Estimating a Process of Health Formation in Older Adults: The Roles of Physical Functioning, Cognition, and Mental Health.

By Robert Millard* and Steven Stern†

November 24, 2025

Abstract

This paper develops a model to examine the dynamic process of health formation in older adults. The transition process allows for complementarities across components and the influence of external factors, including employment, healthy behaviors and economic conditions. Considering an application that specifies health components related to physical functioning, cognitive functioning, and mental health, we estimate the transition process utilizing a comprehensive set of health measures from the Health and Retirement Study. The preliminary results find positive complementarities across all of the health components, with physical functioning showing strong persistence and a significant impact on mental health. Finally, we map the latent components to widely used indicators such as self-reported health, showing that each component is a significant determinant of overall self-assessment. This mapping clarifies what aspects of underlying health dynamics are captured by common summary measures and provides guidance for interpreting empirical work that relies on them.

1 Introduction

A well-established gradient exists between health and economic well-being near the end of working life and into retirement, with causal relationships operating in both directions (Cutler et al., 2008). However,

*Assistant Professor, Department of Economics, Stony Brook University, 100 Nicolls Rd, Stony Brook, New York, USA, 11794. Email: robert.millard@stonybrook.edu.

†Professor, Department of Economics, Stony Brook University, 100 Nicolls Rd, Stony Brook, New York, USA, 11794. Email: steven.stern@stonybrook.edu.

the relative importance of the more nuanced mechanisms driving this relationship is less well understood. Health is a complex, dynamic, and multidimensional concept whose implications for work, retirement, and disability policies vary considerably depending on its manifestation. Furthermore, the estimated economic consequences of health are highly sensitive to the methods used to measure and model it (French and Jones, 2017; Blundell et al., 2023). Despite extensive research on the causes and consequences of health, there lacks a unified consensus on how to best measure and model it in economic studies.

In this paper, we develop a model of health formation that tractably captures the complexity across conditions by disaggregating health into a sparse set of latent components. The framework combines a large set of health measures, mapping them into a smaller and more interpretable set of health components. We model the dynamic interdependence among the components of health to identify both self- and cross-complimentarities in shaping future health outcomes. Moreover, the latent health components are also shaped by individual choices and economic conditions. We apply our analysis to better understand the determinants of health formation and how health drives economic behaviors near the end of working life. The results help to understand the mechanisms linking health to labor market outcomes, yielding important implications for guiding policy aimed at reducing health related disparities that emerge as individuals transition out of work.

We select the health components ex ante with the intent to capture dimensions of health that are most relevant to work capacity, and that have garnered interest in both health and labor studies.[1] The health components are categorized into cognitive functioning, physical functioning, and mental health. Physical functioning encompasses the ability to perform essential tasks in work and daily life, such as getting out of bed, walking, and lifting objects. Similarly, cognitive functioning refers to capabilities related to attention, memory, problem solving, planning, decision-making, and reasoning. Mental health reflects an individual's emotional and psychological well-being, influencing how one thinks, feels, and handles stress. Our choice of health components has an intuitive appeal. First, physical and cognitive health represent two common dimensions of functional capacity with grounded connections to work and retirement decisions (Blundell et al. (2023); Capatina and Keane (2023), and Millard (2025)). Second, there is an emerging economic literature documenting the considerable consequences of mental health challenges in the labor market (e.g., Wang et al. (2023)). Health deterioration in these dimensions can significantly affect work productivity, financial management, and the navigation of complex medical treatment plans. Although recent research has documented the economic consequences of mental health, the economic determinants of its formation remain insufficiently understood.

[1] Some recent examples of these studies include Cunha et al. (2010); Lise and Postel-Vinay (2020); Jolivet and Postel-Vinay (2020); Guvenen et al. (2020); Blundell et al. (2023); Capatina and Keane (2023); Wen (2022); Millard (2025); Darden (2022).

We estimate our model using data from the Health and Retirement Study (HRS), which provides a rich set of objective and subjective health measures, as well as other relevant variables, to inform our understanding of one's health and economic outcomes. We map the many observed indicators into a smaller, interpretable set of latent health components and estimate the transition technology from their joint distribution over time. The resulting structural parameters of the health formation process characterize the cross-elasticities between health components, which measure the degree of complementarity across components in the formation of health. In addition, the process considers the impact of employment, health-related behaviors, and individual heterogeneity on the development of future health outcomes. We then use the framework to map latent components to common summary measures, such as self-reported health and objective conditions, thereby clarifying which dimensions of health drive their estimated effects on outcomes, including employment and mortality.

First, we document stylized facts about the dynamics of health measures, their covariance over age, and their relationship to employment, healthy behaviors, and other health outcomes (e.g., doctor-diagnosed conditions and mortality). We find a positive correlation across all health measures that remains stable as people age. Although physical and cognitive health measures deteriorate as people age, measures of mental health tend to improve at a decreasing rate. Moreover, observed health measures are positively correlated with employment and healthy behaviors (e.g., moderate drinking, exercise), and negatively correlated with unhealthy behaviors (e.g., smoking) and adverse outcomes such as mortality.

Second, we estimate and analyze the technology that governs the formation of physical, cognitive, and mental health in older age. In the current draft, we adopt a linear transition technology. The fully parameterized model formalizes a dual relationship between health and economic conditions, identifying how employment and other health inputs shape future health and, in turn, how health feeds back into these choices. That is, we measure the channels through which, for example, physical health affects mental health and vice versa. A positive cross-dependence may arise if deterioration in physical health restricts daily activities, reduces enjoyment in life and purpose in work, thereby lowers mental health. Conversely, the biological and behavioral consequences of persistently poor mental health (e.g., inflammation, sleep disruption, social withdrawal, substance use) can degrade physical health and increase mortality risk. (Case and Deaton, 2022; Ruhm, 2025).[2]

Our results show significant dependence across the latent health components. Notably, physical health is an important determinant for the development of mental health. Furthermore, our findings indicate

[2] For instance, Ruhm (2025) shows that deteriorating mental health accounts for an estimated 9 % to 29 % of the rise in mortality rates among prime-age Whites in recent years.

that lagged employment positively affects each latent health component, thereby slowing the rate of health decline. This observation aligns with studies indicating that health deteriorates when individuals exit the workforce (Fitzpatrick and Moore, 2018; Black et al., 2018). In contrast, deterioration of health components reduces the probability of employment, with the largest effects from physical health. Declining mental health also significantly lowers the probability of employment, emphasizing the importance of controlling for mental health and its dependence on physical health when investigating the relationship between health and economic outcomes at older ages.

Furthermore, our analysis provides insight into how objective health conditions and an individual's self-perceived health relate to latent health components. Notably, we observe that an individual's self-assessment of their overall health is generated by a combination of these latent health factors. While physical health is the primary determinant of self-reported health, mental and cognitive health also significantly contribute. Decomposing self-reported health into latent components enables us to identify more precisely which aspects of health underlie results in studies that use self-reported measures to examine the relationship between health and economic outcomes. For example, two individuals may assess their overall health similarly, yet one may have low physical health and the other may experience poor mental health. As a result, the estimated effect of self-reported health on employment will fall between the estimated effects of physical and mental health on employment.

This paper makes several novel contributions to the related literature. We fit into a large literature with the specific intention of understanding how individuals arrive at older ages with mental, cognitive, and physical health and financial resources to support them throughout the rest of their lives. We frame our contribution in four broad areas.

Analyzing micro-level relationships between health and economic behavior is inherently complex due to challenges in measuring health itself (for detailed discussions, see Currie and Madrian (1999); O'Donnell et al. (2015); French and Jones (2017), and Blundell et al. (2023)). Our paper presents a comprehensive representation of health that reduces the dependence on single or small set of health indicators. The health economics literature has long explored subjective measures (e.g., Butler et al. (1987); Benítez-Silva et al. (2004); French (2005); Kreider and Pepper (2007); Meyer and Mok (2019)), objective measures (e.g., Bartel and Taubman (1979); Bound (1989); Smith (2004)), and their combinations (e.g., Stern (1989); Blundell et al. (2023)). As a result, the estimated effects of health on employment and related outcomes vary across studies. Our framework accommodates both subjective and objective measures, yielding a tractable yet flexible representation of individual health.

More specifically, our approach aligns with economic literature that employs dynamic factor models to condense a large number of measures into a more sparse and interpretable set of latent variables (Cunha et al., 2010; Bound et al., 2010; Iskhakov, 2010; Poterba et al., 2017; Cunha et al., 2021; Blundell et al., 2023)). Numerous studies utilize unidimensional health indices to investigate health-related disparities, behaviors, or predictive outcomes (French (2005); Bound et al. (2010); Hosseini et al. (2022, 2021); Danesh et al. (2024), and De Nardi et al. (2024)). While these measures have proven useful, such approaches are limited in isolating the specific dimensions of health that drive observed effects, because they compress heterogeneity into a single dimension.[3]

Our model fits among more recent multidimensional frameworks (e.g., Conti et al. (2010); Amengual et al. (2021); Wen (2022), and Cozzi et al. (2024); Darden (2022)) but advances this literature by modeling health as an evolving, richly structured latent process with explicitly dynamic interdependence across components. Our analysis has similarities to Blundell et al. (2023), who apply principal component analysis to summarize a similar set of HRS indices to jointly model the effects of self-reported and cognitive health. They treat health as exogenous, while we focus on the endogenous formation of health. Moreover, we are more granular in distinguishing the health components, and allow for mental health to drive effects on employment among other outcomes.

Second, this study addresses a notable gap in the economic literature concerning the endogenous formation of health. More similar to our framework are studies employing dynamic models of health, where health formation depends on individual decisions and state variables (e.g., Grossman (1972); Gilleskie (1998); Yogo (2016); Michaud and Wiczer (2018); Strulik (2022); Darden (2022)). We advance this line of research by allowing the influence of economic factors on health to vary between its components and the interdependencies across the components. For example, we model the response of mental health to the deterioration of physical health, thereby offering a more comprehensive analysis of the interrelationships within an individual's overall health. Additionally, by modeling the cross-dependencies among health components, our framework mitigates an omitted variable bias present when estimating the effect of one of the components in isolation.

A large body of research in health and retirement treats health formation as exogenous to individual decision making (French and Jones, 2017). This simplification is often justified on the grounds that, for older individuals, who generally have established health histories, long-standing habits, and broad access to public health insurance, health is largely predetermined. Moreover, studies that adopt this assumption often

[3]Univariate measures (e.g., self-reported health) have clear empirical content: they correlate with objective outcomes (such as mortality) and covary with economic behaviors (such as employment) in expected ways. They are also widely available in survey data and relatively less costly to implement in structural models (De Nardi et al., 2024).

report small behavioral effects (De Nardi et al., 2016). However, this assumption may be overly restrictive for certain aspects of health, particularly mental health, which might be more sensitive to changing economic conditions and individual choices. For instance, continued work beyond the traditional retirement age may impose psychic costs that deteriorate mental health.

