm paragraph 3(A) for the SSI program.

[2]Nagi model has been adapted by the World Health Organization in their classification scheme of disability, the latest one is World Health Organization (2001). See Bedirhan et al. (2010) for an application of the above disablement models in WHO's 2001 classification system, and see (Pope and Tarlov, 1991; Institute of Medicine, 2007) for more on this.

[3]Pathology is an interruption of the normal physiological process caused by developmental disorders (such as cerebral palsy, seizure disorders, mental retardation, hearing and vision impairments, autism, PKU, Huntington disease), infection, injury, trauma, metabolic imbalance (such as diabetes), degenerative disease processes (i.e., deterioration over time the functioning or the structure of tissues or organs leading to osteoarthritis, osteoporosis, cancer, Alzheimer or Parkinson's disease) or any other disease processes.

[4]The impairments of body system involve loss or abnormality of an anatomical, physiological, mental, or emotional nature.

[5]Functional limitations include not being able to have one's ADL (activities of daily living) and IADL (instrumental activities of daily living), role activities (such as occupation, parenting, grand-parenting, and student roles), social activities (such as attending church and other group activities and socializing with friends and relatives) and leisure activities (such as sports and physical recreations, reading, and distinct trips).

[6]The definition of disability depends on the purpose of the study and involves a combination of all the above models of disability.

for certain purposes such as for workers' compensation in sports, construction and factories, and there are needs for public policies to lower such risks. A large proportion of disabilities in mid ages are, however, caused by diseases—both physical and mental (see for instance, Case and Deaton, 2015; The US Burden of Disease Collaborators, 2018). For research on public policies to improve such disability and mortality risks of individuals, it is important to study the biomedical processes modulated by genetic, epigenetic and behavioral factors in the manifestation and prognosis of disabling diseases. Diseases leading to disabilities (both developmental disabilities and late age disabilities) and to mortality are the result of modulated biomedical processes, which at the microbiology level are the outcomes of cellular aging. While aging, an individual succumbs to diseases leading to disability or death. Not all individuals experience the same deterministic aging process — some experience faster aging and aging related diseases than others.

Why do some people experience faster aging, incidence of diseases and mortality? At what stage of life, does it all begin — at middle age, at birth, or even earlier at conception? How do various genetic, epigenetic and behavioral factors modulate the aging process, culminating in diseases, disabilities and death? What biomarkers and epigenetic factors (including environmental factors and individual health related behaviors) over the lifespan predict better the process of aging and incidence of diseases, disability, and death? These are all important questions in the disability policy research.

At the cellular level, aging means cellular senescence — i.e., after a certain number of cell divisions, it stops dividing or have defective replications, causing tissues or organs to increasingly deteriorate over time. Senescence leads to incidence of degenerative diseases. The microbiology literature finds that aging of cells, i.e., cellular senescence, and age-related diseases are associated with shortening of telomere length[7] and changes in global methylation, and that stress, smoking, drinking, chemical misuse, and diets are important modulators of these changes. (For telomere mechanism, see for instance, Austad and Fischer, 2016; Blair et al., 1989; Vaupel, 2010; Zarulli et al., 2018, and for methylation mechanism, see for instance, Alisch et al., 2012; Barres and Zierath, 2011; Boks et al., 2009; Esteller, 2008; Hannum et al., 2013; Horvath, 2013)

What are the critical periods or the developmental milestones in lifecycle that program the motions of health developments over the lifespan of an individual? Research along this line began with the striking findings of Barker (Barker, 1990; Barker, 1998) and later of Gluckman et al. (2008). They found strong association between birth weight and many later life chronic dis-

[7]Telomeres are the caps at the end of chromosomes in a DNA sequence. They look like the plastic caps at the end of shoelaces.

eases, including hypertension, coronary artery diseases, type 2 diabetes, and osteoporosis. Many other studies find that much of health developments in later life is determined very early in life — specifically during the prenatal period, right after conception, i.e., in the womb. The effect of an environmental stress in the womb on the occurrence of later life diseases and developmental outcomes is known as *developmental programming*.

The finding of the above microbiology literature — that cellular stressor at various stages of the cell's lifecycle, especially during the prenatal and postnatal development period are important determinants of the speed of cellular aging — is an important milestone in aging research. The cellular stressors are difficult to measure and monitor for a population. For public policy, it is imperative to find socioeconomic factors and health related behavioral factors that modulate the cellular stressors. Many studies in social sciences find that low socioeconomic status (SES) are associated with inflammation, metabolic dysregulation, and various chronic and age-related diseases such as type 2 diabetes, coronary heart disease, stroke, and dementia, and that low SES create epigenetic changes in individuals that lead to faster biological aging even after controlling for health-related behaviors such as diet, exercise, smoking, alcohol consumption, or access to quality healthcare, see Simons et al. (2016) among others for discussions on such epigenetic mechanisms.

The empirical studies along this line in social sciences are emerging. The study by Karakus and Patton (2011) using the HRS dataset found that depression as a measure of stress at a baseline leads to significantly higher risk for developing diabetes, heart problems, and arthritis but no significant effect on developing cancer during the 12 years follow-up period, after controlling for education, race, income, health risk indicators like BMI and smoking, functional limitations like gross motor index, and health limitations for work and income. The study by Renna (2008) using the National Longitudinal Survey of Youth dataset found no significant effect of alcohol use on labor market outcomes like earnings or hours of work. Seib et al. (2014) collected data on a sample of older women in Australia and their study found that severe traumatic life events create strong stress levels that influence them to have unhealthy living and diets measured by BMI and develop stronger and earlier health problems. Conti et al. (2009) utilized the CES-D data in the HRS dataset to construct a measure of depression, and found that depression of men and women has significant negative effect on the employment status, early retirement, and application for DI/SSI benefits. Luo and Waite (2005) using the HRS dataset found that childhood SES and childhood health strongly influence later life health outcomes. Case and Deaton (2015) found a racial reversal in the mortality rates of non-Hispanic white men and women of ages

45-55 between 1993 and 2013, attributing it to growing distress of this population that led to increases in their mortality and morbidity from drugs and alcohol use and suicides.

As pointed out above, the starting point of the existing disablement models is a pathology, i.e., incidence chronic diseases. I extend the existing disablement models by incorporating indicators of policy relevant factors – especially the early childhood factors – that the microbiology literature advocates as determinants of a pathology, disability, and mortality. I postulate that as individuals age or misuse drugs, alcohols or eat less nutrient diets, the homeostatic regulatory mechanism that controls physiological systems (such as respiratory, cardiovascular, neuroendocrine, immune, and metabolic systems) becomes more and more fragile in its ability to face internal and external stressors, leading to early occurrence of diseases, disability or death. More specifically, I use a multi-state time to event statistical framework to estimate the effects of these factors on the probability of following various pathways through the health states of normal health, chronic diseases, disability, and death. The multi-state framework is suitable to study the effect of various covariates that are specific to each intermediate health state on the risk of becoming disabled or dying.

The observed variables in a dataset imperfectly measure the genetic and epigenetic factors that affect the health development process over the lifespan. This leads to unobserved heterogeneity. Ignoring unobserved heterogeneity would cause biased parameter estimates of the included regressors. Much of the literature on the consequence of ignoring unobserved heterogeneity and statistical procedures to correct for unobserved heterogeneity biases is done for the two-state models. Little is known for the multi-state models. In this paper, I extend a technique for the two-state models to the multi-state model to correct for biases in parameter estimates.

The rest of the paper is organized as follows. In Section 2, I provide an extended disablement model. In Section 3, I describe the econometric specifications and estimation methods. In Section 4, I describe the dataset and the creation variables. In Section 6, I present the estimates of the transition probabilities. In Section 6, I estimate the effect of childhood factors on initial middle age health, and the effect of childhood factors, biomarkers, and health behaviors on individual health trajectories. In Section 7, I discuss the policy implications of the estimates. Section 8 concludes the paper.

2 The Model

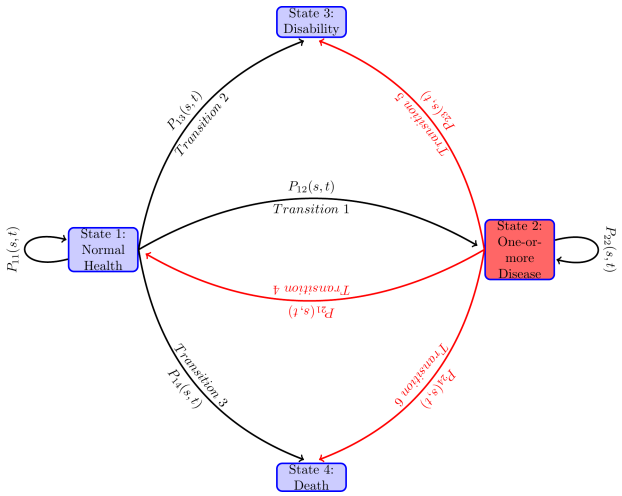
With insights from the disablement modelling literature and the microbiology literature on aging process, I formulate and then estimate an econometric model of pathways to disability enrollment and death before disability enrollment. An individual can be on the disability rolls if the individual has a qualifying disability and has not reached age 65. I assume that an individual's getting on the disability rolls is a terminal event, i.e., the individual does not move to the normal or the diseased health state from the disability health state.[8] After reaching this state, the individual is not followed any further. A competing risk for getting on the rolls is the event of death before age 65. This is a competing risk because an individual cannot be at risk for disability enrollment if the individual is already dead and thus not at risk to get on the disability rolls. In technical terms defined below, I treat health states — disability and death — as absorbing states, i.e., once in that health state, an individual remains in that health state and is removed from the sample for later considerations.

An individual can follow many possible health paths. For instance, beginning with normal health state, an individual can become disabled or die before becoming disabled after some passage of time. Or the individual may first become diseased with one or more diseases, or at the starting point of the study, the individual was in this health state and after some passage of time became disabled or died before becoming disabled. There are many possible sequences of health states that individuals may follow. Even when individuals pass through the same sequence of health states, the duration of stay in each health state (also known as the *waiting time* in stochastic process literature) could vary. A *pathway to disability or death before disability* of an individual consists of the sequence of health states and the waiting times in each health state an individual follows. From the diagram below one can see various health paths that an individual may traverse. The focus of the paper is to estimate the probabilities of various transitions and the duration of stay in each health state.

When time is continuous, the number of paths that an individual may follow is infinite. For an individual, one path maybe more likely than another. The likelihood or risk of following a particular path may depend on the individual's genetic make-up, prior health conditions and health related behaviors. Various factors affect the probability of transition from one health state to another at each point of time. These transition probabilities over time determine the time one

[8]The focus is on the first-time entitlement onto a disability program. A few people, however, recover and move to normal or diseased health state, but more likely they come back later to the disability rolls, see Raut (2017) for estimates of these probabilities.

Figure 1: Path diagram of health trajectories.



takes to get onto the disability rolls from normal or diseased health state.

The goal is to estimate the probability of getting onto the disability rolls and the competing risk of death before disability, sequentially over time as one progresses over life experiencing health shocks and changes in health behaviors. For instance, it is important to get an estimate of the probability of an individual in normal health at age 51 getting onto the disability rolls or die before disability by age 55, 60, or any other age before 65? How do these probabilities change if the individual from normal health at age 51 becomes diseased with one or more chronic diseases say at age 54? How do these probabilities compare for two individuals at age 51, one in normal health state and the other in diseased health state? How do these probabilities depend on an individual's childhood factors such as socioeconomic status, health and education and how do they depend on race and sex? Getting such estimates is important from policy point of view as these probabilities can elucidate the kind of social policies that can reduce disability enrollments or probability of death before disability, and to get quantitative estimates of the effects of such public policies. To that end, one needs to formulate an appropriate statistical model of the pathways through various health states, incorporating the effects of various time-varying covariates and estimate it using a nationally representative sample.

I model pathways through health states as a continuous-time finite-state stochastic process $Z(t), t \in T$, where at each time point t during the study period T, the random variable $Z(t)$ takes a value from a finite set S of health states. I take $T = [51, 65]$. The state space S of the stochastic process contains health states 1 = "healthy or normal health", 2 = "diseased with one or more chronic diseases", 3 = "disabled with DI or SSI qualifying disability" and 4 = "death

before disability". Sometimes I will denote the state space as $S = \{h, i, d, D\}$ in place of $\{1, 2, 3, 4\}$. Denote by $\mathscr{H}(t)$ the history of a set of covariates $X(t)$ including the paths of health states that the individual has gone through and other characteristics such as changes in health-related activities and so on. That is $\mathscr{H}(t) = \{Z(u), X(u)\}_{u=0}^{t}$.

The transition probabilities of our stochastic process $Z(t)$ given the history process $\mathscr{H}(t)$ are defined by,

$$P_{hj}(s, t; \mathscr{H}(s)) = Prob(Z(t) = j \mid Z(s) = h, \mathscr{H}(s)), \tag{1}$$

for all $h, j \in S$, $s, t \in T$, $t \geq s$. The transition probabilities contain important information about policy questions. It is known that the set of transition probabilities and the distribution of initial states completely determine a stochastic process. The statistical parameterization and inference of a stochastic process is generally performed with *transition intensity functions*.