Third, we provide interpretation of the mechanisms through which commonly used health measures and specific conditions affect economic outcomes. A large literature studies the economic consequences of health and health shocks at older ages using diverse indicators and diagnoses (For instance, Bartel and Taubman (1979); Kahn (1998); Moran et al. (2011); Stephens et al. (2018)). Our framework isolates and compares the contributions of distinct latent dimensions of health that underlie these measures. In particular, we map self-reported health and specific diagnoses (e.g., cancer, diabetes) onto the latent components, yielding intuition about which dimensions drive their estimated effects. This helps clarify how concrete health events translate into economic behavior through changes in the multidimensional latent health state.

Finally, this study significantly enhances the understanding of the economic determinants and implications of mental health. Previous research has examined the influence of health dynamics on economic inequalities over the life cycle; however, the endogenous nature of mental health and its interactions with other health dimensions are still not well understood. This work uniquely examines both the determinants and effects of mental health within a unified framework, highlighting its pivotal role in late-life economic disparities. Mental health is increasingly recognized as a critical factor that influences economic outcomes, particularly in the labor market (e.g., Jolivet and Postel-Vinay (2020) and Biasi et al. (2021)). Our analysis complements this perspective by investigating the economic environments that aggravate mental health challenges (e.g., Adhvaryu et al. (2019) and Frank and Glied (2023)). By explicitly modeling endogenous interactions between mental health, physical health, and economic circumstances, our study uncovers mechanisms that have been underexplored in the economics literature, shedding new light on how these dynamics jointly shape economic and health trajectories in later life.

The remainder of the paper is structured as follows. Section 2 details the empirical estimation and identification of the health formation process. Section 3 describes the data used for the analysis. Section 4 reviews the results and Section 5 concludes.

2 Estimating the Technology of Health Formation

This section describes the framework for estimating a process of health formation near the end of working life. The health formation process described below is consistent with a standard dynamic model of health and employment as outlined in Section D of the Appendix. At each age, t, health is represented as a vector of stocks, $x_t = (x_t^p, x_t^c, x_t^{mh})'$, where $\{p, c, mh\}$ corresponds to physic, cognitive, and mental health, respectively. A lower stock of health in any given dimension corresponds to "worse" health, for instance the presence of more severe physical limitations.

At first observation (t=0), individuals are endowed with an initial stock of health, $x_0 = (x_0^p, x_0^c, x_0^{mh})'$. Individuals are first observed at the age of fifty-five, and their health at that time is the product of all previous behavior, early health shocks, and economic factors.[4] Initial conditions depend on one's latent heterogeneity, $b_0 \in \mathbb{R}$, which summarizes the accumulation of factors that affect the formation of one's health from birth until the age at which they are first observed. Latent heterogeneity reflects the factors that have been shown to be important determinants in health formation throughout the life cycle, such as education, childhood health, and socioeconomic status (Currie and Moretti, 2003; Case et al., 2005; Currie, 2009; Conti et al., 2010; Currie et al., 2010; Case and Paxson, 2010; Lundborg et al., 2014; Almond et al., 2018; Adhvaryu et al., 2019; De Nardi et al., 2024).

The stock of each distinct health component evolves dynamically according to a specified production technology, which is a function of the previous period's health stocks, health inputs, and labor market decisions. With three dimensions of health, the system is described by state-space equations, where t represents the discrete time index for the periods we model the health dynamics, $t \in \{1, ..., T_i\}$, and T_i is the oldest age we observe an individual. The dynamic formation of each health component, x_t^k, is determined by

$$x_t^k = f_t(x_{t-1}, I_t, L_{t-1}, \eta_t^k). \tag{1}$$

The function f_t characterizes the health formation technology, revealing how the health of the previous period and the decisions affect the current health stock. For each $k \in \{p, c, mh\}$, x_t^k is determined by $x_{t-1} = (x_{t-1}^p, x_{t-1}^c, x_{t-1}^{mh})'$, capturing both self- and cross-dependencies in health formation across the latent health components. We assume that only health stocks from the preceding period contribute to the formation process, suggesting that prior shocks to any health component impact only the current period's health

[4] As such, the initial endowment of health is the result of an unobserved formation process that occurs before the individual is first observed in the data. This is the initial conditions problem as described by Heckman (1991) and Wooldridge (2005).

through the previous period.[5]

The transition process depends on a vector of observable health inputs/investments, $I_t \in \mathbb{R}^{K_I}$. Health inputs encompass both healthy behaviors, such as exercise, and unhealthy behaviors, such as smoking. Healthy behaviors represent deliberate choices made by individuals based on their current health status. We model health investment decisions as a function of latent health and latent heterogeneity

$$I_t = I(x_t, b_0, t, \epsilon_t^I), \tag{2}$$

Similarly, health transitions depend on the previous period's employment, L_{t-1}. On the one hand, employment can negatively impact health, especially when individuals are subjected to physical stressors, hazardous substances, or psychological pressures. On the other hand, the rate of health decline has been found to increase when exiting the labor force (Black et al., 2018). Unobserved determinants of health investments are captured by ϵ_t^I. We model employment as

$$L_t = L(x_t, b_0, t, W_t, \epsilon_t^L). \tag{3}$$

The value of employment depends on the stocks of latent health, x_t, age, and latent heterogeneity, b_0. Furthermore, additional controls, W_t, encompass observable yet unmodeled determinants of employment.

Finally, η_t^k, capture unobserved determinants of health production. In this model, an initial conditions problem arises as x_t depends on x_0 through successive substitution for x_{t-1} in Equation (1). Hence, there may be unobserved factors in η_t^k (i.e., relevant for determining x_t) that are correlated with x_{t-1} through x_0. Moreover, unobserved factors in η_t^k may be correlated with individuals' decisions: work (L_{t-1}) and health inputs (I_t).[6] Ignoring this dependence introduces an omitted variable problem and results in biased estimates of the parameters in Equation (1). To address this, we express the unobservable factors in equation (1) as

$$\eta_t^k = \gamma^k b_0 + u_t^k, \; k \in \{p, c, mh\}, \tag{4}$$

where $u_t^k = (u_t^p, u_t^c, u_t^{mh})$ are assumed to be mutually independent idiosyncratic shocks to health formation and are exogenous to all other variables in Equation (1). The idea is that b_0 summarizes the relevant factors that jointly affect initial conditions, individual's choices, and productivity of health formation, addressing

[5] That is, serial dependence in health formation follows a first-order Markov process.

[6] These factors may be preferences for health or the productivity of health, such as education, SES, and SES in childhood.

the omitted variables problem.[7]

Finally, we examine how latent health stocks shape mortality risk. Mortality is modeled through a hazard function that captures the probability an individual dies in period t ($d_{it} = 1$) conditional on survival up to $t-1$ ($d_{it-1} = 0$), as

$$H(d_t = 1 | d_{t-1} = 0, x_t, b_0, t). \tag{5}$$

Mortality is also of central policy relevance. It has direct welfare implications, underpins social insurance programs, and drives public expenditures through pensions, health care, and disability insurance. In dynamic models of health, mortality represents the absorbing state of health capital depreciation (à la Grossman (1972)), making it a natural outcome to study when modeling health trajectories. By linking mortality risk to specific latent components, our framework clarifies which dimensions of health drive mortality differences and, in turn, is useful to inform policies aimed at prevention, risk adjustment, and targeting.

2.1 Measurement System

The primary challenge in estimating Equation (1) arises from the latent nature of the underlying health components. Instead, we observe a set of imperfect measures of health. Our approach builds on a substantial body of literature related to the estimation of production functions for latent skills and utilizes methodologies to estimate dynamic factor models (Cunha et al., 2010, 2021; Del Bono et al., 2022; Agostinelli and Wiswall, 2025).

We assume that the measurement system has a dedicated factor structure so that, for each latent health component, we have a set of measures that are generated exclusively by that component. For each $k \in \{p, c, mh\}$ and period $t \in \{0, ..., T_i\}$, we observe M^k noisy measures, $z^k_{t,m}$, $m \in \{1, ..., M^k\}$, generated by

$$z^k_{t,m} = Z^k_m(x^k_t, v^k_{t,m}),$$

where $v^k_{t,m}$ is i.i.d. measurement error. Further, we assume the parameters of the measurement system are age-invariant (Agostinelli and Wiswall, 2025).[8] Our data encompass three distinct types of measurements:

[7] This identification argument follows from the solution to the theoretical life-cycle consumption and labor supply model described in Section D of the Appendix. This model allows for latent heterogeneity in health formation, preferences, and productivity, all of which may be correlated. The model illustrates how health formation depends on endogenous health investment and employment decisions operating through b_0.

[8] The health measures are explicitly designed to be administered consistently across survey waves, allowing researchers to track changes in health as individuals age. This design underlies the assumption that two individuals with the same underlying health vector, but differing in age, will on average record the same measured health. Similar assumptions appear in related

continuous, categorical, and binary.

1. For **continuous measures**, we adhere to the prevailing literature and assume a linear mapping. That is,

$$z^k_{t,m} = \mu^k_m + \lambda^k_m x^k_t + v^k_{t,m}.$$

The measurement parameters μ^k_m and λ^k_m indicate location and scale, respectively.

2. For categorical measures **(ordered)**, we assume they are generated by a latent index $z^{k*}_{t,m} = x^k_t + v^k_{t,m}$, where

$$z^k_{t,m} = j \text{ iff } \tau^k_{m,j-1} < z^{k*}_{t,m} \leq \tau^k_{m,j}$$

where $\tau^k_{m,0} = -\infty$, $\tau^k_{m,J} = \infty$, and the remaining thresholds on the interior are to be estimated.

3. For **binary measures**, we assume they are generated by

$$z^k_{t,m} = \begin{cases} 1 & \text{iff } z^{k*}_{t,m} = x^k_t + v^k_{t,m} > \tau^k_m, \\ 0 & \text{otherwise} \end{cases}$$

where $z^{k*}_{t,m}$ is a latent index and τ^k_m is the threshold to be estimated.

Initial heterogeneity: We have a set of observed measures, $z^{b_0}_m,\ m \in \{1, ..., M^{b_0}\}$, which we assume are exclusively generated by latent initial heterogeneity, b_0. These measures are based on individual characteristics prior to model entry, including education and childhood health.

Supplementary measures: In certain model specifications, we include measures that are generated from a combination of the full set of latent health components, encompassing both objective health conditions and self-reported health. These measures may fall into one of the three types described earlier. A comprehensive description of the complete set of health measures is provided in Section 3 below.

work. For example, Bound et al. (2010) estimate a multi-indicator latent health factor and test whether factor loadings vary with age. Their results cannot reject equality of loadings across ages 50–70, leading them to impose constant loadings and treat the resulting health index as age-invariant.