Let t^- denote the time just before t. A *transition intensity*[9] $\lambda_{hj}(t; \mathscr{H}(t^-))$ of a stochastic process $Z(t)$ is defined as the rate of change in the probability of $Z(t)$ occupying state j at time t given that $Z(t^-)$ is in state h at time t^-. More formally,

$$\lambda_{hj}(t; \mathscr{H}(t^-)) = \lim_{\Delta t \downarrow 0} \frac{P_{hj}(t^-, t^- + \Delta t; \mathscr{H}(t^-))}{\Delta t}, \text{ for all } j \in S. \tag{2}$$

If the process occupies one of the states in S at all times, one gets $\lambda_{hh}(t; \mathscr{H}(t^-)) = -\sum_{j \neq h} \lambda_{hj}(t; \mathscr{H}(t^-))$, for $h = 1, 2, 3, 4$. Denote the matrix of transition intensities by $\Gamma(t; \mathscr{H}(t^-) = \left(\lambda_{ij}(t; \mathscr{H}(t^-)\right)_{i,j=1,2,3,4}$ and the matrix of transition probabilities by $P(s, t; \mathscr{H}(s)) \equiv \left(P_{hj}(s, t; \mathscr{H}(s))\right)_{h,j=1\ldots4}$. For ease of presentation, I suppress the conditioning variable $\mathscr{H}(t)$.

An individual at time t may be in any of the health states in S, the probability of which is known as the *occupation probability*. The occupation probability of a health state j at time t can be viewed as the proportion of population of age t who have that heath state. Denote this occupation probability by $\pi_j(t)$ and all the occupation probabilities together as a column vector $\pi(t) \equiv (\pi_j(t), j \in S)$. The occupation probabilities move over time recursively as follows,

[9] It is also known as the *hazard rate* in the survival analysis literature when exits occur because of only one event, and as the *cause-specific hazard rate* in the competing risk analysis when exits occur because of many events. For instance, see Raut (2017) for a competing risk analysis in a similar context using the SSA Administrative data and compare that with the present framework.

$$\pi(t) = \pi'(s)P(s,t), 0 \leq s < t.$$

Note that given an initial distribution $\pi(0)$ and the transition probabilities $P(s,t),\ 0 \leq s < t, s, t \in T$, one can calculate the occupation probabilities for all time periods in T from the above recursive equation.

Much of statistical estimation methods and their asymptomatic properties are developed for Markov process and I will assume it in the rest of the paper unless otherwise mentioned. A process $Z(t)$ is Markov if it depends on history only at its starting time s, i.e., on $Z(s)$, and $X(s)$ in the case of transition probabilities in Eq. (1) and on $Z(t)$, and $X(t)$ in the case of transition intensities in Eq. (2).

Given that the stochastic health evolution process is parameterized for statistical inference in terms of transition intensities, it is important to know how to get the transition probabilities from the transition intensities? This is accomplished as follows: For a transition $h \to j$, define the *integrated transition intensity* $\Lambda_{hj}(t)$ by $\Lambda_{hj}(t) \equiv \int_0^t \lambda_{hj}(u)du$ and the matrix $\Lambda(t) = \left(\Lambda_{hj}(t)\right)_{h,j=1,2,3,4}$. Let a time interval $[s,t]$ be subdivided into a partition of m sub-intervals with cut-off points $s = t_0 < t_1 < ... < t_m = t$. It can be shown that

$$P(s,t) = \prod_{s \leq u \leq t} (I + d\Lambda(u)) \overset{def}{\equiv} \lim_{\max t_i - t_{i-1} \to 0, m \to \infty} \prod (I + \Lambda(t_i) - \Lambda(t_{i-1})). \tag{3}$$

The entity $\prod_{s \leq u \leq t}(I + d\Lambda(u))$ defined above is known as the *product integral of the matrix function* $\Lambda(u)$.

3 Econometric specifications and estimation strategy

As mentioned in the introduction, the progression of health such as incidence of diseases, disabilities and mortality over life-course is determined by numerous factors. These factors include an individual's genetic make-up, epigenetic configuration modulated by health behaviors and the surrounding environments, and their evolution over time, and the randomness generated by somatic mutations. Since this study focuses on disabilities that individuals apply for and then get entitlement onto a disability program, the incidence of such disabilities also depends on many economic factors that influence these decisions, such influencing factors as the individual's disability insured status, replacement rate and unemployment rate. The included covariates

of a Cox regression model do not include many of those factors. Conditional on the included covariates, there remains individual variations in the risk of diseases, disability and mortality arising from the excluded factors. The variation in risks from excluded factors are known in the economics literature as the *unobserved heterogeneity* and in the epidemiology and statistical literature as *frailty*.

The transition intensity or hazard rate $\lambda_{hj}(t; X(t))$ of an event of transiting from health state h to health state j at time t given the history $X(t)$ is parameterized as follows,

$$\lambda_{hj}\left(t; X(t), V_{hj}\right) = \lambda^0_{hj}(t) V_{hj} e^{X(t)\beta_{hj}} = \lambda^0_{hj}(t) e^{X(t)\beta_{hj} + B_{hj}}. \tag{4}$$

where $V_{hj} \equiv e^{B_{hj}}$. In the above specification, $\lambda^0_{hj}(t)$ is known as the *baseline hazard function*. The specification of transition intensity in Eq. (4) is known as the *proportional hazard model*. It aggregates effects of the regressors linearly $X(t)\beta_{hj}$, known as the *fixed effect* and shifts the baseline hazard function proportionately over time, by the factor $e^{X(t)\beta_{hj}}$. A regression coefficient β of a covariate X in the above specification quantifies the *individual treatment effect* in terms of the change in the risk (measured in log-scale) of encountering transition $h \to j$ at time t from a unit increase in the treatment variable X. More specifically, let h be the healthy state and j be the disability health state and the treatment level $X = x_0$. The quantity $exp(\beta)$, known as the individual level *hazard ratio*, is the ratio of the probability of an individual from healthy state becoming disable at time t if the treatment is increased by a unit $X = x_0 + 1$ over the individual's probability of becoming disabled from healthy state at time t with the treatment level $x = x_0$. Proportionality assumption implies that the individual hazard ratios are time constant. The variable V_{hj} is the aggregate effect of all the unobserved covariates and it is assumed to be a random variable and hence known as the *random effect*. The interpretation of the frailty or unobserved heterogeneity variable V_{hj} is that it imparts a random proportional shift of the baseline intensity by a multiplier of magnitude v_{hj}, the realized value of the random variable V_{hj} for an individual. Individuals with higher realized values of v_{hj} are frailer and will have higher probabilities of transition at any given age. For identification, it is assumed that $E(V_{hj}) = 1$. An individual with $V_{hj} = 1$ will be referred as the *average individual*.

The most widely used statistical procedure for estimation of the transition probabilities $P(s,t)$, $s,t \in T, s < t$ with or without covariates is to plug-in an estimate of $\Lambda(t; X(t))$ in Eq. (3).

There are broadly two types of statistical methods to get an estimate of $\Lambda(t; X(t))$ — parametric method and semi-parametric method. In the next Sub-Section 3.1, I will follow the widely used

Aalen-Johnson-Fleming semi-parametric method. This method uses the Nelson-Aalen estimate of each component transition intensity function of $\Lambda(t; X(t))$ assuming a Cox proportional hazard model to incorporate the time-varying covariate effects. Note that without covariates, this semi-parametric method is, in fact, a nonparametric method.

3.1 Aalen-Johansen-Fleming Estimator for Transition Probabilities

The bulk of developments of statistical methods for inference in the multi-state framework assumes the absence of unobserved heterogeneity or frailty. Two papers (Fleming, 1978; and Aalen and Johansen, 1978) independently extended the Kaplan-Meier nonparametric product limit estimator from survival analysis to the multi-state time to event models.[10] I use this widely used estimation method as a baseline. I first describe the Aalen-Johansen-Fleming estimator for models with no covariates and then I describe the method for the general case with time-varying covariates assuming the absence of unobserved heterogeneity of frailty. Section 3.2 describes the estimation procedure I follow to deal with unobserved heterogeneity of frailty.

To describe the Aalen-Johansen-Fleming estimator, I introduce some concepts and notations. For each individual $i, i = 1, 2, ..., n$ and corresponding to each transient health state h ($h = 1, 2$), define two types of stochastic processes: (1) the counting process $N_{hj,i}(t)$ denoting the **observed** number of transitions from health state h to health state j that individual i has made by time t; (2) $Y_{h,i}(t)$, taking value 1 if individual i is at risk at time t for transition to another possible health state, and taking value 0 otherwise.

For illustration, focus on one transition $h \rightarrow j$. Denote by $\bar{N}_{hj}(t) = \sum_i^n N_{hj,i}(t)$, a counting process measuring the number of transitions of the type $h \rightarrow j$ in the sample at time t, and by $\bar{Y}_h(t) = \sum_i^n Y_{h,i}(t)$, a counting process measuring the number of individuals in the sample at risk for a transition at time t. In any empirical study, data will be available only at discrete times, say in ordered times $0 = t_0 < t_1 < ... < t_m$. For a process $X(t)$ observed at these time points, denote by $\triangle X(t_i) = X(t_i) - X(t_{i-1})$ for $i > 0$ and $\triangle X(t_0) = Z(t_0)$. At each observed time t_i of the dataset, the procedure estimates each component transition intensity function by

$$\hat{\lambda}_{hj}(t_i) = \frac{\triangle \bar{N}_{hj}(t_i)}{\bar{Y}_h(t_i)}, j \neq h \tag{5}$$

The *Nelson-Aalen non-parametric estimate* of the integrated transition intensity functions is

[10] Fleming gave the estimator for complete data, Aalen and Johansen gave the estimator for censored data.

given for each $h = 1, 2$ by

$$\begin{aligned}
\hat{\Lambda}_{hj}(t) &= \sum_{i:t_i \le t} \hat{\lambda}_{hj}(t_i), j \neq h, \\
\hat{\Lambda}_{hh}(t) &= -\sum \hat{\Lambda}_{hj}(t) \text{ and} \\
\hat{\Lambda}_{hj}(t) &= 0, \text{ for } h = 3, 4; \quad j = 1, 2, 3, 4
\end{aligned} \tag{6}$$

The *Aalen-Johansen-Fleming estimator* $\hat{P}(s,t)$ for $s,t \in T, s < t$ for the transition probabilities is obtained by substituting for each component hj the Nelson-Aalen estimates $\hat{\Lambda}_{hj}(t)$ and then applying the product integral formula Eq. (3) as follows

$$\hat{P}(s,t) = \prod_{s<u<t} (I + d\hat{\Lambda}(u)) = \prod_{i:t_i \le t} (I + \triangle\hat{\Lambda}(t_i)). \tag{7}$$

With covariates, one obtains the Cox partial likelihood estimate $\hat{\beta}_{hj}$ separately for each transition $h \to j$ and then computes a weighted risk set defined by

$$\bar{Y}^*_{hj}(t) = \sum_{i=1}^{n} Y_{h,i}(t) \exp\left(X_{(i)}(t)\hat{\beta}_{hj}\right), \tag{8}$$

where $X_{(i)}$ denotes the vector of covariates of the individual i.

The estimate of the intensity functions with covariates $\hat{\lambda}_{hj}(t_i; X)$ is obtained from Eq. (5) by replacing, $\bar{Y}_h(t)$ with $\bar{Y}^*_{hj}(t)$ for each h, j, is known as the *Breslow* estmate of intensity function. The integrated hazard rates and the transition probabilities with regressors are estimated replicating the steps described below Eq. (5).

Aalen-Johnson-Fleming estimator has nice statistical properties. For instance, using Martingale calculus, it can be shown that the estimator is asymptotically unbiased and the normalized estimate is normally distributed (i.e., the central limit theorem holds for normalized parameter estimates) with an asymptotic estimable variance-covariance matrix (see for details, Aalen, Borgan, et al., 2008; Andersen et al., 1993; Fleming and Harrington, 2005; Wreede et al., 2010).

3.2 Unobserved Heterogeneity or Frailty

Much of the research on frailty is carried out for the two-state alive-death type models, i.e., in the notation of this paper, models with $S = \{1, 2\}$, 1 = alive, 2 = death or 2 = disability, treating

it as an absorbing state. Very little is known for multi-state models, as these models are very difficult to handle analytically and numerically. Three types of statistical problems arise when unobserved heterogeneity is present, discussed below.

3.2.1 Biases in parameter estimates

In the presence of unobserved heterogeneity, the parameter estimates of the included covariates become asymptotically biased (Vaupel et al., 1979; Heckman and Singer, 1984a; Aalen, Valberg, et al., 2014). To correct for the unobserved heterogeneity biases in the two-state framework, a few studies incorporated unobserved heterogeneity in various ways to estimate the fixed effects β's, the random effects b's for individuals and the frailty variance θ. I briefly describe the Ripatti and Palmgren (2000) method which I extend to the multi-state framework of this paper. Ripatti and Palmgren (2000) assumed the frailty specification to be as in the last equality of Eq. (4), specializing it to a two-state model. They assumed that the frailty random variable V is log-normally distributed with mean 1 and variance θ. They applied Laplace approximation to the marginal likelihood function[11] of the sample. They decomposed the approximated likelihood into two components, one component allowed one to apply the penalized Cox partial likelihood procedure to estimate β's, b's and their standard errors, given θ fixed, and the other component provided the estimates of θ and its standard errors given the estimated β's and b's, alternating the two-steps iteration until convergence. One fallout of their estimation procedure is that the estimated standard errors of the β's and b's are underestimated as the penalized partial likelihood for estimating those parameters and their standard errors for a fixed value of θ. In their simulation exercise, they showed that the biases are small. This estimation procedure is implemented in the R package `coxme` by Therneau (2022). The package also provides a χ^2 statistic to test the null hypothesis $H_0 : \theta = 0$.[12]

In the multi-state framework, one has a vector of frailties $B = (B_{hj}, h, j \in S, j \neq h)$, each component corresponds to a transition. No estimation procedures are available for a general distribution of B. As a first step, I consider two types of frailty distributions that enable one to apply the Ripatti and Palmgren (2000) technique and the Therneau (2022) `coxme` package to get parameter estimates for the multi-state model.. In one specification, I assume that the frailty random effects are independent across transitions. This allows one to apply the Ripatti and Palmgren

[11] *Marginal likelihood function* is the expectation of the conditional likelihood function given frailty level over the distribution of the frailty. This is also known as the *mixture likelihood function*.