2.2 Implementation

In the current version of the paper, we specify a *linear law of motion* for each component of latent health. Each component evolves over time as a function of other health components, latent heterogeneity, and observed factors as

$$x_{t+1}^k = a_0^k + a_1^k x_t^p + a_2^k x_t^c + a_3^k x_t^{mh} + a_4^k b_0 + c_I^k I_t + c_e^k L_t + \nu_t^k, \tag{6}$$

$$\nu_t^k \sim iid\ \mathcal{N}(0, \sigma_{\nu^k}^2), \quad \text{for all } k \in \{p, c, mh\}. \tag{7}$$

The vector $c_I^k \in \mathbb{R}^{K_I}$ captures the effect of health inputs, I_t, and c_e^k captures the effect of employment status on health formation.[9]

We use a nonstructural approximating model for health input and employment policy functions. Health input decisions, I_t, are choices made conditional on health. The decision regarding health input $j \in K_I$ is modeled as

$$\begin{aligned} I_{jt}^* &= \beta_0^{I_j} + \beta_1^{I_j} x_t^p + \beta_2^{I_j} x_t^c + \beta_3^{I_j} x_t^{mh} + \beta_4^{I_j} b_0 + \beta_5^{I_j} t + \nu_t^{I_j}, \\ \nu_t^{I_j} &\sim iid\ \mathcal{N}(0, 1), \\ I_{jt} &= \mathbb{1}(I_{jt}^* \geq 0). \end{aligned} \tag{8}$$

The latent value of healthy behavior $j \in K_I$ is I_{jt}^*.

In a similar manner, we model the employment decision, L_t as

$$\begin{aligned} L_t^* &= \beta_0^L + \beta_1^L x_t^p + \beta_2^L x_t^c + \beta_3^L x_t^{mh} + \beta_4^L b_0 + \beta_5^L W_t + \beta_6^L t + \nu_t^e, \\ \nu_t^e &\sim iid \mathcal{N}(0, 1), \\ L_t &= \mathbb{1}(L_t^* \geq 0), \end{aligned} \tag{9}$$

where L_t^* reflects the latent value of employment. The control variables, W_t, include marital status, total work experience at initial observation, and overall non-housing wealth as in Blundell et al. (2023). Further, we assume that ν_t^k, $\nu_t^{I_j}$, and ν_t^e are mutually independent for all $k \in \mathbb{R}^3$, $j \in K_I$ and over time.

For individuals that have survived to $t-1$, the probability they die in period t ($d_t = 1$) is determined

[9] The linear specification can be interpreted as the logarithm of a Cobb–Douglas production function. In particular, let $x_t = ln(\tilde{x}_t)$, where $\tilde{x}_t$ denotes the true latent stock. In this formulation, the observed measures are generated by the log of the latent stocks, and the distributional assumptions are imposed on the logarithmic rather than the level representation of these stocks.

by a latent variable, d_t^*, where

$$
\begin{aligned}
d_t^* &= \beta_0^s + \beta_1^s x_t^p + \beta_2^s x_t^c + \beta_3^s x_t^{mh} + \beta_4^s b_0 + \beta_5^s t + \epsilon_t^s, \qquad (10)\\
\epsilon_t^s &\sim iid\, \mathcal{N}(0,1),\\
d_{it} &= \mathbb{1}\{d_t^* \geq 0\}.
\end{aligned}
$$

Finally, we estimate the model using an unbalanced panel, where individuals, and attrition is prevalent in the sample. However, we model attrition from the sample as endogenous to ones health.[10] We denote the value of attriting from the sample in period t as a_t^*, then we model the decision to attrit ($a_{it} = 1$) as,

$$
\begin{aligned}
a_t^* &= \beta_0^s + \beta_1^a x_t^p + \beta_2^a x_t^c + \beta_3^a x_t^{mh} + \beta_4^a b_0 + \beta_5^a t + \epsilon_t^a, \qquad (11)\\
\epsilon_t^a &\sim iid\, \mathcal{N}(0,1),\\
a_{it} &= \mathbb{1}\{a_t^* \geq 0\}.
\end{aligned}
$$

Moreover, we accommodate temporary gaps and right-censoring in outcomes, treating all censoring as independent of the underlying process (non-informative)..

As individuals are first observed at age 55 and substantial health dynamics occur prior to model entry, we allow initial conditions to be jointly distributed. Let $w_0 = (x_0^p, x_0^c, x_0^{mh}, b_0)'$ denote the vector of initial conditions. We assume $w_0 \sim N(0, \Sigma_w)$, and the diagonal of the covariance matrix, Σ_w, is normalized to unity,

$$
\Sigma_w = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{1b} \\ \rho_{12} & 1 & \rho_{23} & \rho_{2b} \\ \rho_{13} & \rho_{23} & 1 & \rho_{3b} \\ \rho_{1b} & \rho_{2b} & \rho_{3b} & 1 \end{bmatrix}.
$$

The vector of correlation coefficients, $\rho = (\rho_{12}, \rho_{13}, \rho_{23}, \rho_{1b}, \rho_{2b}, \rho_{3b})'$ characterizes the dependence among initial health stocks and latent heterogeneity. This flexible structure captures the complex joint distribution of health upon entry into the sample. The assumed distribution on w_0 and Equation (6) imply that, for each $t \in \{1, .., T\}$, the distribution of the latent health vector, x_t, is

$$
x_t \sim \mathcal{N}(\mu_{x_t}, \Sigma_{x_t}).
$$

[10] For analysis on attrition in the HRS, see Banks et al. (2011)

The off-diagonal elements of Σ_{x_t}, capturing the covariance structure among the latent health components, manifest from ρ and the parameters of Equation (6).

2.3 Identification of Linear Formation Technology

Identifying and estimating the technology function is challenging, as both inputs and outputs are only observed through proxies. Inputs may be endogenous, and unobserved components in the input equations may be correlated with unobservables in the technology function. To show identification of the model, we impose additional restrictions on the measurement system that follow common practice in the related literature (Cunha et al., 2021).

1. $v^k_{t,m}$ are mean zero for all $k \in \{p, c, mh, b\}$, $t \in \{1, .., T_i\}$, and $m \in \{1, ..., M^k\}$. The variance of measurement errors for discrete and categorical measures is normalized to one.

2. $v^k_{t,m}$ is independent of $(x^p_\tau, x^c_\tau, x^{mh}_\tau, b_0)$ for all $t, \tau \in \{1, .., T_i\}$; $m \in \{1, ..., M^k\}$; and $k \in \{p, c, mh, b\}$.

3. $v^k_{t,m}$ is independent of $v^l_{\tau,n}$ for all $t, \tau \in \{1, .., T_i\}$, $t \neq \tau; m \in \{1, ..., M^k\}$; and $n \in \{1, ..., M^l\}$ where $m \neq n$, $k \in \{p, c, mh, b\}$, and $k \neq l$.

4. $v^k_{t,m}$ and $v^k_{t,n}$ are allowed to freely covary for $m, n \in \{1, ..., M^k\}$.

Under these assumptions, three time periods and a single measure generated by each latent health component is sufficient to identify the model's parameters. A step-by-step description of identification in such a scenario is provided in Section A of the Appendix. Here we describe the intuition for identification with many dedicated measures of each latent health component, as is the case in this paper.

First, we consider identification of the correlation of initial endowments, ρ, and the measurement system parameters. The identification intuition is straightforward with age-invariance of the measurement system parameters and the normalization on the distribution of the latent endowments, w_0. The threshold parameters of binary measures (μ^k_m) simply calibrate to match $Pr(z^k_{t,m} = 1)$ in the observed data. A similar argument holds for the categorical measures, except the thresholds ($\tau^k_{j,m}$) calibrate to fit the density for each category j, $P(z^k_{t,m} = j)$, for all $j \in \{1, ..., J\}$. The normalization on w_0 also helps to pin down both the location and scale of the latent factors, which identifies the intercept and scale parameters of the continuous measures. First, the location parameters, μ^k_m, are identified by $E(z^k_{m,1})$. The scales of continuous measures, λ^k_m, and the variance of measurement errors, ν^k_m, are separately identified using the covariance structure across continuous measures and periods, as in Cunha and Heckman (2008). With the measurement system parameters pinned down, the correlation coefficients for initial endowments, ρ, are backed out by covariances observed measures in the first period, $z^{k_1}_{0,m}$ and $z^{k_2}_{0,m}$ for $k_1, k_2 \in \{c, mh, p, b\}$ and $k_1 \neq k_2$.[11]

The latent health components evolve according to Equation 6. If the latent states were observed, their law of motion could be estimated directly from serial covariances of the states and controls (employment, health behaviors). But since the states are unobserved, we must infer their covariance structure indirectly through the measurement system. With the parameters of the measurement system identified, the serial variance-covariance structure across all latent factors and all time periods is identified from the serial variance-covariance structure of observed measures. The key insight is that the covariance structure of the observed measures (both discrete and continuous) can be written in terms of the underlying latent covariances and the measurement parameters. With only continuous outcomes, this mapping is direct (for example, see Cunha and Heckman (2008), Agostinelli and Wiswall (2025)). With discrete measures, the mapping involves more algebra, but conceptually the idea is the same.

That is, the estimators for the parameters of Equation 6 are functions of the serial variance-covariances of the latent factors and the covariances the observed variables, L_{t-1} and I_{t-1}, and the latent factors. Covariance of the observed measures with the observed variables provides the covariance of latent components with the observed measures and the measurement parameters. Hence, the parameters of Equation 6 are identified. A similar argument holds for the parameters of Equations 8, 9, and 10. Finally, we can back out the variance of production shocks via the variance-covariance matrix period $t = 2$ as in STEP X of Appendix Section C.

2.4 Estimation

We estimate the model using maximum simulated likelihood, which is particularly useful when the likelihood involves high-dimensional integrals over latent states that cannot be computed in closed form. In our model the observed-data likelihood involves integrating out the latent states $\{x^1_t, x^2_t, x^3_t,\}_{t=1\ldots T}$ and the latent heterogeneity, b_0. The approximated likelihood contribution for each individuals is constructed using the particle filter. This simulation approach replaces this difficult integral with an average over simulated

[11] For instance, consider k_1, k_2 to be two latent random variables that are assumed normally distributed and observed by measures $z^{k_1}_1$ and $z^{k_2}_2$. The correlation coefficient for k_1 and k_2 that are both observed by continuous measures, we have $\rho_{k_1,k_2} = cov(z^{k_1}_{1,m}, z^{k_2}_{1,m_2})/\lambda^{k_1}_{m_1}\lambda^{k_2}_{m_2}$. For k_1 and k_2 that are both observed by binary measures we have $\rho_{k_1,k_2} = \rho^{12}s^1s^2$, where ρ^{12} is the tetrachoric correlation of the standardized index of these measures and identified directly from the data, as in step 1.a in Appendix Section A, and $s^k = \sqrt{var(k) + var(v_0)} = \sqrt{2}$ in the first period. Last, if k_1 is observed by a continuous measure and k_2 is observed by a binary measure we have $\rho = \frac{s^j_s}{\phi(\hat{d}^j)} cov(z^i, z^j)/\lambda^{k_1}_{m_1}$, where $\phi(\cdot)$ is the PDF of a standard normal, and $d^j = \frac{\tau^j - E(k^j)}{s^j}$ is the standardized threshold.

draws from the joint distribution of the latent states. We maximize the approximated log-likelihood with respect to parameters θ using a combination of Nealder-Meade and the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm. A detailed description of the likelihood is located in Section B of the Appendix. Further details on the simulation and estimation steps are located in Section C of the Appendix.

3 Health and Retirement Study

We estimate the technology of health formation using the Health and Retirement Study (HRS). HRS is a biennial panel survey of noninstutionalized individuals and their spouses living in the United States. HRS includes an abundant set of measures of one's health, along with variables related to income, demographics, and other economic conditions. The HRS began in 1992, and respondents enter the survey between the ages 51-61. We focus on white male respondents aged 55 or 56 that are observed for at least two total periods. We exclusively use waves 3-14 of the survey, as there are nontrivial differences in the questionnaire starting in the third wave. Our sample is an unbalanced panel of 3,429 individuals (23,331 total observations).

Table 1 shows the number of observations at for each age grouping. We can see that we lose a considerable number of individuals from the sample as individuals age. The dominant factor causing this loss is due to right censoring. Next, we detail the various sets of variables used to measure health and to estimate the model.

Table 1: Sample Counts

Age Group	Observed	Dead	Attrition	Censor
1	3,378	0	0	51
2	3,321	3	0	105
3	3,117	17	132	163
4	2,789	47	237	356
5	2,476	70	356	527
6	2,144	102	462	721
7	1,815	139	582	893
8	1,488	193	691	1057
9	1,136	233	810	1250
10	848	266	893	1422
11	568	295	981	1585
12	251	316	1055	1807
Totals:	23,331	1,681	6199	9937

Note: Age groups correspond to two year bins starting from age 55/56, 57/58, 59/60, ...

Mental Health

Mental health is measured using a set of questions derived from the Center for Epidemiological Studies Depression (CESD) scale. This scale is based on the sum of eight indicators. Six of these indicators are "negative," where respondents report "yes" or "no" to experiencing the following sentiments all or most of the time: depression, everything is an effort, restless sleep, feeling alone, feeling sad, and not being able to get going. The remaining two indicators are "positive," where respondents report "yes" or "no" to experiencing the following sentiments all or most of the time: feeling happy and enjoying life. The commonly applied CESD scale is based on aggregating these responses, giving a score that ranges from 0 to 8, with higher scores indicating worse mental health. In our framework, we treat each indicator separately.

Cognitive Functioning

Cognitive functioning is measured using six cognitive health indices that are consistently collected in all survey waves. These measures correspond to scores on tests administered to survey respondents and self-reported functional limitations.

First, we use two measures of fluid intelligence (the capacity to think logically and solve problems in novel situations, independent of acquired knowledge). Prior research suggests that fluid intelligence is strongly associated with labor market outcomes (e.g., Heineck and Anger (2010)).[12] To measure fluid intelligence, respondents complete two standardized word recall tests: immediate and delayed word recall. The immediate word recall score captures the number of words out of ten that are correctly recalled immediately after presentation, while the delayed word recall score reflects the number of words recalled correctly following a delay of approximately five minutes.

Second, we incorporate two measures derived from tests of basic arithmetic ability. The first is the serial sevens test, in which respondents are asked to sequentially subtract 7 from 100 across five trials. Scores range from 0 to 5, based on the number of correct subtractions. The second measure involves backward counting, where respondents are instructed to count backward from 20 and from 86 for ten consecutive numbers. Scoring is based on performance: a score of 2 is assigned if the respondent completes the task correctly on the first attempt, 1 if correct on the second attempt, and 0 if unsuccessful in both attempts.

Finally, we include three subjective measures of functional limitations related to cognition. Respondents are asked whether they have difficulty using a map and whether they have difficulty managing money.

[12] Fluid intelligence is distinguished from crystallized intelligence, which relies more on retrieving information from long-term memory.

Additionally, they self-report their memory on a five-point scale. These questions reflect everyday cognitive functioning limitations that may not be fully captured by formal test scores.

Physical Functioning

Physical functioning is assessed using a set of functional limitation indices, each constructed by aggregating responses to questions about difficulty performing specific everyday tasks. For each task, respondents report whether they experience difficulty that is expected to last at least three months. Responses are coded as 1 for difficulty and 0 for no difficulty. Each functional limitation index represents the sum of difficulties within a group of tasks that involve similar types of physical functioning, with higher scores indicating greater functional limitation.

Three indices are used to capture different areas of physical functioning: mobility, large muscle function, and activities of daily living (ADLs). The mobility index includes five tasks: walking several blocks, walking one block, walking across a room, climbing several flights of stairs, and climbing one flight of stairs. The large muscle index includes four tasks: sitting for two hours, getting up from a chair, stooping, kneeling or crouching, and pushing or pulling a large object. The ADL index includes five tasks: bathing, eating, dressing, getting out of bed, and using the toilet.

Employment and Healthy Behaviors

We model the health transition process as a function of employment status. An indicator for whether an individual is employed is derived from their reported labor force status, with individuals flagged as employed if they report working either part time or full time.[13]

HRS respondents are surveyed on a set of questions regarding health-related behaviors. Respondents are asked how often they engage in vigorous physical activity—such as aerobics, running, swimming, or bicycling. Following common practice in the literature, individuals are flagged as engaging in regular exercise if they report participating in such activities three or more times per week.

In addition, respondents are asked whether they consume alcohol, and if so, how frequently and how many drinks they typically consume. They are also asked whether they smoke cigarettes. To capture these behaviors as inputs in the health formation process, we construct indicator variables for engaging in regular

[13]A respondent can give evidence of working, being retired, and disability alone or in combination with other statuses. If the respondent is working full-time or part-time and there is no mention of retirement in the previous two years, they are considered employed. Working 35+ hours per week, 36+ weeks per year is considered full-time. Less than this is considered part-time. The hours and weeks from both the main and second job are considered in determining whether the respondent is working full-time or part-time.

vigorous physical activity, drinking alcohol, and smoking status.

Measures of Latent Heterogeneity

Respondents enter the sample at age 55, meaning their initial observed health reflects a lifetime of prior choices and experiences. These accumulated factors may influence both the formation of health and other individual decisions, such as engagement in healthy behaviors or employment. Importantly, unobserved characteristics—such as early-life socioeconomic status—may be jointly correlated with health outcomes and behavioral choices. For instance, individuals from disadvantaged backgrounds may have received fewer investments in health-promoting behaviors (e.g., regular exercise) or in the development of skills that enhance workplace productivity.

To account for this, we construct an initial heterogeneity variable based on information collected prior to age 55. This variable summarizes pre-existing individual differences at the start of the observation window that are plausibly related to both subsequent health and employment trajectories (Currie and Moretti, 2003; Case et al., 2005; Currie, 2009; Conti et al., 2010; Currie et al., 2010; Case and Paxson, 2010; Lundborg et al., 2014; Almond et al., 2018; Adhvaryu et al., 2019). The pre-55 variables are used as observed measures in a factor analytic framework to recover the initial heterogeneity, b_0.

First, we include the respondent's own education level and his mother's education level. These variables relate to the efficient producer hypothesis, which posits that more educated individuals are more efficient at producing and maintaining health (Grossman, 1972). Second, we incorporate a measure of childhood socioeconomic status (SES) based on the respondent's self-reported assessment of their family's financial situation during childhood (categorized as "well off," "about average," or "poor"). This variable reflects early-life resource constraints and social environment, consistent with the early-life adversity or health capital frameworks, which emphasize the long-term effects of early deprivation.

We also include the respondent's self-rated health at age 16 (reported on a five-point scale: excellent, very good, good, fair, poor), which captures early-life health endowments that may shape later-life health outcomes and labor market capacity.

Finally, we include indicators of early-life risk exposure, such as whether the respondent ever smoked prior to age 50, and whether either parent used alcohol or drugs to an extent that caused problems in the family. These variables proxy for exposure to adverse family environments and behavioral risk factors that may have long-run impacts on both health and socioeconomic outcomes.

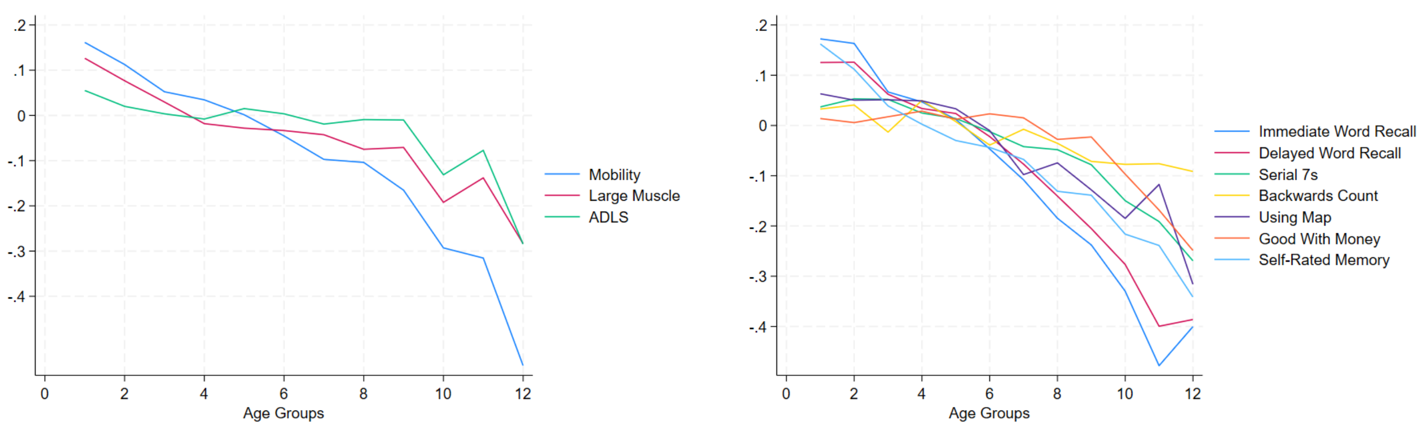
Additional Health Measures

Our model incorporates a set of additional health measures as functions of all the latent health components. Including these measures enables us to examine how each latent health dimension relates to commonly used indicators of health status. First, we include self-reported health, a widely used summary measure in empirical research on health and its economic consequences. Self-reported health is a simple and readily available measure of health that has been shown to be a strong predictor of mortality (Idler and Benyamini, 1997), labor supply (Bound, 1989; Stern, 1989), as well as exhibiting a strong correlation with other health measures.[14]. In our analysis, we use respondents' self-rated health, measured on a five-point categorical scale where "5" indicates excellent health.

Second, we incorporate a set of objective health indicators, capturing whether the respondent has ever been diagnosed with high blood pressure, diabetes, cancer, lung disease, heart disease, stroke, psychiatric problems, or arthritis. Linking these diagnoses to the latent health components reveals how specific conditions manifest in the latent structure. The estimated factor loadings on these objective conditions indicate how each latent health component maps onto particular diagnoses.