[12] The specification in both Ripatti and Palmgren (2000) and Therneau (2022) are more general. The framework is capable of estimating shared frailty models, groups of records sharing common frailties by specifying appropriate design vectors for the frailty term in Eq. (4).

(2000) technique to each transition separately. These parameter estimates together with the χ^2 test statistics for each transition are reported in Table 7 and Table 8. I found that the parameter estimates are all slightly higher in magnitude compared to the parameter estimates of the Cox models without including unobserved heterogeneity. In the second specification, I assume that the random effects across transitions are identical, i.e., a common or shared random effect across all transitions for everyone. These parameter estimates are used for the policy analysis of Section 7 and the estimates of the fixed effects and the χ^2 statistic to test if the common frailty has variance 0 are reported in Table 12. Again, I found that the estimates for β's have slightly higher magnitude than the Cox partial likelihood estimates without unobserved heterogeneity (not reported).

3.2.2 Dynamic selection and mixed transition intensity and transition probabilities

The second set of problems with unobserved heterogeneity is that the estimates of transition intensities and transition probabilities computed even with the bias corrected regression coefficients of the covariates in Eq. (5) - Eq. (8) are for an average individual, i.e., for an individual with frailty level equals the population average frailty level. They will give biased estimates of the population averages because of the dynamic selection. The higher is the bias, the higher is the variance of the random effect. I explain it with the above two-state model. Let the Laplace transform of the frailty random variable V_{12} be denoted as $\mathscr{L}_{12}(c) \stackrel{def}{=} \int_0^\infty e^{-cv} dG_{12}(v;\theta)$, where $G_{12}(v;\theta)$ is the distribution function of V_{12}.[13] The superscript m on an entity will be used to represent the entity's marginal distribution or the population average. Note that the survival function of an individual with unobserved heterogeneity or frailty level $V_{12} = v$ is given by $P_{11}(t;X,v) = exp(-v\Lambda_{12}(t;X))$. The population survival function is a mixture of the individual survival functions and is given by,

$$P_{11}^m(t;X) \equiv \int_0^\infty exp(-v\Lambda_{12}(t;X)) dG_{12}(v;\theta) = \mathscr{L}_{12}(\Lambda_{12}(t;X)) \tag{9}$$

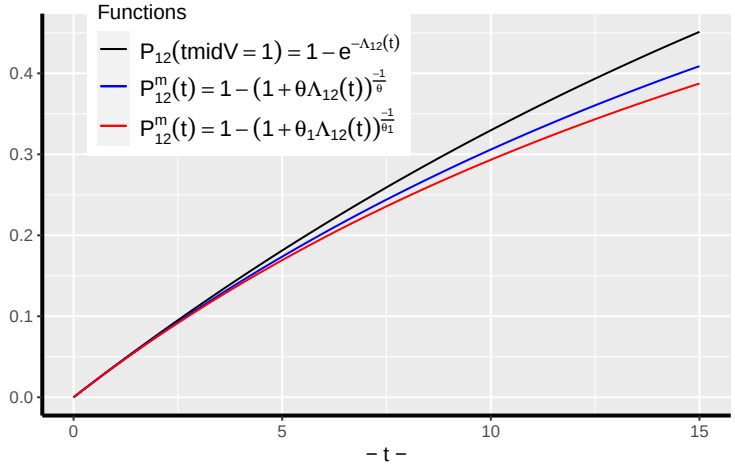
The population-average transition intensity or the hazard rate is given by

$$\begin{aligned} \lambda_{12}^m(t;X) &\equiv -dlnP_{11}^m(t;X)/dt = -\lambda_{12}(t;X)\frac{\mathscr{L}'_{12}(\Lambda_{12}(t;X))}{\mathscr{L}_{12}(\Lambda_{12}(t;X))} \\ &= \lambda_{12}(t;X)\int_0^\infty v dG_{12}(v \mid X, T \geq t;\theta) \end{aligned} \tag{10}$$

[13] When $G(.)$ is a gamma distribution, with mean 1 and variance θ, one can derive that $\mathscr{L}(c) = (1+\theta c)^{-1/\theta}$.

An individual with higher frailty level will have higher probability of encountering a transition. As time progresses, the transition free population of given characteristics X consists of higher proportion individuals of lower frailties as compared to an earlier time. This is what is meant by *dynamic selection*. This dynamic selection will make the average frailty level of the transition-free population over time — i.e., the value of the integral in the second line of Eq. (10) — become smaller and smaller, except at $t = 0$ when they are equal. A fallout of this is that the proportionality assumption of the individual intensity functions will not hold for the population-average intensity function. As a result, an estimated regression coefficient of a covariate will give biased estimate of its hazard ratio or the average treatment effect at the population level. Another fallout of the dynamic selection is that the average of the individual transition probabilities of a population $P_{12}^m(t;X)$ will be smaller and smaller over time than the transition probability of the average individual, $P_{12}(t;X,V_{12}=1)$. Furthermore, the higher the variance θ of V_{12}, the smaller becomes $P_{12}^m(t;X)$ for all time $t > 0$. When the frailty variance $\theta \to 0$, the average transition probability $P_{12}^m(t;X) \to P_{12}(t;X,V_{12}=1)$ the transition probability of the average individual. Figure 2 illustrates the above for the case of a constant baseline hazard function without covariates for the gamma frailty distribution with two values for θ.

Figure 2: Comparing the probability of a transition for an individual with average frailty level, $P_{12}(t \mid V = 1)$, with population average $P_{12}^m(t)$ for $\lambda = .04$, and $\theta = 0.492$, $\theta_1 = 0.8$



For general multi-state models such as ours, no tractable numerical algorithms are currently available to compute the population-averages of transition intensities or transition probabilities of individuals. In the policy analysis Section 7, I will use the transition intensities and transition probabilities of the average individual, i.e., the individual with all $V_{hj} = 1$ to study treatment effects of policies.

3.2.3 Indentification problems

Without covariates, Eq. (10) becomes $\lambda_{12}^m(t) = \lambda_{12}^0(t) \cdot g(t)$, where $g(t) = \Lambda_{12}'(t)/\Lambda_{12}(t)$, a function of t. From a random sample one can only estimate $\lambda_{12}^m(t)$, which is consistent with both a specification of individual intensities with individual frailty and a specification without frailty in Eq. (4). Unobserved heterogeneity cannot be identified without having further structure. Elbers and Ridder (1982) were the first to show this identification problem. They also showed that identification can be achieved when there are covariates in the proportional hazard specification and the unobserved heterogeneity has finite mean (see also Heckman and Singer (1983) and Heckman and Singer (1984b) for other identification conditions in proportional hazard specifications).

Notice that even with covariates, it will not be possible to distinguish between two specifications of individual transition intensities — one with proportionality and unobserved heterogeneity specified in Eq. (4) and non-proportional intensity function without unobserved heterogeneity of the form $\lambda_{12}^m(t; X)$ in Eq. (10)

An important point to note is that a non-parametric estimate of the transition intensity already encompasses the effect of unobserved heterogeneity if there is any. Thus, the non-parametric estimates I present below in Section 5 are free from unobserved heterogeneity biases.

4 The Dataset and the Variables

4.1 The dataset

Table 1: Summary of the health status of the individuals in the sample over the survey years.

Survey Year	Alive:normal health	Alive: diseased	Became disabled	Died before disability	Attained FRA	Total
1992	3,016	6,403	92	0	0	9,511
1994	2,591	6,608	76	144	0	9,419
1996	2,291	6,623	139	146	0	9,199
1998	1,726	5,478	134	123	727	8,188
2000	1,279	4,313	86	113	741	6,532
2002	848	3,148	54	58	759	4,867
2004	468	1,893	34	48	781	3,224
2006	132	636	4	14	795	1,581

I use the Health and Retirement Study (HRS) dataset. A lot has been reported on the family of HRS datasets — about its structure, purpose, and various modules collecting data on genetics, biomarkers, cognitive functioning, and more, see for instance (Juster and Suzman, 1995; Sonnega et al., 2014; Fisher and Ryan, 2017). I will briefly describe the survey relevant for this paper. The first survey was conducted in 1992 on a representative sample of individuals living in households i.e., in non-institutionalized, community dwelling, in the United States from the population of cohort born during 1931 to 1941 and their spouses of any age. "The sample was drawn at the household financial unit level using a multistage, national area-clustered probability sample frame. An oversample of Blacks, Hispanics (primarily Mexican Americans), and Florida residents was drawn to increase the sample size of Blacks and Hispanics as well as those who reside in the state of Florida" (Fisher and Ryan, 2017).

The number of respondents in the first survey was 13,593. Since 1992, the surveys were repeated every two years, each is referred to as a *survey wave*. New cohorts were added in 1993, 1998, 2004 and 2010, ending up with a sample size of 37,495 respondents and 23,000 households in wave 12 in 2014. For ease of use, the RAND Corporation created many variables from the original HRS data. I create all the variables (with a few exceptions noted below) from the RAND HRS dataset version P. The details of the RAND HRS version P can be found in Bugliari et al. (2016). I use the original cohort first interviewed in 1992 so that we have a homogeneous group of individuals with data for many years to avoid cohort effects in our estimates. This cohort has the largest number of respondents.

As mentioned in the introduction, I define the disability health state to be the one that qualifies one to be on the SSA's OASDI or SSI disability programs. The data on disability is self-reported. Later I plan to use the Social Security Administration's matched administrative data on this variable and earnings variables not included here. The matched data will, however, reduce the sample size to half, as only 50 percent of the respondents are used for matching HRS data with SSA Administrative data. The HRS collected information on if and when the doctor diagnosed that the respondent has any of the severe diseases like high blood pressure, diabetes, cancer, lung disease, heart attack, stroke, psychiatric disorder and severe arthritis.

I drop respondents who were enrolled onto disability programs before the first survey year 1992 and I also drop the spouses in the sample who were not born between 1931 to 1941, so that the respondents in our sample are between ages 51 to 61 and they are not disabled or dead by the first survey year 1992. I ended up with the final sample size of 9511 for this analysis. Table 1 provides summary health statistics of the cohort in our sample over the survey years.

Table 1 reports these statistics only up to the survey year 2006, as after this survey year, the individuals exited the study because of disability, death before disability, or are censored because they are over age 65. The table shows that the first period of this study 1992 has 3016 individuals in normal health, which is 32 percent of the sample respondents, 6403 individuals (i.e., 67 percent) in diseased health state with one-or-more chronic diseases and 92 individuals (i.e., 1 percent) left the study as they became disabled. No individuals died or were censored because of ages higher than 65 — this is the result of sample selection criterion mentioned above. In the next survey year 1994, out of 9419 non-exited individuals, 144 died without any disability. In the survey year 1998 for the first time, 727 individuals in the sample left our study because they reached ages above 65. The total number of individuals during the last survey round of 2006 before they all become older than 65 is 1581, i.e., about 17 percent of the original sample.

4.2 Variables

I have noted earlier the importance of the early childhood factors such as childhood socioeconomic status, childhood health status, cognitive and non-cognitive skills in determining the health developments in the middle age. Other important factors are biomarkers measuring the initial physical and mental health status in the middle age and health related behaviors. Furthermore, the health development may vary by race and sex. I describe the construction of these variables in this subsection.

I use the Item Response Theory (IRT) from the latent variable analysis literature to construct an aggregate measure of childhood socioeconomic status, **Childhood SES**.

IRT techniques are not commonly used in Economics. Originally the IRT techniques were used in the psychometry literature to measure latent traits such as cognitive ability and personality of individuals. More recently this technique has been used in healthcare fields to measure health status of individuals in clinical trials and treatments. In this procedure, the latent trait, known as *score*, is assumed to be a continuous variable and individuals differ in the levels they possess. The procedure uses responses on several test items usually with true/false or with multiple choices to estimate the level of the latent trait that an individual possesses. The probability of a particular response to an item depends on the individual's trait level and on item characteristics such as difficulty level to answer a question objectively or the imperfection of the item question to measure the trait, or an individual might be guessing a response. The IRT procedure specifies a probability model of the responses to each item as a function of the level of the latent trait and item characteristics. The procedure uses various statistical methods to estimate the latent trait

level and the characteristics of the item. Mainly three statistical estimation procedures are used in the literature — the maximum likelihood (ML) procedure, Bayesian maximum a posteriori (MAP) procedure and expected a posteriori (EAP) procedure. I have used a two parameter IRT model (which includes the well-known Rasch model as a special case) of the probabilities of item responses and the MAP procedure to estimate the individual scores and the set of item parameters. I did this in SAS. See Embretson and Reise (2000) for a lucid exposition of the basic one-dimensional IRT models and the above three estimation procedures, see Cai et al. (2016) for a survey of IRT models of multi-dimensional traits and extensions to dynamic scoring, and see An and Yung (2014) for details on the SAS IRT procedure and general introduction to various IRT procedures that SAS can perform.

The demographic variables **White** and **Female** have the standard definitions. The variable **College+** is a binary variable taking value 1 if the respondent has education level of completed college or higher (does not include some college), i.e., has a college degree and more and taking value 0 otherwise.

CES-D: I used the score on the Center for Epidemiological Studies Depression (CES-D) measure in various waves that is created by RAND release of the HRS data. RAND creates the score as the sum of five negative indicators minus two positive indicators. "The negative indicators measure whether the Respondent experienced the following sentiments all or most of the time: depression, everything is an effort, sleep is restless, felt alone, felt sad, and could not get going. The positive indicators measure whether the Respondent felt happy and enjoyed life, all or most of the time." I standardize this score by subtracting 4 and dividing by 8 the RAND measure. The wave 1 had different set of questions so it was not reported in RAND HRS. I imputed it to be the first non-missing future CES-D score. I refer the variable as CES-D in the paper. I use this variable as a measure of stress and depression, see Steffick (2000) for discussions and its validity as a measure of stress and depression.