3.1 Descriptive Statistics of Health Measures

Figure 1: Mean of Standardized Health Measures by Age

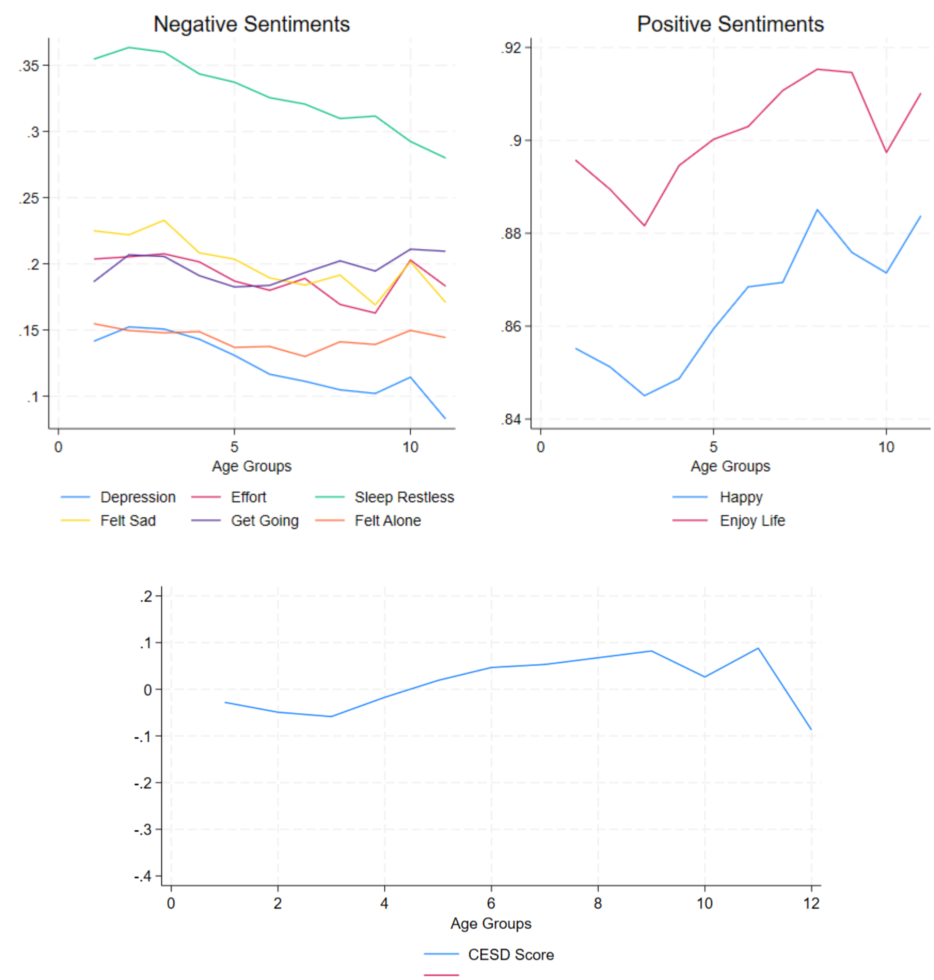


Note: ages are gouped into two-year bins. That is, age group 1 is 55-56 years old, age group 2 is 57-58 years old, etc..

[14]See White (2023) and Blundell et al. (2023) for further discussion on the use of self-reported health in economics

This section describes the dynamics of the observed health measures and their joint correlation structure. Figure 1 displays the averages of physical health measures (left), cognitive health measures (right). To facilitate comparison, each measure is standardized and signed so that higher values indicate "better" health. The x-axis shows distinct two-year age groupings. Figure 1 illustrates that the mean dynamic profile of physical and cognitive health, which are more typically considered in related studies (Poterba et al. (2017); Hosseini et al. (2021); Capatina and Keane (2023)), deteriorate with age. This deterioration has been used to explain health consequences for labor supply, medical expenditures, early retirement, and other variables.

Figure 2: Mean of Mental Health Measures by Age



Note: ages are grouped into two-year bins. That is, age group 1 is 55-56 years old, age group 2 is 57-58 years old, etc...

Next, Figure 2 shows the average responses by age groupings for each separate mental health index with negative indices on the left and positive indices on the right. The bottom graph shows the age-trend in the overall CESD score, which is constructed by aggregating over all CESD questions to construct the CESD index. The figures report the same standardizing and signing of the CESD index. In contrast to the age profiles for physical and cognitive health, mental health has a concave shape that improves then gradual declines after age group 67/68. The rise in mental health prior to retirement is consistent with evidence showing the life-cycle profile of mental health is U-shaped (Dijk and Mierau, 2023). However, the change in direction could be caused by an accumulation of health deficits. An important empirical question arises regarding the extent to which this improvement in mental health is influenced by work and the degree to which mental health deterioration exacerbates the effects of declining health in other dimensions on economic decisions.

Figure 3: Correlation of Physical and Cognitive Measures with Mental Health by Age

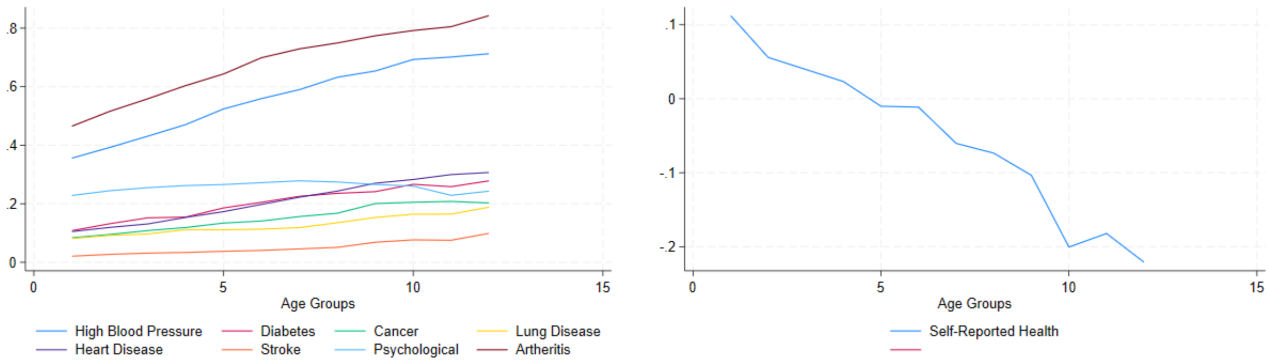
Note: ages are gouped into two-year bins. That is, age group 1 is 55-56 years old, age group 2 is 57-58 years old, etc..

Figure 3 illustrates the relationship between physical and cognitive measures and mental health over time. Mental health exhibits a positive correlation with both physical and cognitive health. Considering the varying age profiles of physical, cognitive, and mental health, it can be inferred that a subset of individuals experiencing declines in physical or cognitive health may also suffer from deteriorating mental health.

Next, Figure 4 shows the average probability of self-reporting an objective health condition (left), and the evolution of standardized self-reported health (right). Both measures, frequently examined in related literature, demonstrate a similar age-trend as physical and cognitive measures. The probability of reporting an objective health condition rises with age across all categories. Second, self-reported health tracks physical measures closely. Both figures are consistent with the accumulated decline in health as individuals age into and beyond retirement.

The final set of descriptives examine the relationship between the health measures, employment, and

Figure 4: Mean of Objective Health Measures and Self-Reported Health by Age



Note: ages are gouped into two-year bins. That is, age group 1 is 55-56 year olds, age group 2 is 57-58 year olds, etc.

Table 2: Descriptive correlations of health measures and labor outcomes

	CESD	Self-Rated Memory	Immediate Word Recall	Delayed Word Recall	Mobility	Large Muscle	ADLS
Employment	0.125	0.162	0.158	0.138	0.232	0.206	0.165
Vigorous Activity	0.057	0.051	0.045	0.037	0.092	0.091	0.041
Smokes	-0.124	-0.051	-0.043	-0.046	-0.087	-0.067	-0.062
Drinks	0.150	0.137	0.163	0.155	0.234	0.209	0.130
Mortality	-0.059	-0.040	-0.057	-0.055	-0.123	-0.067	-0.109

Note: Correlations are calculated using the sample pooled across all age groups.

healthy behaviors. Table 2 presents correlations over age between the health measures and employment (top row) and three health-related behaviors. Each measure is positively correlated with employment, with the strongest associations arising for the physical health indicators, consistent with evidence linking health to labor market outcomes (Bound, 1989; Currie and Madrian, 1999). Vigorous exercise is also positively correlated with the health measures, whereas smoking is negatively correlated, in line with prior work on risky health behaviors (Hai and Heckman, 2022). Interestingly, drinking shows a positive correlation with the health measures, consistent with earlier findings that moderate alcohol use can be positively associated with health and labor market outcomes (Mullahy and Sindelar, 1996).

4 Estimation Results

Table 3: Model Estimates: Covariance Matrix of Initial Endowments

	x_0^c	x_0^{mh}	x_0^p	b_0
x_0^c	1.000	0.336***	0.408***	0.678***
		(0.026)	(0.035)	(0.047)
x_0^{mh}		1.000	0.654***	0.393***
			(0.023)	(0.023)
x_0^p			1.000	0.458***
				(0.022)
b_0				1.000

Note: Standard errors presented in parentheses below point estimates.

This section reviews the estimated parameters of the health formation process. Table 3 presents the covariance matrix of the latent health components at $t = 0$, (x_0^c, x_0^{mh}, x_0^p), along with individual heterogeneity, b_0. The endowments of the three health components have positive correlations, with the strongest association between physical and mental health. Moreover, initial heterogeneity positively correlates with each health component, indicating that individuals with higher endowments of b_0 tend to exhibit better health outcomes.

Table 4: Conditional Mean of Latent Heterogeneity, b_0

	Education		Mother's Education		SES
None	-0.92	None	-0.55	Poor	-0.76
Above Bachelors	1.12	College +	1.13	Rich	1.31
	Self Reported Health in Childhood		Ever-Smoked		Parents Drink/ Drug
Poor	-1.53	No	0.67	No	0.30
Excellent	0.55	Yes	-0.48	Yes	-0.94

Table 4 presents the mean of b_0 conditional on a subset of realizations of the measures used to recover its distribution, offering intuition to the sorts of characteristics this variable captures. Initial heterogeneity is measured by retrospective characteristics of the individual and their parents. First, b_0 has a positive correlation with both the individual's and their mother's education levels, as well as with childhood socio-economic status and self-reported health during childhood. Additionally, b_0 is lower among individuals who have smoked before the age of 55 and among those whose parents engaged in excessive drinking or drug use.

These patterns support the efficient-producer hypothesis (Grossman, 1972), which asserts that early investments in human capital, encompassing both personal and parental investments, improve the efficiency of health production in later life. Furthermore, they align with findings on the long-term health consequences related to adverse childhood environments and risky adolescent behaviors (Case and Paxson, 2010; Conti et al., 2010).

Table 5: Model Estimates: Health Formation Process

	x_t^c	x_t^{mh}	x_t^p
Intercept	-0.200***	-0.052***	-0.202***
	(0.005)	(0.007)	(0.005)
x_{t-1}^c	0.294***	0.088***	0.080***
	(0.005)	(0.006)	(0.006)
x_{t-1}^{mh}	0.008***	0.225***	0.269***
	(0.001)	(0.008)	(0.007)
x_{t-1}^p	0.039***	0.287***	0.522***
	(0.003)	(0.006)	(0.008)
b_0	0.198***	0.188***	0.134***
	(0.005)	(0.013)	(0.011)
L_{t-1}	0.120***	0.028***	0.274***
	(0.008)	(0.004)	(0.011)
Drink	0.199***	0.230***	0.022***
	(0.009)	(0.011)	(0.007)
Smoke	-0.173***	-0.134***	-0.025
	(0.011)	(0.010)	(0.017)
Exercise	-0.111	-0.107	0.397***
	(0.130)	(0.170)	(0.010)

Note: Standard errors presented in parentheses below point estimates.

Table 5 presents the estimates of the transition process. The columns distinguish estimates for cognitive, mental-health, and physical-health, respectively. The first row details the time trend, indicating that cognitive and physical health decline with age, while the rate of decline in mental health is significantly smaller in magnitude. Rows 2–4 capture both own-lag and cross-lag effects. As indicated on the diagonal, each component displays persistence, with the most pronounced persistence observed in physical health.

Furthermore, physical health has the largest cross-effect on mental health, highlighting its crucial role as a precondition for mental well-being. Mental health is also an important determinant of physical health, which suggests the increased psychological distress has physiological effects on physical health. The effect of b_0 on health formation is displayed in Row 5. Latent heterogeneity positively affects health formation for all components. This finding indicates that certain demographic groups, such as individuals with higher education, improved socioeconomic status during upbringing, or better self-reported health in childhood, experience a slower rate of health decline in older age.

Next, we consider the effect of the lagged employment on health formation. Having worked in the previous period raises subsequent health in all three domains. This is consistent with the finding in related studies that health is negatively affected by labor market exit into retirement or disability insurance (Fitzpatrick and Moore, 2018; Black et al., 2018). Employment has the greatest impact on the preservation of physical health, and its effect on mental health is an order of magnitude smaller.

Rows 6-8 provide estimates of the effects of healthy behaviors. Smoking is found to significantly diminish all health stocks, with the most substantial impact on cognitive health. Conversely, vigorous exercise enhances physical health, but is negatively related to cognitive and mental health. Interestingly, alcohol consumption appears to have a positive effect on health formation. This phenomenon may stem from the use of a binary indicator for "any drinking." Future work will consider both the extensive and intensive margins of drinking.

4.0.1 Effects of health on employment, mortality, and healthy behaviors.

Despite the extensive literature on the effects of health on employment, a lack of consensus on the magnitude of these effects and the specific dimensions of health most predictive of particular outcomes remains. This lack of consensus may, in part, arise from the variety of empirical methodologies and datasets utilized to evaluate these effects. Table 6 presents estimates of employment determinants, illustrating the relative impacts of various health components. Improved health correlates positively with an increased probability of employment across all components. Importantly, physical health exerts the largest significant effect on employment, while the effects of cognitive and mental health are smaller, although still statistically significant. These estimates corroborate the findings in Blundell et al. (2023), who show that cognition has modest effects on employment when also accounting for health. These findings lend support to the use of physical health measures as the primary indicators for analyzing of the effects of health on employment. However, Table 5 indicates mental health to be an important determinant of physical health. Ignoring the relationship between health components can result in potentially bias estimates of health's employment effects.

Table 6: Model Estimates: Employment, Mortality, and Healthy Behaviors

	L_t	H_t	Drink	Smoke	Vigorous Activity
Intercept	-0.022***	-2.484***	-0.417***	-0.913***	-1.102***
	(0.002)	(0.042)	(0.015)	(0.022)	(0.020)
x_t^c	0.015***	0.431***	0.172***	-0.025***	0.049***
	(0.001)	(0.034)	(0.011)	(0.003)	(0.005)
x_t^{mh}	0.035***	-0.545***	0.109***	-0.194***	-0.140***
	(0.005)	(0.036)	(0.015)	(0.010)	(0.015)
x_t^p	0.191***	-0.145***	0.204***	-0.006	0.151***
	(0.01)	(0.014)	(0.012)	(0.004)	(0.013)
b_0	0.135***	0.008***	-0.010***	-0.040***	-0.053***
	(0.01)	(0.003)	(0.000)	(0.007)	(0.005)
Age	-0.292***	0.004***			
	(0.002)	(0.001)			
wealth	-0.025				
	(1.481)				
experience	0.048				
	(0.117)				
married	-0.014				
	(0.365)				

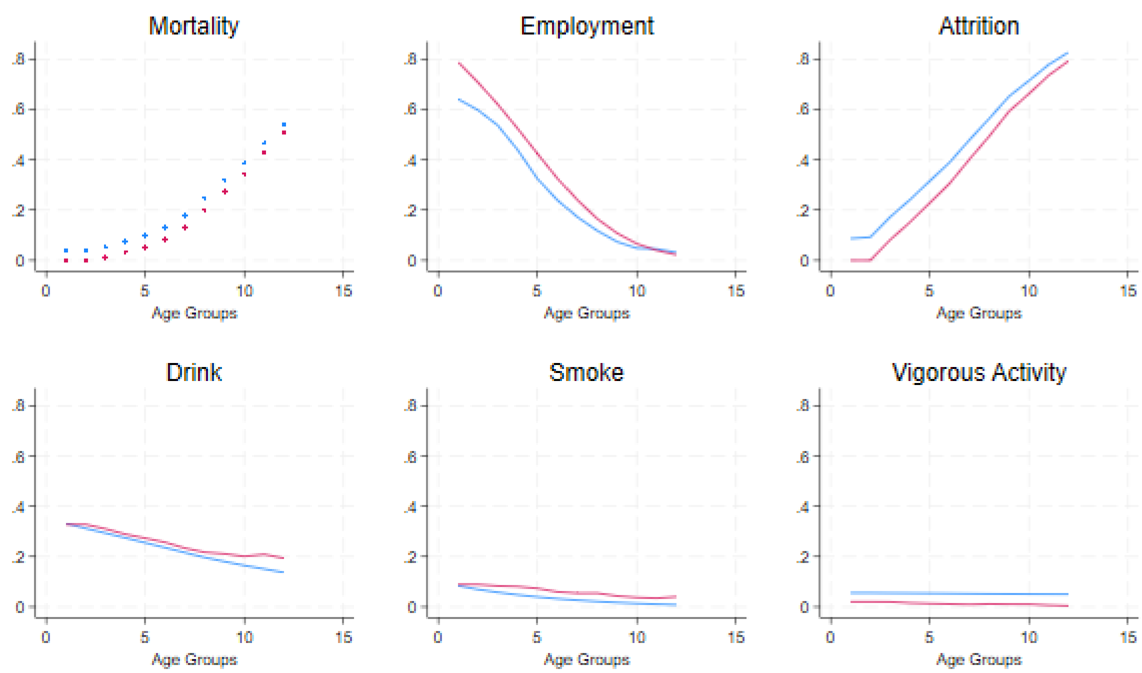
Note: Standard errors presented in parentheses below point estimates.

Next, we examine mortality, a crucial element in an individual's decision-making process, as it determines the expected decision horizon when making choices. For example, a decline in health that raises mortality risk may prompt individuals to reassess their employment decisions in light of a shortened expected lifespan. The second column of Table 6 illustrates that the probability of mortality increases with age. Additionally, individuals in poorer physical and mental health exhibit a higher likelihood of mortality. Surprisingly, cognitive health positively affects the likelihood of mortality.

Finally, we investigate the predictive relationship between latent health components and individual health behaviors. Individuals exhibiting good health across all latent dimensions are more inclined to engage in moderate drinking and less inclined to smoke. Better physical and cognitive health increases the likelihood of participation in vigorous exercise. Interestingly, mental health negatively affects the likelihood of engaging in vigorous physical activity.

4.1 Fit of the Empirical Model

Figure 5: Fit of Empirical Model



Note: ages are grouped into two-year bins. That is, age group 1 is 55-56 years old, age group 2 is 57-58 years old, etc. Red lines correspond to predictions from the simulated model and blue lines correspond to the observed data.

Figure 5 shows how well the predicted behavior simulated from the empirical model fits the average behavior and outcomes observed in the data over age groups. The red lines correspond to the simulated data and the blue lines correspond to the observed data. The model over predicts employment at younger ages and under predicts attrition. Moreover, the models over predicts drinking at older age groups. Overall, we see that the model fits the overall age-trajectories of the observed data fairly well.

4.2 Additional Results: Objective Health Measures and Self-Reported Health

The final set of results examine a version of the model that includes additional objective health conditions, such as binary measures indicating if the respondent has cancer or diabetes, and a respondent's self-assessment of their overall health, measured on a scale of five. Each additional health measure is modeled as a function of all latent health stocks. The estimated parameters reveal the extent to which each latent health component relates to both objective health conditions and self-reported health. In other words, for objective health conditions, we estimate how each condition manifests itself in terms of the latent stocks. For self-reported health, we estimate how each latent health component determines an individual's self-perception of their overall health.

Table 7: Objective Health Conditions

	High Blood Pressure	Diabetes	Cancer	Lung Disease	Heart Problem	Stroke	Psychological Problem	Arthritis
Intercept	-0.423***	-1.379***	-1.414***	-1.686***	-1.460***	-2.410***	-0.788***	-0.013
	(0.017)	(0.022)	(0.021)	(0.040)	(0.025)	(0.051)	(0.020)	(0.024)
x^c	-0.134***	-0.365***	0.021	-0.147**	-0.011	-0.198***	-0.087**	0.257***
	(0.033)	(0.049)	(0.04)	(0.069)	(0.043)	(0.079)	(0.041)	(0.052)
x^{mh}	-0.130***	-0.313***	0.000	-0.207***	-0.156***	-0.237**	-0.484***	0.193***
	(0.051)	(0.055)	(0.087)	(0.059)	(0.05)	(0.102)	(0.050)	(0.042)
x^p	-0.405***	-0.235***	-0.131	-0.607***	-0.514***	-0.423***	-0.509***	-1.160***
	(0.049)	(0.046)	(0.099)	(0.051)	(0.046)	(0.068)	(0.045)	(0.029)
Age	0.077***	0.054***	0.053***	0.033***	0.069***	0.062***	-0.014***	0.061***
	(0.003)	(0.004)	(0.004)	(0.006)	(0.004)	(0.007)	(0.004)	(0.004)

Note: Standard errors presented in parentheses below point estimates.