Total Cognitive Scores: This variable is a measure of cognitive functioning. RAND combined the original HRS scores on cognitive function measure which includes "immediate and delayed word recall, the serial 7s test, counting backwards, naming tasks (e.g., date-naming), and vocabulary questions' '. Three of the original HRS cognition summary indices—two indices of scores on 20 and 40 words recalls and a third index is a score on the mental status index which is sum of scores "from counting, naming, and vocabulary tasks"—are added together to create this variable. Again, due to non-compatibility with the rest of the waves, the score in the first wave was not reported in the RAND HRS. I have imputed it by taking the first future non-missing value

of this variable.

BMI: The variable body-mass-index (BMI) is the standard measure used in the medical field and HRS collected data on this for all individuals. If it is missing in 1992, I have imputed it with the first future non-missing value for the variable. Following the criterion in the literature, I create the variable BMI taking value 1 if BMI > 25 and value 0 otherwise.

Now I describe the construction of the behavioral variables.

Behavior: Smoking: This variable is constructed to be a binary variable taking value 1 if the respondent has reported yes to ever smoked question during any of the waves as reported in the RAND HRS data and then repeated the value for all the years.

Behavior: Exercising: The RAND HRS has data on whether the respondent did vigorous exercise three or more days per week. This variable takes value 1 if the respondent did vigorous exercise three or more days per week in any of the waves, otherwise it takes value 0. The values are repeated for all the years.

Childhood SES: This variable is a binary variable measuring childhood SES. I constructed it using the IRT procedure as follows. From the HRS data I created four binary variables using the original categorical data on whether family moved for financial reason, family usually got financial help during childhood, father unemployed during childhood, father's usual occupation during childhood. For each variable, I assigned 0, if it is disadvantageous and 1 if it is advantageous; three tertiary variables two on each parent's education levels assigning 0 = High School dropout, 1 = some college, 2 = completed college and higher and the third on family financial situation (0 = poor, 1 = average, 2 = well-off). I used these seven variables as items in the IRT procedure to first compute a continuous score estimate and then I define **Childhood SES** = 1 if the score is above mean plus one standard deviation of the scores and 0 otherwise. I will discuss more about it in Section 6.

Childhood Health is a binary measure of childhood health constructed from the self-reported qualitative childhood health variable in HRS. I define **Childhood Health** = 1 if the respondent reported very good or excellent health, and zero otherwise.

Init.HLTH is a categorical variable denoting the initial health state of a respondent right before the respondent entered the Health and Retirement Study. It takes value 1, if the respondent was in normal health and value 2 if the respondent had one-or-more serious diseases.

To get an idea about distribution of the variables among different groups, Table 2 provides the averages of the above variables for different groups, and Table 3 provides the polychoric correlations among the childhood factors (cFactors), and with biomarkers and health behaviors around the age when joining the HRS.

Table 2: Averages of the above variables for various groups.

Sample	cSES	cHLTH	College+	CESD	cogtot	BMI	Smoking	Exercising
non-white	0.165	0.598	0.126	-0.285	21.168	0.711	0.599	0.736
White	0.306	0.677	0.189	-0.355	24.509	0.615	0.628	0.779
Male	0.284	0.667	0.218	-0.366	23.299	0.691	0.728	0.789
Female	0.271	0.655	0.139	-0.319	24.312	0.584	0.527	0.754
Childhood SES: poor	0.000	0.612	0.114	-0.322	23.201	0.653	0.627	0.748
Childhood SES: rich	1.000	0.787	0.340	-0.387	25.331	0.586	0.610	0.829
Childhood Health: poor	0.174	0.000	0.126	-0.299	22.848	0.626	0.636	0.642
Childhood Health: good	0.330	1.000	0.202	-0.359	24.231	0.639	0.615	0.837
No College+	0.222	0.640	0.000	-0.324	23.254	0.646	0.636	0.761
College+	0.535	0.758	1.000	-0.417	26.520	0.581	0.558	0.815

Table 3: Polychoric correlations among childhood factors and with biomarkers and health behaviors prior to joining the HRS study.

cFactors	cSES	cHLTH	College+	CESD	cogtot	BMI	Smoking	Exercising
White	0.257	0.121	0.148	-0.199	0.358	-0.150	0.044	0.080
Female	-0.025	-0.021	-0.191	0.125	0.138	-0.177	-0.328	-0.071
cSES	1.000	0.294	0.458	-0.197	0.269	-0.106	-0.026	0.164
cHLTH	0.294	1.000	0.187	-0.164	0.162	0.021	-0.034	0.367
College+	0.458	0.187	1.000	-0.289	0.395	-0.096	-0.114	0.103

From the average and correlation figures, one finds that whites have higher percent of individuals with richer Childhood SES—31 percent white and 17 percent nonwhite have rich Childhood SES with a correlation coefficient of 0.257 between these two variables. Similarly, a lower percent of individuals of poor Childhood SES has good childhood health as compared to their peers—61 percent for the poor Childhood SES and 79 percent for the rich Childhood SES with a correlation coefficient of 0.294 between these two variables. A striking difference for these two Childhood SES groups is in the percent of attaining College+ education level – 11 percent for the poor

Childhood SES and 34 percent for the rich Childhood SES with a correlation coefficient of 0.458 between these two variables. Similarly, the differences in the distribution of other variables and groups can be read from the tables.

5 Transition probability estimates

I have used the R package *mstate* that implemented most of the estimation methods with or without covariates described in Section 3. In this section, I report the estimated transition probabilities for the overall population without covariates. Denote by $P_{ij}(t)$ the probability of transition to state j by time t starting from state i at the base age 51 (the base period of this paper is age 51). For all ages including ages 51 and 65, I report corresponding estimated probabilities in Table 4. I plot these probabilities in Figure 3 from two transient states — state 1 (normal health status) in panel (a) and state 2 (diseased health status) in panel (b). The starting age in the transient state of each panel is the base age 51.

Table 4: Estimated transition probabilities by duration of stay.

Age	1⟶1	2⟶2	1⟶2	2⟶1	1⟶3	2⟶3	1⟶4	2⟶4
51	1.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
52	0.995	0.972	0.000	0.000	0.000	0.017	0.005	0.010
53	0.898	0.953	0.086	0.012	0.005	0.021	0.011	0.014
54	0.837	0.930	0.135	0.016	0.014	0.033	0.013	0.021
55	0.762	0.900	0.202	0.032	0.018	0.041	0.018	0.027
56	0.703	0.877	0.251	0.041	0.025	0.048	0.022	0.034
57	0.635	0.854	0.308	0.050	0.029	0.056	0.028	0.041
58	0.582	0.830	0.347	0.054	0.037	0.068	0.033	0.049
59	0.529	0.805	0.389	0.062	0.044	0.078	0.038	0.055
60	0.489	0.782	0.416	0.067	0.051	0.087	0.044	0.064
61	0.442	0.763	0.448	0.067	0.058	0.096	0.052	0.074
62	0.400	0.744	0.473	0.067	0.068	0.106	0.060	0.084
63	0.362	0.730	0.500	0.068	0.072	0.111	0.066	0.091
64	0.360	0.719	0.493	0.067	0.074	0.114	0.073	0.100
65	0.299	0.711	0.548	0.0686	0.075	0.115	0.0778	0.105

Notes: Recall that health state codes represent 1 = normal health, 2 = diseased, 3 = disabled, 4 = death.

The estimated probabilities show some general characteristics of the population. For instance, an individual of age 51 with normal health has the probability of getting onto the disability

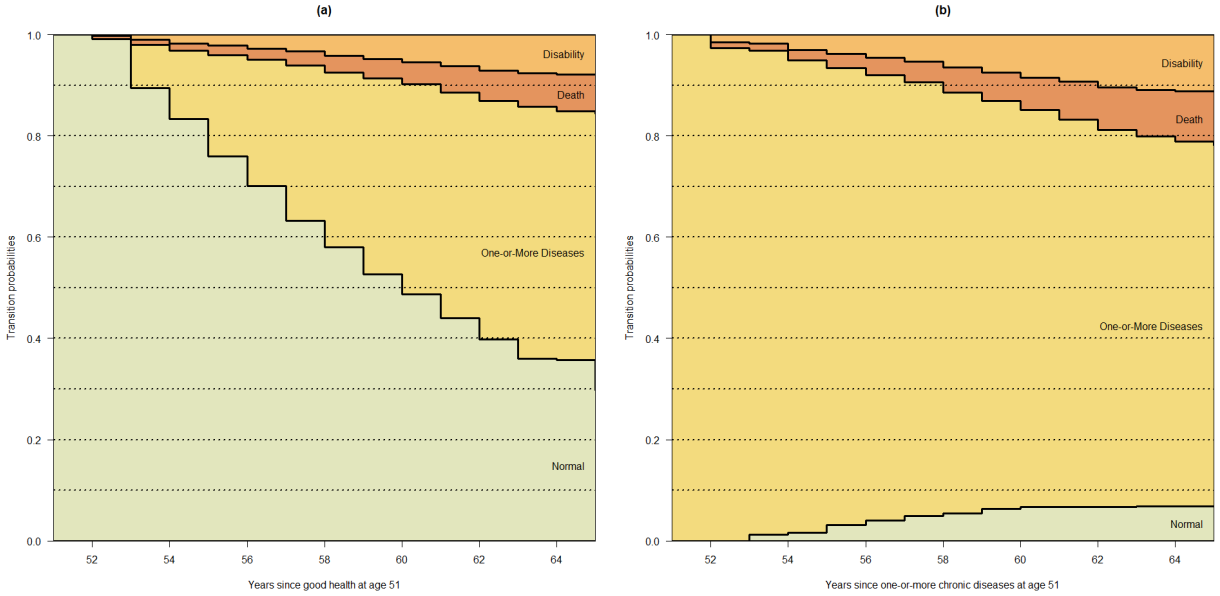
rolls 0.051 by age 60 and 0.075 by age 65. Similarly, an individual of age 51 with one or more chronic diseases has the corresponding probability of getting onto the disability rolls 0.087 by age 60 and 0.115 by age 65. The probabilities in the diseased health states are higher than the the probabilities in the normal health state. It is expected the second group to have higher probabilities. The statistical model is able to provide a quantitative estimate of how much higher, which is important for disability programs like the OASDI and SSI. Similarly, the estimates of the competing risk of death before disability starting from these two transient health states can be read from the lower part of the figure. From the estimates it appears that from normal health the probability of getting onto disability is higher than probability of death without disability by age 60. But by age 65 the probability of death before disability becomes higher than getting onto the disability. From the diseased health state, however, this pattern over age does not reverse. Comparing the transition probabilities $P_{12}(t)$ and $P_{21}(t)$, we see that by age 60 around 42 percent of the population in normal health who did not die or become disabled by then has the probability of becoming diseased (i.e., joining the pool of higher disability risk population) and the probability of reverse transition of diseased health state to normal health is about 0.069.

The multi-state time-to-event framework is also useful to estimate the effects of a policy that directly affects transition intensity of one transition on the transition probabilities of all other transitions. For instance, suppose a policy reduces only the transition intensity of death before disability of individuals in normal health and it does not affect any other transition intensities. This will directly affect the transition probability $P_{14}(t)$. It will also affect all other transition probabilities. The changed effects can be obtained by plugging-in the new transition intensity in Eq. (6), and then using it to change the Aalen-Johansen-Fleming estimate in Eq. (7). This illustrates the usefulness of a multi-state time to event model for policy analysis.

Figure 3 and Table 4 give more details of the nature of these transition probabilities age by age.

From the transition probabilities in Table 4, one can compute the probability of various transition paths. For instance, suppose one is interested in estimating the risk of disability by age 65 for a person who is in normal health state at age 60. How does that probability compare with the probability of one who is in diseased health state at age 60? Or one is interested in estimating such probabilities for an individual who joined the Health and Retirement Study at age 60. These can be obtained if we can compute $P(60, 65)$. To compute this transition probability matrix $P(60, 65)$ from the computed probabilities in Table 4, note that applying the Chapman-Kolmogorov Eq. (1), one has $P(51, 65) = P(51, 60) \cdot P(60, 65)$. This implies, $P(60, 65) = [P(51, 60)]^{-1} \cdot P(51, 65)$, (if the inverse exists).

Figure 3: Transition probabilities (a) from normal health state, (b) from one-or-more diseased health state.



The components of the estimated transition probability matrix $P(60, 65)$ and $P(51, 65)$ (the last row of Table 4) are shown below for comparison.

Table 5: Transition probabilities starting from different ages

$P(s,t)$	$1 \to 1$	$2 \to 2$	$1 \to 2$	$2 \to 1$	$1 \to 3$	$2 \to 3$	$1 \to 4$	$2 \to 4$
P(60,65)	0.579	0.877	0.374	0.038	0.021	0.034	0.026	0.051
P(51,65)	0.299	0.711	0.548	0.069	0.075	0.115	0.078	0.105

Going back to the comparison of the risk of disability by age 65 for a person who is in normal health state at age 60 with the probability of one who is in diseased health state at age 60, these probabilities from the above table turn out to be 0.021 and 0.034. Intuitively these are expected to be lower than the respective probabilities 0.075 and 0.115 from the starting age 51. The advantage of the multi-state statistical model is that it can provide quantitative comparisons such as how much lower are the probabilities in this comparison. The quantitative estimates of the transition probabilities are useful for quantitative policy analysis.

Following the above procedure, probabilities of other paths can be easily computed and compared.

6 Childhood factors and mid-age health outcomes

As mentioned in the introduction, the molecular biology literature points out that stressors of the body cells are important factors during early development and later life health progression.[14] While cellular level stressors during early cell developments cannot be directly observed or measured, many socioeconomic factors that modulate the cell level stressors can be observed. Those early life socioeconomic factors thus affect early life health outcomes. Furthermore, early life health developments together with later life health behaviors determine later life health outcomes. Health behaviors are also partly determined by cognitive abilities. Education level, an indicator of cognitive ability, can thus affect health behaviors and health developments in later life. Education is also an important determinant of earnings, which affect health related expenditures and thus health outcomes.