Table 7 presents the estimated loadings for the objective health measures. Each dependent variable is coded as one if the respondent reports receiving a doctor's diagnosis of the condition, and zero otherwise. The model for each condition controls for age. Several notable patterns emerge in these estimates. First, the probability of all health conditions increases with age, with the exception of psychological problems. Physical health tends to be the most important correlate of all objective health conditions, with the exception of diabetes. Mental health tends to be more closely related to the objective health conditions relative to cognitive health, with the exception of diabetes and high blood pressure. Arthritis is strongly associated with poor physical health, confirming the close link between arthritis and physical limitations. However, surprisingly, arthritis is positively associated with cognitive and mental health. While psychological issues correlate with poor mental health (-0.484), they demonstrate a stronger relationship with physical health than with cognitive health, which is counterintuitive. Heart problems are closely linked to poor physical and mental health. Conversely, cancer does not display a significant association with cognitive or mental health, whereas lung disease and stroke are strongly associated with all three latent health components.

Last, in Table 8, we examine the relationship between latent health components and self-reported health. The loading associated with physical health is larger than that of cognitive and mental health, indicating that respondents tend to prioritize somatic limitations in their overall health assessments. Nonetheless, all

Table 8: Results: Self-Reported Health

	Self-Reported Health
x^c	0.473***
	(0.028)
x^{mh}	0.555***
	(0.037)
x^p	0.625***
	(0.030)
Age	0.013***
	(0.003)

Note: Standard errors presented in parentheses below point estimates.

components significantly influence self-reported health, indicating that individuals with varying perceptions of latent health components may still self-report a comparable overall rating of their health. Self-reported health is frequently utilized in applied research, partly due to its accessibility and standardized measurement across various micro-datasets (French, 2005; Capatina, 2015; De Nardi et al., 2024). Furthermore, numerous studies show that it correlates strongly with other health measures and predicts future mortality, even when controlling for various factors (Pijoan-Mas and Ríos-Rull, 2014). However, aggregating health components can obscure the mechanisms by which health affects economic behavior. This is illustrated in Table 6, where the probability of employment differs across latent health dimensions. These findings underscore the necessity for further investigation into the causes and consequences of health dimensions, which remains the primary objective of this working paper. Such analysis is essential for predicting the distributional effects of policy changes, such as alterations to the retirement age, and for developing more targeted interventions to address health-related inequalities.

5 Conclusion

This research investigates the formation and dynamics of health in older adults, focusing on the interplay among physical, cognitive, and mental health components. We develop and estimate a model using data from the Health and Retirement Study to understand how these health dimensions evolve over time and influence economic behaviors, such as employment and retirement decisions. Our key findings suggest positive dynamic complementarities between health components, with physical functioning showing strong persistence and a significant impact on mental health. The study also examines how latent health components relate to commonly used health measures and predict outcomes like medical expenditures in retirement.

At its current stage, this paper lays the groundwork for a deeper understanding of health formation and its role in shaping health-related economic inequalities at older ages. Ongoing development of this research builds on three main areas. First, we are relaxing distributional assumptions on the health components, independence assumptions on unobservables, and functional form assumptions within the formation process, to better capture the underlying structure of health dynamics. Second, given the credible identification of how economic factors affect health, we are working to strengthen the other causal direction; how health affects economic outcomes and other commonly used measures, notable self-reported health. While the current analysis is largely descriptive, advancing this causal link will enhance the interpretability of both the latent health measures and their implications. Finally, we aim to incorporate the estimated health formation technology into the theoretical model presented in Section 2. This will structurally link health dynamics with economic behavior and enable counterfactual policy analyses, such as evaluating the impacts of retirement age reforms.

References

Adhvaryu, A., Fenske, J., and Nyshadham, A. (2019). Early life circumstance and adult mental health. *Journal of Political Economy*, 127(4):1516–1549.

Agostinelli, F. and Wiswall, M. (2025). Estimating the technology of children's skill formation. *Journal of Political Economy*, 133(3):846–887.

Almond, D., Currie, J., and Duque, V. (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature*, 56(4):1360–1446.

Amengual, D., Bueren, J., and Crego, J. A. (2021). Endogenous health groups and heterogeneous dynamics of the elderly. *Journal of applied econometrics*, 36(7):878–897.

Banks, J., Muriel, A., and Smith, J. P. (2011). Attrition and health in ageing studies: evidence from elsa and hrs. *Longitudinal and life course studies*, 2(2):10–14301.

Bartel, A. and Taubman, P. (1979). Health and labor market success: The role of various diseases. *The Review of Economics and Statistics*, pages 1–8.

Benítez-Silva, H., Buchinsky, M., Man Chan, H., Cheidvasser, S., and Rust, J. (2004). How large is the bias in self-reported disability? *Journal of Applied Econometrics*, 19(6):649–670.

Biasi, B., Dahl, M. S., and Moser, P. (2021). Career effects of mental health. Technical report, National Bureau of Economic Research.

Black, B., French, E., McCauley, J., and Song, J. (2018). The effect of disability insurance receipt on mortality. In *Memorandum 19/2012*. University College London.

Blundell, R., Britton, J., Dias, M. C., and French, E. (2023). The impact of health on labor supply near retirement. *Journal of Human Resources*, 58(1):282–334.

Borella, M., Bullano, F. A., De Nardi, M., Krueger, B., and Manresa, E. (2024). Health inequality and health types. Technical report, National Bureau of Economic Research.

Bound, J. (1989). The health and earnings of rejected disability insurance applicants.

Bound, J., Stinebrickner, T., and Waidmann, T. (2010). Health, economic resources and the work decisions of older men. *Journal of Econometrics*, 156(1):106–129.

Butler, J. S., Burkhauser, R. V., Mitchell, J. M., and Pincus, T. P. (1987). Measurement error in self-reported health variables. *The Review of Economics and Statistics*, pages 644–650.

Capatina, E. (2015). Life-cycle effects of health risk. *Journal of Monetary Economics*, 74:67–88.

Capatina, E. and Keane, M. (2023). Health shocks, health insurance, human capital, and the dynamics of earnings and health. Technical report, Federal Reserve Bank of Minneapolis.

Case, A. and Deaton, A. (2022). The great divide: education, despair, and death. *Annual Review of Economics*, 14(1):1–21.

Case, A., Fertig, A., and Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of health economics*, 24(2):365–389.

Case, A. and Paxson, C. (2010). Causes and consequences of early-life health. *Demography*, 47(1):S65–S85.

Conti, G., Heckman, J., and Urzua, S. (2010). The education-health gradient. *American Economic Review*, 100(2):234–38.

Cozzi, G., Mantovan, N., and Sauer, R. M. (2024). How important are mental and physical health in career and family choices? Technical report, IZA Discussion Papers.

Cunha, F. and Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of human resources*, 43(4):738–782.

Cunha, F., Heckman, J. J., and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.

Cunha, F., Nielsen, E., and Williams, B. (2021). The econometrics of early childhood human capital and investments. *Annual Review of Economics*, 13(1):487–513.

Currie, J. (2009). Healthy, wealthy, and wise: Is there a causal relationship between child health and human capital development? *Journal of Economic Literature*, 47(1):87–122.

Currie, J. and Madrian, B. C. (1999). Health, health insurance and the labor market. *Handbook of labor economics*, 3:3309–3416.

Currie, J. and Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, 118(4):1495–1532.

Currie, J., Stabile, M., Manivong, P., and Roos, L. L. (2010). Child health and young adult outcomes. *Journal of Human Resources*, 45(3):517–548.

Cutler, D. M., Lleras-Muney, A., and Vogl, T. (2008). Socioeconomic status and health: dimensions and mechanisms.

Danesh, K., Kolstad, J. T., Spinnewijn, J., and Parker, W. D. (2024). The chronic disease index: Analyzing health inequalities over the lifecycle. Technical report, National Bureau of Economic Research.

Darden, M. E. (2022). Cognitive decline and dynamic selection. Technical report, National Bureau of Economic Research.

De Nardi, M., French, E., Jones, J. B., and McCauley, J. (2016). Medical spending of the us elderly. *Fiscal Studies*, 37(3-4):717–747.

De Nardi, M., Pashchenko, S., and Porapakkarm, P. (2024). The lifetime costs of bad health. *Review of Economic Studies*, page rdae080.

Del Bono, E., Kinsler, J., and Pavan, R. (2022). Identification of dynamic latent factor models of skill formation with translog production. *Journal of Applied Econometrics*, 37(6):1256–1265.

Dijk, H. and Mierau, J. (2023). Mental health over the life course: Evidence for a u-shape? *Health Economics*, 32(1):155–174.

Fitzpatrick, M. D. and Moore, T. J. (2018). The mortality effects of retirement: Evidence from social security eligibility at age 62. *Journal of Public Economics*, 157:121–137.

Frank, R. G. and Glied, S. A. (2023). America's continuing struggle with mental illnesses: Economic considerations. *Journal of Economic Perspectives*, 37(2):153–178.

French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. *The Review of Economic Studies*, 72(2):395–427.

French, E. and Jones, J. B. (2017). Health, health insurance, and retirement: a survey. *Annual Review of Economics*, 9(1):383–409.

Gilleskie, D. B. (1998). A dynamic stochastic model of medical care use and work absence. *Econometrica*, pages 1–45.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2):223–255.

Guvenen, F., Kuruscu, B., Tanaka, S., and Wiczer, D. (2020). Multidimensional skill mismatch. *American Economic Journal: Macroeconomics*, 12(1):210–244.

Hai, R. and Heckman, J. J. (2022). The causal effects of youth cigarette addiction and education. Technical report, National Bureau of Economic Research.

Heckman, J. J. (1991). Identifying the hand of past: Distinguishing state dependence from heterogeneity. *The American Economic Review*, 81(2):75–79.

Heineck, G. and Anger, S. (2010). The returns to cognitive abilities and personality traits in germany. *Labour economics*, 17(3):535–546.

Hosseini, R., Kopecky, K. A., and Zhao, K. (2021). How important is health inequality for lifetime earnings inequality? Technical report, Working paper.

Hosseini, R., Kopecky, K. A., and Zhao, K. (2022). The evolution of health over the life cycle. *Review of Economic Dynamics*, 45:237–263.

Idler, E. L. and Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of health and social behavior*, pages 21–37.

Iskhakov, F. (2010). Structural dynamic model of retirement with latent health indicator. *The Econometrics Journal*, 13(3):S126–S161.

Jolivet, G. and Postel-Vinay, F. (2020). A structural analysis of mental health and labor market trajectories.

Kahn, M. E. (1998). Health and labor market performance: the case of diabetes. *Journal of Labor Economics*, 16(4):878–899.

Kreider, B. and Pepper, J. V. (2007). Disability and employment: Reevaluating the evidence in light of reporting errors. *Journal of the American Statistical Association*, 102(478):432–441.

Lise, J. and Postel-Vinay, F. (2020). Multidimensional skills, sorting, and human capital accumulation. *American Economic Review*, 110(8):2328–2376.

Lundborg, P., Nilsson, A., and Rooth, D.-O. (2014). Adolescent health and adult labor market outcomes. *Journal of Health Economics*, 37:25–40.

Meyer, B. D. and Mok, W. K. (2019). Disability, earnings, income and consumption. *Journal of Public Economics*, 171:51–69.

Michaud, A. and Wiczer, D. (2018). Occupational hazards and social disability insurance. *Journal of Monetary Economics*, 96:77–92.