The HRS dataset does not have prenatal or postnatal data on individuals. It has a few variables on childhood socioeconomic status, which are correlates of the stressors of cell developments. How does one quantify childhood SES (denoted as Childhood SES now on)? There is no consensus on what exactly constitutes Childhood SES. Some studies use different sets of variables to represent Childhood SES. For instance, Heckman and Raut (2016) and a few other studies used parents' education as a measure of childhood SES in modelling the attainment of college degree. Luo and Waite (2005) used Father's and Mother's education and the Family financial well-being as regressors without aggregating them into a single measure to examine how these variables affect middle age health outcomes for the HRS sample. It is, however, useful to have a single measure of Childhood SES. Some studies used the latent variable approach to come up with a statistically defined measure of Childhood SES. For instance, Vable et al. (2017) used the Mplus software to create their latent variable measure of Childhood SES using several childhood variables from the HRS dataset. I have used a slightly different latent variable statistical procedure IRT on a set of parental characteristics during the childhood of the respondents described in Section 4.2. I use this variable and a few other variables in the Logistic regression models of the childhood factors described below.

Childhood health status (Childhood Health) is an important factor for later life health outcomes and educational attainments. The Childhood SES variable influences the stressors of the cells' environment and thus will affect Childhood Health. Apart from Childhood SES, other factors such as nutrition and pediatric healthcare are also important factors. We do not have data on

[14] Genetic make-up also controls gene expressions for producing proteins that create diseases but the epigenetic factors creating the stressors are important as well.

those variables. In the next subsection I will specify a Logit model of Childhood Health with childhood socioeconomic status and other observable characteristics as regressors.

Cognitive skill or Education level is an important factor for later life health outcomes as it determines various health related choices an individual makes throughout life. It is also an important determinant of earnings, and employment with or without covered health insurance. Similar to many studies, I use a binary education level, College+. Many factors—such as innate IQ, family background, preschool inputs, prenatal and postnatal stressors for brain development, the childhood health status, and mother's time input—determine the College+ variable. See, Heckman (2008) and Raut (2018) for recent literature on the biology of brain development and the roles played by socioeconomic factors, and Heckman and Raut (2016) for a Logit model of College+ in which a measure of IQ, family background measured with parents' education, preschool inputs and non-cognitive skills play important roles. HRS does not have data on many such variables. I use Childhood SES and Childhood Health, together with a few other demographic variables as regressors in the Logistic regression specifications College+ in the next subsection.

I examine two types of middle age health outcomes: (1) initial health status, Init.HLTH, of the respondents in their early *50s* when they first participated in the HRS study; and (2) pathways through the health states that they traversed starting from the initial health state. Both types of health outcomes are modeled as function of childhood factors, Childhood SES, Childhood Health and College+. The Section 6.1 below has the first model, and the Section 6.2 has the second model. In Section 6.3, I estimate the second model adding biomarkers and health behaviors as regressors.

6.1 Models of childhood socioeconomic status, health, and initial middle age health

In this subsection, I estimate three sets of Logistic regression models for Childhood Health, College+ and Init.HLTH. In each set, I have two specifications of Logistic regression models: in one, I include the Childhood SES measure that I created in this paper, and in the second, I include in its place three family background variables used in Luo and Waite (2005)—Father's Education, Mother's Education and Father's job situation during the respondent's childhood, controlling for other common regressors in both models, as can be seen in Table 6. I then examine if the coefficient estimates and their significance levels of the common covariates of the models are similar. If they are similar, then the single measure Childhood SES of the paper is validated as a single measure of Childhood SES. I have also calculated the pseudo R^2 defined as $R^2 = (1 - deviance/nulldeviance)$. It turns out to be the case that the parameter estimates of the

common regressors mostly do not differ in statistical significance levels and numerical magnitudes. The R^2 for the models with the regressor Childhood SES is slightly higher or close to the R^2 of the competing models. Therefore, the measure Childhood SES constructed in the paper is validated with respect to these three Logistic regression models and will be used as a childhood socioeconomic status variable.

From the statistically significant parameter estimates of the variable Childhood SES in the models with Childhood SES as a regressor, we see that Childhood SES has positive effect on childhood health, and on the probability of attaining College+ education and on the probability of possessing normal health state as opposed to diseased health state up to one's early *50s*.

The estimated College+ and Init.Health models show that a better childhood health leads to a higher probability of College+ education and a higher probability of being in normal health in one's early *50s*. A College+ education also has a significant positive effect on the probability of possessing normal health in one's early *50s*.

The estimates also show that the race variable White has positive effect on the probabilities of achieving better childhood health, attaining College+ education, and possessing normal health in one's early *50s*. The gender variable Female has no significant effect on childhood health. But females have a lower probability of completing college and remaining in good health in their early *50s*.

It is possible that even after controlling for Childhood SES, Childhood Health and College+, the race and gender variables might be picking up the effects of health behaviors. I cannot control for health behaviors in the models of this subsection as the HRS data does not have data on health behaviors prior to the survey years. In the next two subsections, I will examine the effects of race and gender variables, controlling for the effects of childhood factors, biomarkers, and health behaviors on the health developments through middle ages, more specifically on probabilities of following different health trajectories starting at age *51*.

6.2 Childhood factors, and middle age health pathways

I examine the effect of childhood factors, race and sex on the probabilities of various health pathways. The parameter estimates are shown in Table 7.

From Table 7, the estimated hazard ratio for the College+ variable is 0.309 ($\equiv exp(-1.173)$), which means that the probability of a college graduate of normal health at age 51 becoming

disabled at any age is 0.309 times the probability of a person of normal health at age 51 without a college degree becoming disabled at that age. In Cox proportional hazard models, the hazard ratio is time constant, i.e., this ratio is same at all ages.

From the estimates in Table 7 it is clear that childhood health has the most significant effect for transition $1 \rightarrow 4$ with a proportional hazard ratio of 0.08 $\equiv exp$(-2.492), i.e., the probability of dying before disability of an individual in normal health at age 51 and normal childhood health is 0.08 times the probability (s)he would have had, had (s)he had bad childhood health. A better childhood health leads to a lower probability of death before disability and a higher probability of recovering from diseased health state. A better childhood health also has a higher probability of becoming diseased but given that better childhood health also leads to higher probability recovery from diseased health state, comparing the magnitudes of the parameter estimates for transitions $1 \rightarrow 2$ and $2 \rightarrow 1$, we see that a better childhood health leads to a lower net transition from $1 \rightarrow 2$.

An individual from disadvantaged SES during childhood has a higher probability of disease incidence and a higher probability of death before disability after they become diseased.

As far as getting onto the disability rolls, the only variable that has a significant effect is College+.

A white individual has a lower probability of death before disability both from the normal health state and the diseased health state and has a lower probability of getting onto disability or recovering after being diseased. Female has a significantly lower probability of death before disability from diseased health state with a hazard ratio of 0.52 $\equiv exp$(-0.662) and no significant effect from normal health state. Furthermore, females have a lower probability of recovering from diseased health state with a hazard ratio of 0.66 $\equiv exp$(-0.414). The favorable effects of the dummy variables white and female could be their better quality of health, measured by biomarkers of health, and healthy living styles, not necessarily genetic effects. I examine these next in Section 6.3.

6.3 Childhood factors, health behaviors, biomarkers, and middle age health pathways

In this subsection, I include biomarkers and health behaviors in the model of the previous subsection. I examine the effects of biomarkers and health behaviors on pathways to disability and death before disability, and after controlling for these effects, what effects sex, race and childhood factors impart on pathways to disability and death before disability. The parameter estimates are shown in Table 8.

First, consider the biomarkers. The variable CES-D, measuring depression and stress, has the statistically most significant coefficients in Table 8. CES-D is measured in a unit such that it varies from -0.5 to 0.5. CES-D is associated with a higher probability of getting onto the disability rolls from both normal health state and diseased health state. For instance, a 0.1 unit increase in CES-D is associated with a $1.26 \equiv exp(2.336 \times 0.1)$ times higher probability of getting onto disability rolls for individuals in the normal health state and a $1.11 \equiv exp(1.073 \times 0.1)$ times higher probability of getting onto the disability rolls for individuals in the diseased health state.

For individuals in the normal health state, while CES-D does not have significant relation with the probability of death before disability, a 0.1 unit increase in CES-D, however, is associated with a $1.08 \equiv exp(0.736 \times 0.1)$ times higher probability of becoming diseased. For individuals in the diseased health state, a 0.1 unit increase in CES-D seen to be associated with a $0.93 \equiv exp(-0.73 \times 0.1)$ times lower probability of recovery to normal health, and a $1.06 \equiv exp(0.593 \times 0.1)$ times higher probability of death before disability.

The estimates in the table also show that higher cognitive total scores are associated with a lower risk of going onto the disability rolls from both the normal health state and the diseased health state—the reduction is much higher from the normal health than from the diseased health state. BMI has no significant relation with the probability of getting onto the disability rolls or the probability of death before disability. But a higher BMI is associated with a 1.01 times higher probability of becoming diseased and 1.01 times lower probability of recovery to normal health from the diseased health state. Even though BMI does not show significant effect on the probability of getting onto the disability rolls or the probability of death before disability for individuals in normal health, it will have indirect positive effects on the probabilities of these two events since individuals from diseased health state have higher probabilities of these two events that we saw in Figure 3 and Table 4.

Consider next the health behavior variables. The most important factors are smoking, showing significant adverse effects, and exercising three or more days a week regularly, showing significant favorable associations in most transitions. The variable alcohol use has no significant negative correlation in any of the transitions. Instead, it is associated with a lower risk of death before disability for people in the diseased health state. The preventive care variable has counter intuitive coefficient estimates on most transitions except that preventive care is associated with lower probabilities of death before disability. The counter intuitive result might be indicating that there is endogeneity bias since it could be the case that individuals who are in poor health might be more concerned about preventive care, or that the variable has high measurement

errors. More work with better data will be useful.

Note that the favorable effects of the White race variable have almost wiped out when we controlled for health biomarkers and health behaviors. The Female sex variable has now much lower significant effects compared to the estimates in Table 7.

Finally consider the effects of childhood factors. While some of the significant effects of these variables from the previous table ceased to be significant, they still have significant effects on a few transitions. For instance, College+ education lowers the probability of getting onto the disability rolls from diseased health state 0.76 folds. A better childhood health reduces probability of death before disability 0.19 folds for the individuals in the normal health state, 0.49 folds for the individuals in the diseased health state and increases the probability of recovery from the diseased health state 1.16 folds. Moving from disadvantaged Childhood SES to an advantaged Childhood SES lowers the probability of becoming diseased 0.89 folds.

The conducive health effects of the childhood factors will be even higher when one takes into account the indirect effects on Childhood Health, College+ and Init.HLTH (see Table 6) and on health behaviors (see correlations in Table 3).

Table 6: Effects of childhood factors, race, and sex on childhood health, college education and initial health in early 50s.

	Childhood Health		College+		Init.HLTH	
	(1)	(2)	(1)	(2)	(1)	(2)
Intercept	0.236 ***	-1.303 ***	-0.030	0.013	-1.098 ***	-1.129 ***
	(0.053)	(0.072)	(0.062)	(0.071)	(0.070)	(0.077)
White	0.293 ***	0.252 ***	0.809 ***	0.709 ***	0.201 ***	0.192 ***
	(0.053)	(0.059)	(0.056)	(0.057)	(0.058)	(0.058)
Female	-0.021	-0.111 *	-0.080	-0.048	-0.212 ***	-0.213 ***
	(0.044)	(0.049)	(0.050)	(0.051)	(0.044)	(0.044)
childhood SES	0.820 ***		1.506 ***		0.204 ***	
	(0.056)		(0.081)		(0.052)	
Father High School or some college		0.378 ***		0.791 ***		0.071
		(0.072)		(0.089)		(0.062)
Father college+		0.148		1.411 ***		0.186
		(0.134)		(0.255)		(0.111)
Mother High School or some college		0.327 ***		1.069 ***		0.135 *
		(0.067)		(0.082)		(0.059)
Mother college+		0.124		2.146 ***		0.150
		(0.151)		(0.368)		(0.127)
Family moved due to financial difficulties		0.751 ***		0.109		0.036
		(0.064)		(0.074)		(0.066)
Father unemployed during childhood		0.505 ***		-0.013		0.032
		(0.057)		(0.064)		(0.056)
Family got financial help in childhood		1.081 ***		-0.346 ***		0.009
		(0.067)		(0.081)		(0.070)
Father's Occupation		0.291 **		0.718 ***		-0.011
		(0.101)		(0.138)		(0.080)
Childhood Health			0.461 ***	0.462 ***	0.193 ***	0.164 **
			(0.050)	(0.057)	(0.048)	(0.053)
High School or some college					0.160 **	0.146 **
					(0.056)	(0.056)
College+					0.266 ***	0.234 **
					(0.073)	(0.075)
N	9511	9511	9511	9511	9511	9511
logLik	-5984.12	-5167.57	-4906.97	-4712.32	-5948.69	-5946.09
AIC	11976.24	10357.14	9823.94	9448.64	11911.38	11920.17

Table 7: Unobserved heterogeneity bias corrected estimates of Cox regression models with childhood factors as regressors.

	1→2	1→3	1→4	2→1	2→3	2→4
White	0.047	0.056	-0.727 **	0.335 **	-0.228 *	-0.242 *
	(0.063)	(0.280)	(0.225)	(0.108)	(0.109)	(0.095)
Female	0.047	-0.372	-0.446 *	-0.414 ***	-0.096	-0.662 ***
	(0.046)	(0.216)	(0.215)	(0.078)	(0.096)	(0.086)
Childhood SES	-0.089	-0.082	-0.604	-0.040	-0.064	-0.639 ***
	(0.053)	(0.253)	(0.307)	(0.092)	(0.127)	(0.135)
Childhood Health	0.326 ***	-0.072	-2.492 ***	0.360 ***	0.061	-1.684 ***
	(0.054)	(0.226)	(0.284)	(0.090)	(0.102)	(0.095)
High School or some college	-0.172 **	-0.314	0.465	0.011	-0.508 ***	0.009
	(0.060)	(0.247)	(0.258)	(0.099)	(0.104)	(0.095)
College+	-0.248 **	-1.173 **	-0.133	0.093	-1.328 ***	-0.263
	(0.075)	(0.405)	(0.395)	(0.127)	(0.206)	(0.156)
#obs	3566	3566	3566	7969	7969	7969
#events	1918	91	94	667	436	564
θ	0.1678	0.0004	1.0363	0.6188	0.0004	0.9006
χ^2	217.64	0.51	2.77	19.30	4.48	21.65
χ^2-pvalue	0.00	0.48	0.10	0.00	0.03	0.00
R^2	0.096	0.004	0.065	0.057	0.009	0.112
logLik	13989.481	-673.190	-579.556	-5410.973	-3677.818	-4346.786
AIC	27990.962	1358.380	1171.112	10833.946	7367.636	8705.573

***** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.**

Notes: The statistic χ^2 is to test the null hypothesis, H_0: $\theta = 0$, i.e., no unobserved heterogeneity and χ^2-pvalue is its p-value. Recall that health states numbers represent 1 = normal health, 2 = diseased, 3 = disabled, 4 = death.

Table 8: Unobserved heterogeneity bias corrected estimates of Cox regression models with childhood factors, biomarkers, and health behaviors.

	1→2	1→3	1→4	2→1	2→3	2→4
White	0.080	0.426	-0.821 *	0.234 *	-0.156	-0.302 *
	(0.066)	(0.320)	(0.404)	(0.114)	(0.114)	(0.146)
Female	0.075	-0.339	-0.508	-0.486 ***	-0.140	-0.413 **
	(0.050)	(0.245)	(0.413)	(0.085)	(0.104)	(0.130)
Childhood SES	-0.117 *	0.069	0.267	-0.141	-0.071	-0.084
	(0.054)	(0.268)	(0.423)	(0.094)	(0.129)	(0.158)
Childhood Health	0.025	-0.016	-1.660 ***	0.149	-0.045	-0.705 ***
	(0.056)	(0.253)	(0.375)	(0.094)	(0.105)	(0.126)
High School or some college	-0.090	0.231	0.244	-0.109	-0.272 *	0.149
	(0.066)	(0.289)	(0.470)	(0.109)	(0.117)	(0.154)
College+	-0.096	-0.348	-1.071	-0.079	-0.924 ***	-0.430
	(0.086)	(0.473)	(0.862)	(0.143)	(0.224)	(0.264)
CES-D	0.736 ***	2.336 ***	-0.827	-0.730 ***	1.073 ***	0.593 **
	(0.117)	(0.397)	(1.179)	(0.209)	(0.168)	(0.228)
Total cognitive scores	0.007	-0.087 ***	-0.008	0.012	-0.017	0.001
	(0.006)	(0.026)	(0.042)	(0.010)	(0.011)	(0.014)
BMI: Under-weight	-0.728 *	0.186	1.489	-0.303	0.507	0.698
	(0.292)	(1.023)	(1.070)	(0.505)	(0.421)	(0.463)
BMI: Overweight	0.174 **	0.047	0.282	-0.385 ***	-0.044	-0.214
	(0.053)	(0.254)	(0.410)	(0.090)	(0.127)	(0.148)
BMI: Obese	0.487 ***	0.308	0.357	-0.701 ***	0.284 *	-0.229
	(0.067)	(0.308)	(0.554)	(0.114)	(0.128)	(0.163)
Behavior: Smoking	0.104 *	0.222	2.237 **	0.047	0.298 **	0.844 ***
	(0.049)	(0.244)	(0.739)	(0.086)	(0.110)	(0.161)
Behavior: Exercising	-0.053	-0.692 *	-0.850 *	0.674 ***	-0.539 ***	-0.999 ***
	(0.071)	(0.272)	(0.418)	(0.139)	(0.107)	(0.129)
#obs	3191	3191	3191	7079	7079	7079
#events	1824	81	31	639	420	266
θ	0.1286	0.5474	0.0004	0.5236	0.0965	0.0004
χ^2	230.94	1.16	0.25	18.26	5.90	3.20
χ^2-pvalue	0.00	0.28	0.62	0.00	0.02	0.07
R^2	0.103	0.034	0.018	0.066	0.030	0.026
logLik	13064.626	-540.293	-201.946	-5109.364	-3426.463	-2192.774
AIC	26155.251	1106.587	429.891	10244.727	6878.926	4411.548

*** **p < 0.001; ** p < 0.01; * p < 0.05.**

Notes: The statistic χ^2 is to test the null hypothesis, H_0: θ =0, i.e., no unobserved heterogeneity and χ^2-pvalue is its p-value. Recall that health states numbers represent 1 =

7 Policy implications

In Section 6.3, I have shown that childhood factors, biomarkers and health behaviors are all important determinants of pathways to diseases, disability, and death. In this section, I use the estimated final model— Table 12 of the previous subsection—to compute quantitative effects of these factors on the risk of various health pathways, especially on the risk of disability in middle ages for various groups defined below.

Individuals with values of the childhood factors, Childhood SES = 0, Childhood Health = 0, College+ = 0 will be referred as *disadvantaged childhoods*, and as *advantaged childhoods* if these variables take value 1; individuals with values for biomarkers CES-D, Total Cognitive Scores at their mean values, and BMI = 1 (i.e., high BMI) as *average biomarker*, and health behaviors—Behavior:Smoking = 1, Behavior:Exercising = 0—as *poor health practices*, and as *good health practices* if these variables take the opposite values.

I consider four types of individuals, all having average values of the biomarkers and currently of age 51:

- Type 1: Disadvantaged childhoods and poor health practices.
- Type 2: Advantaged childhoods and poor health practices.
- Type 3: Advantaged childhoods and good health practices.
- Type 4: Disadvantaged childhoods and good health practices.

Consider first the white male population of the above four types. Using the parameter estimates from Table 12 in Eq. (7), I have computed the predicted transition probabilities $P(t)$, $t = 51, ..., 65$ for each of the above four types. These probabilities are plotted in Figure 4 and are shown in Table 9 for selected ages 55, 60 and 65.

Table 9: Predicted transition probabilities of white males of type τ and age a, shown as Type τ: a, where $\tau = 1, 2, 3, 4$, and $a = 55, 60, 65$.

Type: a	$P_{11}(a)$	$P_{22}(a)$	$P_{12}(a)$	$P_{21}(a)$	$P_{13}(a)$	$P_{23}(a)$	$P_{14}(a)$	$P_{24}(a)$
Type 1: 55	0.707	0.951	0.263	0.019	0.029	0.028	0.000	0.002
Type 2: 55	0.747	0.968	0.235	0.020	0.018	0.011	0.000	0.001

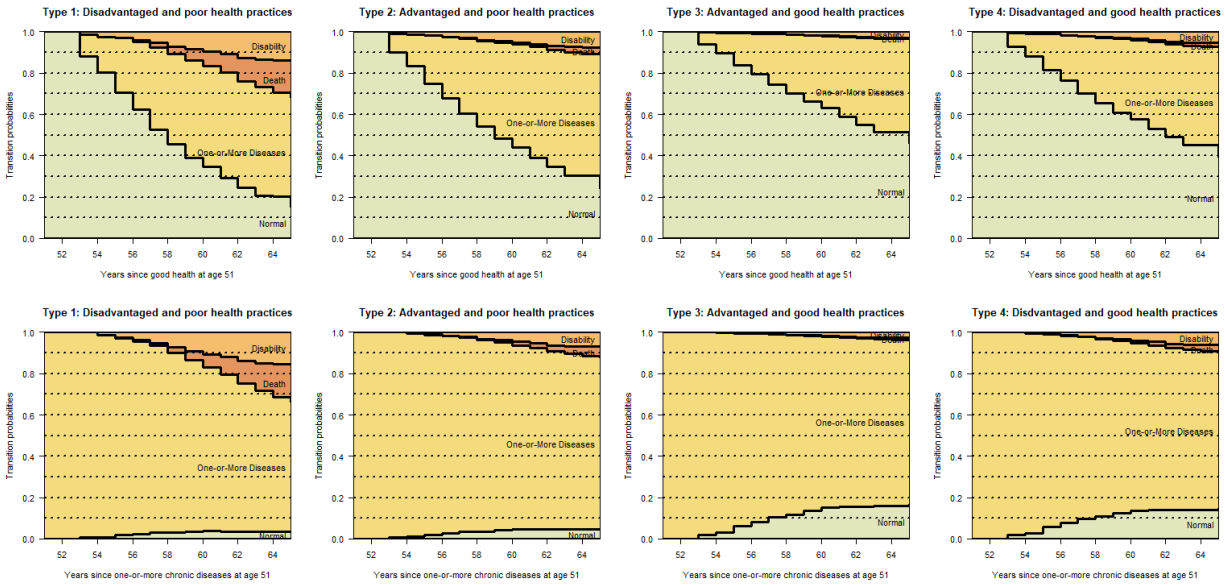
Type: a	$P_{11}(a)$	$P_{22}(a)$	$P_{12}(a)$	$P_{21}(a)$	$P_{13}(a)$	$P_{23}(a)$	$P_{14}(a)$	$P_{24}(a)$
Type 3: 55	0.839	0.935	0.155	0.061	0.007	0.004	0.000	0.000
Type 4: 55	0.813	0.930	0.176	0.060	0.011	0.010	0.000	0.000
Type 1: 60	0.347	0.794	0.489	0.038	0.094	0.106	0.070	0.062
Type 2: 60	0.441	0.893	0.496	0.044	0.052	0.046	0.010	0.017
Type 3: 60	0.627	0.833	0.352	0.146	0.020	0.018	0.001	0.003
Type 4: 60	0.576	0.809	0.383	0.138	0.035	0.041	0.006	0.011
Type 1: 65	0.155	0.632	0.529	0.032	0.141	0.157	0.175	0.180
Type 2: 65	0.242	0.832	0.644	0.044	0.078	0.071	0.036	0.053
Type 3: 65	0.456	0.799	0.510	0.164	0.030	0.028	0.004	0.009
Type 4: 65	0.395	0.754	0.530	0.148	0.054	0.063	0.021	0.036

Consider public policies that can improve childhood factors and later health behavior. Parameter estimates in Table 9 indicate that for white males of normal health status at age 51, the probability of getting onto the disability rolls by age 65, estimated as 0.141 for those with disadvantaged childhoods and poor health practices (Type 1), is reduced to 0.078 for those with advantaged childhoods and poor health practices (Type 2), reduced still farther to 0.03 for those with both advantaged childhoods and good health practices (type 3) and reduced to 0.054 for those with disadvantaged childhoods and good health practices (type 4).

Similarly, for white males of diseased health status at age 51, the probability of getting onto the disability rolls by age 65, estimated as 0.157 for those with disadvantaged childhoods and poor health practices (Type 1), is reduced to 0.071 for those with advantaged childhoods and poor health practices (Type 2), reduced still farther to 0.028 for those with both advantaged childhoods and good health practices (type 3) and reduced to 0.063 for those with disadvantaged childhoods and good health practices (type 4).

Parameter estimates in Table 9 indicate similar patterns for the event of death before disability by age 65.

Figure 4: The transition probabilities of white males of four types – top panel from normal health states and bottom panel from diseased health state.



It is important to observe that the above reductions in disability enrollment rates from introduction of public policies are after considering that the policies would reduce the death before disability rates, and thus increase the population susceptible to disability incidence after the introduction of such policies.

The improvements in health practices will also improve the scores of the biomarkers and thus, the above improvements in probabilities will be even larger.

Improvements of childhood factors involve policies for disadvantaged children to complete college (see Heckman (2008), Heckman and Raut (2016) and Raut (2018) among others for policies on this) and to achieve better health involves improvements in healthcare access of the disadvantaged children. The policies that bring changes in health behaviors will involve public health education and regulations in food industry. The latter type of policies is relatively more difficult to design and implement. Even only with the policies that improve the childhood factors for disadvantaged children, some of these disadvantaged children will have improvements in the biomarkers and in health behaviors (see Table 3 for correlations of childhood factors with the biomarkers and health behaviors), and thus, the positive effects mentioned above for the type 1 individuals will be even higher.

Furthermore, the public policies that turn type 1 individuals to type 2 individuals will also im-

prove the likelihood of having normal health in their early middle ages from 0.290 to 0.416. This will also additionally reduce the probability of death before disability and the probability of disability in the population.

Table 10: Public Policies and improvements in risks of getting onto disability and death before disability.

Main group	$P_{13}^{(1)}(65)$	$P_{13}^{(2)}(65)$	$P_{23}^{(1)}(65)$	$P_{23}^{(2)}(65)$	$P_{14}^{(1)}(65)$	$P_{14}^{(2)}(65)$	$P_{24}^{(1)}(65)$	$P_{24}^{(2)}(65)$
White Male	0.141	0.077	0.156	0.072	0.175	0.036	0.181	0.052
Non-White Male	0.124	0.066	0.180	0.084	0.284	0.053	0.233	0.069
White Female	0.121	0.062	0.145	0.065	0.122	0.025	0.126	0.036
Non-White Female	0.112	0.055	0.169	0.077	0.197	0.036	0.163	0.047

To examine the effects of the above types of policies for other race and sex, I have computed the transition probabilities $P_{ij}^{(\tau)}(65)$ for health states $i = 1, 2, j = 3, 4$, type $\tau = 1, 2$ by age 65 shown in Table 10. The table shows that the benefits are larger for all disadvantaged groups. With respect to the probability of getting onto disability and the probability of death before disability by age 65, there are racial and gender gaps: from normal health, whites have higher probabilities than nonwhites, and from the diseased health state, nonwhites have higher probabilities than whites. Probabilities of death before disability are higher for nonwhites than for whites from both health states at any given age. As for gender differences, within each race, females have lower probability of disability and lower probability of death before disability than males from both normal and diseased health states at any given age. The reduction in the probability of getting onto disability rolls by age 65 is the largest for the nonwhite diseased male individuals – reduction is from 0.18 to 0.084.

Parameter estimates in Table 6 of Section 6.1 show that improvement in Childhood SES has significant positive effects on childhood health, probability of college completion and probability of normal health in the early middle age. All these in turn have significant effects in lowering the probability of diseases, disability, and death along the middle age health pathways. From policy perspective, it is important to examine which factors are relatively more important for Childhood

SES. To that end, note that our measure of childhood Childhood SES has polychoric correlations with the component variables (all are ordered categorical variables, with higher values mean better condition) as follows: family financial situation = 0.59, family moved for financial reason = 0.49, family usually got financial help during childhood = 0.40, father unemployed during childhood = 0.43, father's usual occupation during childhood = 0.73, father's education = 0.95, mother's education = 0.84.

From the estimates we see that the most important factors to improve Childhood SES are policies that help fathers to have steady jobs during their children's childhood and parents to have higher education. Out of those two factors, policies that make disadvantaged individuals complete college is more important because it has direct individual effect as it improves the probability one's attaining College+ education; it also has intergenerational effect as it makes their children's Childhood SES improve.

Table 11: Effects of various policies on disadvantaged nonwhite males

Policy type	$P_{11}(65)$	$P_{22}(65)$	$P_{12}(65)$	$P_{21}(65)$	$P_{13}(65)$	$P_{23}(65)$	$P_{14}(65)$	$P_{24}(65)$
Type 1	0.132	0.521	0.427	0.024	0.133	0.196	0.309	0.260
only	0.196	0.624	0.536	0.036	0.132	0.186	0.137	0.154
Childhood								
Health								
only College+	0.212	0.667	0.542	0.035	0.073	0.104	0.173	0.194
Childhood	0.237	0.681	0.505	0.034	0.072	0.093	0.186	0.192
SES &								
College+								
all cFactors	0.282	0.757	0.567	0.045	0.070	0.087	0.080	0.111

8 Conclusion

The main purpose of this paper is to explore factors that affect the probability of getting onto the Social Security's Disability Insurance program or the Supplemental Security Income program and the probability of death before disability and to explore policies that affect these probabilities. To that end, the paper extends the existing disablement models of disability using insights from the biomedical literature on aging and formulates a statistical multi-state time-to-event model of pathways through episodes of normal health and diseased health states ending-up in disability, or death before disability. It studies such pathways for individuals in their mid-ages 51 - 65. The paper uses the statutory definition of disability that the Social Security Administration applies in administering these two large national programs and uses the Health and Retirement Studies

(HRS) dataset to statistically estimate the model.

The paper views aging as depletion of one's homeostatic regulatory health levels that control physiological body systems such as respiratory, cardiovascular, neuroendocrine, immune, and metabolic body systems. The higher rate of depletion of one's homeostatic health levels makes the individual become more and more frail in his/her ability to face internal and external stressors. Too much depletion of the homeostatic health levels lead to one-or-more chronic diseases such as diabetes, cancer, vascular, musculoskeletal, cognitive and mental disorders, and to disability or death.

To identify factors that modulate individual variations in pathways through health states, the paper reviews the biomedical literature on the genetic and epigenetic mechanisms at the molecular (i.e., cellular) level of the aging process culminating in age-related diseases, disability, and death. The dominant consensus of the literature utilized in the paper is that much of an individual's later life health outcomes is programmed at an early stage of life—as early as the prenatal stage, and most importantly right after the conception stage. The programming is strongly modulated by the epigenetic inputs created by the environment in the mother's womb. The genetic predisposition also matters. But epigenetic factors modulate quite strongly the programming for later life health developments. The most important epigenetic factor is stress of any kind, especially the psychological, financial, social, and chemical stresses. Other significant factors are diet, smoking, substance use, and exercising. These modulating factors are important throughout life, with stronger effects imparted in early stages of life.

Using information in the HRS dataset, and applying the Item Response Theory (IRT) in some cases to aggregate multiple measurements on some factors, the paper creates demographic factors (race and sex), measures of childhood factors (College+, Childhood SES, and Childhood Health), measures of health behaviors (smoking propensity and regular exercising) and measures of internal homeostatic health level, depression and stress level using biomarkers (body mass index (BMI), Center for Epidemiologic Studies Depression Scale (CES-D) and mid age cognitive scores).

The paper corrects for unobserved heterogeneity (or frailty) biases for the parameter estimates for multi-state models, extending a statistical procedure developed for the two-state models, and use these bias corrected parameter estimates for inference and policy analysis in the paper.

The paper finds that the childhood factors, biomarkers, and health behaviors are all statistically significant determinants of mid-age health pathways; consistent with the developmental pro-

gramming hypothesis mentioned above, the childhood factors are found to be most important, especially the Childhood SES variable which also found to have significant positive effects on Childhood Health and College+ variables and on the health state of the individual until reaching early 50s. The statistically most significant effect is from the variable CES-D measuring the level of stress and depression. For instance, a 0.1 unit increase in CES-D—the variable CES-D is measured in a scale of −0.5 to 0.5—increases the probability of getting onto the disability program 1.26 times higher for one in normal health state and 1.11 times higher for one in diseased health state, holding all other factors constant. Out of the health behaviors, smoking has statistically significant detrimental effects and regular exercising has conducive effects on the probabilities of diseases, disability, and death. After controlling for health behaviors, and early childhood factors, the paper finds that the often-cited significant good effects for women and whites turn out to be not so significant.

The paper uses the estimated statistical model to estimate quantitative policy effects on the probability of getting onto the disability rolls and on the probability of dying before disability by age 65 for various types of individuals. Individuals with low values of childhood factors are referred to as *disadvantaged childhood* and having all high values of the childhood factors as *advantaged childhood*. They are further divided by the type of health practices they follow—good health practices and poor health practices–and by their initial early 50s health status—normal initial health or diseased initial health with one-or-more chronic diseases. The paper considers public policies that turn a disadvantaged childhood individual into an advantaged childhood individual, public policies that help individuals follow good health practices and public policies that can achieve both.

For a disadvantaged childhood individual of normal initial health following bad health practices, the model predicts that policies that turn him to an advantaged individual with no changes in his health practices will reduce his probability of getting onto the disability rolls by age 65 from about 0.141 to about 0.078. The policies that change him from following bad health practices to following good health practices without changing his disadvantaged status will lower the probability to about 0.054. Policies that change both his disadvantage status and health practices, reduce the probability even further to about 0.03. It follows that the policies that improve health practices will also lower the disability entitlement probability of an advantaged individual following poor health practices by about 0.048. Similarly, for a disadvantaged childhood individual of diseased initial health following bad health practices, the probability of his getting onto the disability rolls decreases from about 0.157 to about 0.071 with policies changing only his disad-

vantaged status, to about 0.063 with policies changing only his health practices, and even further to about 0.028 with policies changing both.

The effects of the above types of policies on the probability of death before disability by age 65 are similar. For instance, for a disadvantaged childhood individual of normal initial health following bad health practices, the probability reduces from about 0.175 to about 0.036 from policies improving his childhood factors only, to about 0.021 from policies improving his health practices only, and to about 0.004 from both types of policies. Such policies will also improve biomarkers, and thus, the above improvements in probabilities will be even larger.

The quantitative effects on the disadvantaged population of other race and sex are slightly larger than the effects mentioned above for the white males. There are, however, significant racial and gender gaps in the probabilities of various transitions. For instance, with respect to the probability of getting onto disability rolls by age 65, whites have higher probabilities than nonwhites if they are in normal health at any given age, and nonwhites have higher probabilities than whites if they are in diseased health state at any given age. The probability of death before disability is higher for nonwhites from both diseased and normal current health states. As for gender differences, within each race, females have lower probability of getting onto disability rolls and lower probability of death before disability from both normal and diseased health states at any given age. The reduction in the probability of getting onto disability rolls by age 65 is the largest for the nonwhite males if they are in diseased health state at any given age. For instance, at age 51, the reduction in probability is from about 0.18 to about 0.084.

Improvements of childhood factors involve policies for disadvantaged children to complete college (see Heckman (2008), Heckman and Raut (2016) and Raut (2018) among others for policies on this) and to achieve better health involves improvements in healthcare access of the disadvantaged children. The policies that bring changes in health behaviors will involve public health education and regulations in food industry (see, Mozaffarian (2016), Timpel et al. (2019), Budreviciute et al. (2020) among others). The latter types of policies are relatively more difficult to design and implement. Even only with the policies that improve the childhood factors for disadvantaged children, some of the children will also have improvements in their biomarkers, health behaviors, and the likelihood of having normal health in their early middle ages. These will additionally reduce the probability of getting onto the disability rolls and the probability of death before disability in the population.

9 Appendix

Table 12: Unobserved heterogeneity bias corrected estimates of Cox regression models with childhood factors, biomarkers and health behaviors - common random effect across transitions.

	1→2	1→3	1→4	2→1	2→3	2→4
White	0.058	0.433	-0.843 *	0.234 *	-0.178	-0.296 *
	(0.066)	(0.320)	(0.402)	(0.114)	(0.114)	(0.145)
Female	0.081	-0.330	-0.449	-0.489 ***	-0.097	-0.411 **
	(0.050)	(0.244)	(0.404)	(0.085)	(0.104)	(0.129)
Childhood SES	-0.130 *	0.104	0.303	-0.153	-0.116	-0.071
	(0.054)	(0.264)	(0.420)	(0.094)	(0.127)	(0.157)
Childhood Health	0.014	-0.006	-1.655 ***	0.139	-0.070	-0.712 ***
	(0.056)	(0.252)	(0.373)	(0.093)	(0.105)	(0.125)
College	-0.026	-0.568	-1.264	0.023	-0.701 ***	-0.565 *
	(0.063)	(0.396)	(0.757)	(0.106)	(0.202)	(0.227)
CES-D	0.785 ***	2.289 ***	-0.853	-0.686 **	1.143 ***	0.607 **
	(0.116)	(0.391)	(1.166)	(0.207)	(0.166)	(0.226)
Total cognitive scores	0.004	-0.081 **	-0.005	0.009	-0.027 **	0.005
	(0.006)	(0.024)	(0.039)	(0.010)	(0.010)	(0.013)
BMI	0.269 ***	0.065	0.298	-0.482 ***	0.087	-0.224
	(0.049)	(0.228)	(0.377)	(0.082)	(0.110)	(0.131)
Behavior: Smoking	0.092	0.199	2.241 **	0.058	0.314 **	0.858 ***
	(0.049)	(0.243)	(0.738)	(0.085)	(0.110)	(0.161)
Behavior: Exercising	-0.081	-0.724 **	-0.891 *	0.686 ***	-0.569 ***	-1.029 ***
	(0.071)	(0.270)	(0.416)	(0.139)	(0.107)	(0.128)
#obs	3192	3192	3192	7080	7080	7080
#events	1824	81	31	639	420	266
θ	0.242					
χ^2	290.28					
χ^2-pvalue	0.00					
R^2	0.024	0.020	0.017	0.020	0.023	0.026
logLik	-13306.642	-563.470	-202.833	-5283.710	-3456.878	-2196.111
AIC	26633.284	1146.940	425.665	10587.420	6933.755	4412.223

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Notes: The statistic χ^2 is to test the null hypothesis, H_0: $\theta = 0$, i.e., no unobserved heterogeneity and χ^2-pvalue is its p-value. Recall that health states numbers represent 1 = normal health, 2 = diseased, 3 = disabled, 4 = death.

References

[1] Aalen, O. O., Borgan, Ø., et al. Survival and Event History Analysis. Springer New York, 2008. DOI: 10.1007/978-0-387-68560-1 (cit. on p. 15).

[2] Aalen, O. O. and Johansen, S. An Empirical Transition Matrix for Non-Homogeneous Markov Chains Based on Censored Observations, *Scandinavian Journal of Statistics*, **5**, no. 3 (1978), 141–150 (cit. on p. 14).

[3] Aalen, O. O., Valberg, M., et al. Understanding variation in disease risk: the elusive concept of frailty, *International Journal of Epidemiology*, **44**, no. 4 (Dec. 2014), 1408–1421. DOI: 10.1093/ije/dyu192 (cit. on p. 16).

[4] Albrecht, G. L. and Verbrugge, L. M. "The global emergence of disability", *Handbook of Social Studies in Health and Medicine*. Ed. by G. L. Albrecht et al. London: Sage, 2003, 293–307 (cit. on p. 4).

[5] Alisch, R. S. et al. Age-associated DNA methylation in pediatric populations, *Genome Research*, **22**, no. 4 (2012), 623–632. DOI: 10.1101/gr.125187.111 (cit. on p. 6).

[6] Altman, B. M. "Definitions, Models, Classifications, Schemes, and Applications", *Handbook of Disability Studies*. SAGE Publications, Inc., 2001, 97–122. DOI: 10.4135/9781412976251.n4 (cit. on p. 4).

[7] An, X. and Yung, Y.-F. Item response theory: what it is and how you can use the IRT procedure to apply it, *SAS Institute Inc. SAS364-2014* (2014) (cit. on p. 22).

[8] Andersen, P. K. et al. Statistical Models Based on Counting Processes. Springer-Verlag, New York, 1993. DOI: 10.1007/978-1-4612-4348-9 (cit. on p. 15).

[9] Austad, S. N. and Fischer, K. E. Sex Differences in Lifespan, *Cell Metabolism*, **23**, no. 6 (2016), 1022–1033. DOI: 10.1016/j.cmet.2016.05.019 (cit. on p. 6).

[10] Barker, D. J. P. In utero programming of chronic disease, *Clinical Science*, **95**, no. 2 (1998), 115–128. DOI: 10.1042/cs0950115 (cit. on p. 6).

[11] Barker, D. J. P. The fetal and infant origins of adult disease. *BMJ: British Medical Journal*, **301**, no. 6761 (1990), 1111 (cit. on p. 6).

[12] Barres, R. and Zierath, J. R. DNA methylation in metabolic disorders, *The American Journal of Clinical Nutrition*, **93**, no. 4 (2011), 897S–900S. DOI: 10.3945/ajcn.110.001933 (cit. on p. 6).

[13] Bedirhan, O. T. et al. Developing the World Health Organization Disability Assessment Schedule 2.0, *Bulletin of the World Health Organization*, **88**, no. 11 (2010), 815–823. DOI: 10.2471/blt.09.067231 (cit. on p. 5).

[14] Blair, S. N. et al. Physical Fitness and All-Cause Mortality: A Prospective Study of Healthy Men and Women, *JAMA*, **262**, no. 17 (1989), 2395–2401. DOI: 10.1001/jama.1989.03430170057028 (cit. on p. 6).

[15] Boks, M. P. et al. The Relationship of DNA Methylation with Age, Gender and Genotype in Twins and Healthy Controls, *PLoS ONE*, **4**, no. 8 (Aug. 2009). Ed. by J. Najbauer, e6767. DOI: 10.1371/journal.pone.0006767 (cit. on p. 6).

[16] Budreviciute, A. et al. Management and Prevention Strategies for Non-communicable Diseases (NCDs) and Their Risk Factors, *Frontiers in Public Health*, **8** (Nov. 2020). DOI: 10.3389/fpubh.2020.574111 (cit. on p. 44).

[17] Bugliari, D. et al. *RAND HRS Data Documentation, Version P*. Tech. rep. 2016 (cit. on p. 20).

[18] Cai, L. et al. Item Response Theory, *Annual Review of Statistics and Its Application*, **3**, no. 1 (June 2016), 297–321. DOI: 10.1146/annurev-statistics-041715-033702 (cit. on p. 22).

[19] Case, A. and Deaton, A. Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century, *Proceedings of the National Academy of Sciences*, **112**, no. 49 (2015), 15078–15083. DOI: 10.1073/pnas.1518393112 (cit. on pp. 6, 7).

[20] Conti, R. M. et al. "Early Retirement and DI/SSI Applications: Exploring the Impact of Depression", *Health at Older Ages: The Causes and Consequences of Declining Disability among the Elderly*. NBER Chapters. National Bureau of Economic Research, Inc, 2009, 381–408 (cit. on p. 7).

[21] Elbers, C. and Ridder, G. True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model, *The Review of Economic Studies*, **49**, no. 3 (July 1982), 403–409. DOI: 10.2307/2297364 (cit. on p. 19).

[22] Embretson, S. and Reise, S. Item Response Theory. Item Response Theory for Psychologists. Taylor & Francis, 2000 (cit. on p. 22).

[23] Esteller, M. Epigenetics in Cancer, *New England Journal of Medicine*, **358**, no. 11 (2008), 1148–1159. DOI: 10.1056/nejmra072067 (cit. on p. 6).

[24] Fisher, G. G. and Ryan, L. H. Overview of the Health and Retirement Study and Introduction to the Special Issue, *Work, Aging and Retirement*, **4**, no. 1 (2017). Ed. by M. Wang, 1–9. DOI: 10.1093/workar/wax032 (cit. on p. 20).

[25] Fleming, T. R. Nonparametric Estimation for Nonhomogeneous Markov Processes in the Problem of Competing Risks, *The Annals of Statistics*, **6**, no. 5 (1978), 1057–1070 (cit. on p. 14).

[26] Fleming, T. R. and Harrington, D. P. Counting Processes and Survival Analysis. Wiley, 2005 (cit. on p. 15).

[27] Gluckman, P. D. et al. Effect of In Utero and Early-Life Conditions on Adult Health and Disease, *New England Journal of Medicine*, **359**, no. 1 (2008), 61–73. DOI: 10.1056/nejmra0708473 (cit. on p. 6).

[28] Hahn, H. Toward a politics of disability: Definitions, disciplines, and policies, *The Social Science Journal* (1985) (cit. on p. 4).

[29] Hannum, G. et al. Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates, *Molecular Cell*, **49**, no. 2 (2013), 359–367. DOI: 10.1016/j.molcel.2012.10.016 (cit. on p. 6).

[30] Heckman, J. J. Schools, Skills and Synapses, *Economic Inquiry*, **46**, no. 3 (July 2008), 289–324. DOI: 10.1111/j.1465-7295.2008.00163.x (cit. on pp. 29, 39, 44).

[31] Heckman, J. J. and Raut, L. K. Intergenerational long-term effects of preschool-structural estimates from a discrete dynamic programming model, *Journal of Econometrics*, **191**, no. 1 (2016), 164–175. DOI: 10.1016/j.jeconom.2015.10.001 (cit. on pp. 28, 29, 39, 44).

[32] Heckman, J. J. and Singer, B. A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data, *Econometrica*, **52**, no. 2 (1984), 271–320. DOI: 10.2307/1911491 (cit. on p. 16).

[33] Heckman, J. J. and Singer, B. The Identifiability of the Proportional Hazard Model, *The Review of Economic Studies*, **51**, no. 2 (1984), 231–241. DOI: 10.2307/2297689. (cit. on p. 19).

[34] Heckman, J. J. and Singer, B. "The identification problem in econometric models for duration data", *Advances in Econometrics*. Cambridge University Press, Feb. 1983, 39–78. DOI: 10.1017/cbo9781139052160.002 (cit. on p. 19).

[35] Horvath, S. DNA methylation age of human tissues and cell types, *Genome Biology*, **14**, no. 10 (2013), 3156. DOI: 10.1186/gb-2013-14-10-r115 (cit. on p. 6).

[36] Institute of Medicine. Improving the Social Security Disability Decision Process. Ed. by J. D. Stobo et al. Washington, DC: The National Academies Press, 2007. DOI: 10.17226/11859 (cit. on p. 5).

[37] Juster, F. T. and Suzman, R. An Overview of the Health and Retirement Study, *The Journal of Human Resources*, **30** (1995), S7. DOI: 10.2307/146277 (cit. on p. 20).

[38] Karakus, M. C. and Patton, L. C. Depression and the Onset of Chronic Illness in Older Adults: A 12-Year Prospective Study, *The Journal of Behavioral Health Services & Research*, **38**, no. 3 (2011), 373–382. DOI: 10.1007/s11414-011-9234-2 (cit. on p. 7).

[39] Luo, Y. and Waite, L. J. The Impact of Childhood and Adult SES on Physical, Mental, and Cognitive Well-Being in Later Life, *The Journals of Gerontology: Series B*, **60**, no. 2 (Mar. 2005), S93–S101. DOI: 10.1093/geronb/60.2.s93 (cit. on pp. 7, 28, 29).

[40] Marks, D. Models of disability, *Disability and Rehabilitation*, **19**, no. 3 (1997), 85–91. DOI: 10.3109/09638289709166831 (cit. on p. 4).

[41] Mozaffarian, D. Dietary and Policy Priorities for Cardiovascular Disease, Diabetes, and Obesity, *Circulation*, **133**, no. 2 (Jan. 2016), 187–225. DOI: 10.1161/circulationaha.115.018585 (cit. on p. 44).

[42] Nagi, S. Z. An Epidemiology of Disability among Adults in the United States, *The Milbank Memorial Fund Quarterly. Health and Society*, **54**, no. 4 (1976), 439. DOI: 10.2307/3349677 (cit. on p. 5).

[43] Nagi, S. Z. "Disability in America: Toward a national agenda for prevention", ed. by A. M. Pope and A. R. Tarlov. Washington, DC: National Academies Press, 1991. Chap. Disability concepts revised: Implication for prevention, 309–327. DOI: 10.17226/1579 (cit. on p. 5).

[44] Nagi, S. Z. "Some conceptual issues in disability and rehabilitation", *Sociology and Rehabilitation*. Ed. by M. B. Sussman. 100-113: American Sociological Association, 1965 (cit. on p. 5).

[45] Pope, A. M. and Tarlov, A. R. Disability in America: Toward a national agenda for prevention. Ed. by A. M. Pope and A. R. Tarlov. National Academies Press, 1991. DOI: 10.17226/1579 (cit. on p. 5).

[46] Raut, L. K. Exits from Disability: Estimates from a Competing Risk Model, *Social Security Bulletin*, **77**, no. 3 (2017), 15–38. DOI: https://www.ssa.gov/policy/docs/ssb/v77n3/v77n3p15.html (cit. on pp. 9, 11).

[47] Raut, L. K. Long-term Effects of Preschool on School Performance, Earnings and Social Mobility, *Studies in Microeconomics*, **6**, no. 1-2 (June 2018), 24–49. DOI: 10.1177/2321022218802023 (cit. on pp. 29, 39, 44).

[48] Renna, F. Alcohol Abuse, Alcoholism, and Labor Market Outcomes: Looking for the Missing Link, *ILR Review*, **62**, no. 1 (2008), 92–103. DOI: 10.1177/001979390806200105 (cit. on p. 7).

[49] Ripatti, S. and Palmgren, J. Estimation of Multivariate Frailty Models Using Penalized Partial Likelihood, *Biometrics*, **56**, no. 4 (2000), 1016–1022. DOI: https://doi.org/10.1111/j.0006-341X.2000.01016.x (cit. on p. 16).

[50] Seib, C. et al. Stress, Lifestyle, and Quality of Life in Midlife and Older Australian Women: Results From the Stress and the Health of Women Study, *Women's Health Issues*, **24**, no. 1 (2014), e43–e52. DOI: 10.1016/j.whi.2013.11.004 (cit. on p. 7).

[51] Simons, R. L. et al. Economic hardship and biological weathering: The epigenetics of aging in a U.S. sample of black women, *Social Science & Medicine*, **150** (2016), 192–200. DOI: 10.1016/j.socscimed.2015.12.001 (cit. on p. 7).

[52] Snyder, A. R. et al. Using Disablement Models and Clinical Outcomes Assessment to Enable Evidence-Based Athletic Training Practice, Part I: Disablement Models, *Journal of Athletic Training*, **43**, no. 4 (2008), 428–436. DOI: 10.4085/1062-6050-43.4.428 (cit. on p. 4).

[53] Sonnega, A. et al. Cohort Profile: the Health and Retirement Study (HRS), *International Journal of Epidemiology*, **43**, no. 2 (2014), 576–585. DOI: 10.1093/ije/dyu067 (cit. on p. 20).

[54] Steffick, D. E. Documentation of affective functioning measures in the Health and Retirement Study, *Ann Arbor, MI: University of Michigan* (2000) (cit. on p. 22).

[55] The US Burden of Disease Collaborators. The State of US Health, 1990-2016: Burden of Diseases, Injuries, and Risk Factors Among US States, *JAMA*, **319**, no. 14 (2018), 1444–1472. DOI: 10.1001/jama.2018.0158 (cit. on p. 6).

[56] Therneau, T. M. Mixed effects Cox models, *CRAN repository* (2022) (cit. on p. 16).

[57] Timpel, P. et al. What should governments be doing to prevent diabetes throughout the life course?, *Diabetologia*, **62**, no. 10 (Aug. 2019), 1842–1853. DOI: 10.1007/s00125-019-4941-y (cit. on p. 44).

[58] Vable, A. M. et al. Validation of a theoretically motivated approach to measuring childhood socioeconomic circumstances in the Health and Retirement Study, *PLOS ONE*, **12**, no. 10 (Oct. 2017). Ed. by A. Fraser, e0185898. DOI: 10.1371/journal.pone.0185898 (cit. on p. 28).

[59] Vaupel, J. W. Biodemography of human ageing, *Nature*, **464**, no. 7288 (Mar. 2010), 536–542. DOI: 10.1038/nature08984 (cit. on p. 6).

[60] Vaupel, J. W. et al. The impact of heterogeneity in individual frailty on the dynamics of mortality, *Demography*, **16**, no. 3 (Aug. 1979), 439–454. DOI: 10.2307/2061224 (cit. on p. 16).

[61] Verbrugge, L. M. and Jette, A. M. The disablement process, *Social Science & Medicine*, **38**, no. 1 (1994), 1–14. DOI: 10.1016/0277-9536(94)90294-1 (cit. on p. 5).

[62] Verbrugge, L. M., Latham, K., et al. Aging With Disability for Midlife and Older Adults, *Research on Aging*, **39**, no. 6 (2017), 741–777. DOI: 10.1177/0164027516681051 (cit. on p. 5).

[63] World Health Organization. *International Classification of Functioning, Disability and Health*. Tech. rep. Geneva, 2001 (cit. on p. 5).

[64] Wreede, L. C. de et al. The mstate package for estimation and prediction in non- and semi-parametric multi-state and competing risks models, *Computer Methods and Programs in Biomedicine*, **99**, no. 3 (2010), 261–274. DOI: 10.1016/j.cmpb.2010.01.001 (cit. on p. 15).

[65] Zarulli, V. et al. Women live longer than men even during severe famines and epidemics, *Proceedings of the National Academy of Sciences*, **115**, no. 4 (2018), E832–E840. DOI: 10.1073/pnas.1701535115 (cit. on p. 6).