Millard, R. G. (2025). Heterogeneous disability shocks and the dynamics of income, employment, and partial insurance. *Employment, and Partial Insurance*.

Moran, J. R., Short, P. F., and Hollenbeak, C. S. (2011). Long-term employment effects of surviving cancer. *Journal of health economics*, 30(3):505–514.

Mullahy, J. and Sindelar, J. (1996). Employment, unemployment, and problem drinking. *Journal of health economics*, 15(4):409–434.

O'Donnell, O., Van Doorslaer, E., and Van Ourti, T. (2015). Health and inequality. In *Handbook of income distribution*, volume 2, pages 1419–1533. Elsevier.

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

Poterba, J. M., Venti, S. F., and Wise, D. A. (2017). The asset cost of poor health. *The Journal of the Economics of Ageing*, 9:172–184.

Ruhm, C. J. (2025). Mental health and mortality trends in the united states. *Journal of Health Economics*, 102:103015.

Smith, J. P. (2004). Unraveling the ses: health connection. *Population and development review*, 30:108–132.

Stephens, M., Toohey, D. J., et al. (2018). The impact of health on labor market outcomes: experimental evidence from mrfit. Technical report, National Bureau of Economic Research.

Stern, S. (1989). Measuring the effect of disability on labor force participation. *Journal of human Resources*, pages 361–395.

Strulik, H. (2022). A health economic theory of occupational choice, aging, and longevity. *Journal of Health Economics*, 82:102599.

Wang, B., Frank, R., and Glied, S. (2023). Lasting scars: The impact of depression in early adulthood on subsequent labor market outcomes. *Health Economics*, 32(12):2694–2708.

Wen, J. (2022). Occupational retirement and social security reform: the roles of physical and cognitive health. *Available at SSRN 4235450*.

White, M. N. (2023). Self-reported health status and latent health dynamics. *Journal of Health Economics*, 88:102727.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of applied econometrics*, 20(1):39–54.

Yogo, M. (2016). Portfolio choice in retirement: Health risk and the demand for annuities, housing, and risky assets. *Journal of monetary economics*, 80:17–34.

Appendix to "Estimating the Process of Health Formation in Older Adults: The Roles of Physical Functioning, Cognition, and Mental Health." by Robert Millard and Steven Stern

A Identification

Given the assumptions outlined in the main text, a single measure for each latent factor, along with $t \geq 3$ periods, is sufficient to identify the model parameters. In this context, consider representing the model such that $X_t \in \mathbb{R}^3$ are the latent health components, $W_t \in \mathbb{R}^k$ are observable inputs to the production technology, such as employment, and U_t are additional control variables, such as age. For now, suppose that b_0 is observed and contained in U_t. We show how the parameters associated with b_0 are recovered at the end. With this, the model can be written as

$$X_t = a + BX_{t-1} + C_1 W_{t-1} + C_2 U_t + e_t, \tag{12}$$

$$W_t = D_1 X_t + D_2 U_t + \nu_t, \tag{13}$$

where $e_t \sim iid\mathcal{N}(0, \Sigma_e)$ represents the unobservable component of the production function, and $\nu_t \sim iid\mathcal{N}(0, \Sigma_\nu)$ denotes the unobservable determinants of W_t. We assume that e_t is independent of $(X_1, \{\nu_s\}_{s=1,..,T_i})$ and ν_t is independent of $(X_1, \{e_s\}_{s=1,...T_i})$.

We do not directly observe X_t. Instead, we observe a single noisy measurement for each of it's components. These measurements can be categorized into two types: continuous measures associated with the factors, and binary measures modeled through a probit link. For notational clarity, we denote $j, r \in \mathcal{B}$ as the index for the x's measured by binary indicators, and $i, l \in \mathcal{C}$ as the index for the x's measured by continuous indicators. We define the set of binary measures as $\mathcal{B}$ and the set of continuous measures as $\mathcal{C}$, with $|\mathcal{C}| \geq 1$ and $|\mathcal{B}| \geq 1$ to ensure we have at least one of each type.

- For x_t^i we observe $z_t^i = \mu^i + \lambda^i x_t^i + v_t^i$, where $v_t^i \sim iidN(0, \sigma^2_{v^i})$, $i \in \mathcal{C}$,
- For x_t^j we observe $z_t^j = \mathbb{1}\{x_t^j + v_t^j > \tau\}$, where $v_t^j \sim iidN(0, 1)$, $j \in \mathcal{B}$.
- v_t^i and v_s^j are assumed independent for all $i \in \mathcal{C}$, $j \in \mathcal{B}$, and for all $t, s = 1, ..., T_i$. Moreover, v_t^i and v_t^j are independent of ν_s, for all $t, s = 1, .., T$

- We denote the covariance matrix of measurement errors as Σ_v, which is a diagonal matrix, whose elements corresponding to the variance of discrete measures is one.

First, the control variables shift the means only and we can transform the model to partial out U_t. The data moments (variances, covariances, etc.) used to identify the model parameters are computed conditional on the relevant control variables, $\mathcal{U}_{st} = (U_s, ..., U_t)$. The coefficients for U_t, C_2, D_2, are recovered following identification of the other model parameters. We transform X_t and W_t as follows,

$$
\begin{aligned}
W_t^c &= W_t - E(W_t|U_t) \\
&= D_1 X_t^c + \nu_t, \\
X_t^c &= X_t - E(X_t|U_t) \\
&= B X_{t-1}^c + C_1 W_{t-1}^c + e_t.
\end{aligned} \tag{14}
$$

We can simplify the model further though substitution of W_{t-1}^c into Equation (14) to obtain

$$
\begin{aligned}
X_t^c &= a + B^* X_{t-1}^c + u_t, \\
B^* &= (B + C_1 D_1), \\
u_t &= C_1 \nu_{t-1} + e_t \ \sim N(0, \Sigma_u).
\end{aligned}
$$

Define $\Sigma_t = var(X_t|U_t)$ and we normalize $\Sigma_1(i,i) = 1$, for $i = 1, .., 3$, as in the main text. The off-diagonal elements of Σ_1, which represent the covariance of the initial latent factors, are parameters to be identified. Additionally, we define the conditional population moments for the serial covariance of X_t as

$$
\Gamma_{t,s} \equiv cov(X_t^c, X_{t-1}^c | \, \mathcal{U}_{ts}),
$$

With equation (12) we can write the serial covariance matrices in terms of the model parameters,

$$
\Gamma_{t,t-1} = B^* \Sigma_{t-1} \tag{15}
$$

$$
\Gamma_{t,1} = B^* \Gamma_{t-1,1} \tag{16}
$$

$$
\Sigma_t = B^* \Sigma_{t-1} B^{*\prime} + \Sigma_e. \tag{17}
$$

With these preliminaries, identification of the model is outlined in subsequent steps. The parameters to identify are the coefficients in Equations 12, $\{a, B, C_1, C_2, D_1, D_2\}$, the parameters of the measurement system, $\{\lambda, \mu, \tau, \Sigma_v\}$, the distributional parameters for the latent variables, $\{\Sigma_t, \Gamma_{t,s}, \Sigma_u, \Sigma_e\}$, and the correlation structure of initial endowments, Σ_1.

1. First, we can characterize the elements of serial covariances $\Gamma_{t,t-1}$ in terms of parameters and observed data. The elements of $\Gamma_{t,t-1}$ take one of three functional forms depending on the types of measures for the respective elements of X_t. Then we use the dynamics of Equation (12) to write parameters in terms of observables only.

   (a) **Discrete Measure Moments:** For each discrete measure, j, define $var(x_{jt} + v_{jt}|U_t) \equiv s_t^j = \sqrt{var(x_t^j|U_t) + 1}$.[15] Note that s_t^j is a function of U_t and can be written as $s_t^j(U_t)$, but we keep this dependence implicit for notational convenience. For each j, we derive a standardized conditional index as

$$V_t^j = \frac{x_t^j - E(x_t^j|U_t) + v_t^j}{s_t^j},\ V_t^j \sim N(0,1).$$

   Then, the discrete measures can be expressed as $z_t^j = \mathbb{1}\{V_t^j > k_t^j\}$, where $k_t^j = \frac{\tau^j - E(x_t^j|U_t)}{s_t^j}$ is the conditional standardized threshold. We identify $k_t^j,\ \forall j \in \mathcal{B},\ t = 1, ..., T_i$ from the conditional expectation of z_t^j:

$$\begin{aligned} E(x_t^j|U_t) &= P(z_t^j = 1|U_t) \\ &= 1 - \Phi(k_t^j) \\ \Rightarrow \hat{k}_t^j &= \Phi^{-1}\big(1 - P(z_t^j = 1|U_t)\big). \end{aligned}$$

   where Φ is the standard normal CDF. Similarly, we identify $\tau^j,\ \forall j \in \mathcal{B}$ from z_1^j given the normalization imposed in the first period, $\hat{\tau}^j = \sqrt{2}\hat{k}_1^j$.

   Next, we recover the conditional tetrachoric correlations using the conditional binary-binary joint probabilities across time.[16] For any two time periods $t \neq s$, (V_t^j, V_s^j) is bivariate standard

[15] Note that $var(v_t^j|U_t) = var(v_t^j) = 1$ and $E(v_t^j|U_t) = E(v_t^j) = 0$.

[16] The tetrachoric correlation is a statistical method used to estimate the correlation between two dichotomous (binary) variables that are assumed to be derived from an underlying continuous and normally distributed latent variable. Essentially, it estimates what the correlation would be if the variables were measured on a continuous scale rather than being artificially categorized.

normal with

$$\begin{aligned}\rho_j^{ts} &= corr(V_t^j, V_s^j |\, \mathcal{U}_{ts}) \\ &= \frac{cov(x_t^j, x_s^j |\, \mathcal{U}_{ts})}{s_t^j s_s^j}.\end{aligned}$$

To pin down ρ_j^{ts}, consider the joint probability

$$\begin{aligned}P(z_t^j = 1, z_s^j = 1 |\, \mathcal{U}_{ts}) &= P(V_t^j > k_t^j,\ V_s^j > k_s^j | \mathcal{U}_{ts}) \\ &= 1 - \Phi(k_t^j) - \Phi(k_s^j) + \Phi_2(k_t^j, k_s^j; \rho_j^{ts}),\end{aligned}$$

where Φ_2 is a bivariate standard normal CDF with correlation ρ_j^{ts}. The LHS of this equation is observed, and the RHS is a strictly increasing function in ρ_j^{ts} for any fixed $\hat{k}_t^j, \hat{k}_s^j$ (which are identified from above). Hence, we can rearrange and invert Φ_2 to identify the tetrachoric correlation coefficient,

$$\hat{\rho}_j^{ts} = \Phi_2^{-1}\Big(P(z_t^j = 1, z_s^j = 1 |\, \mathcal{U}_{ts}) - 1 + \Phi(\hat{k}_t^j) + \Phi(\hat{k}_s^j); \hat{k}_t^j, \hat{k}_s^j\Big),$$

Doing this for all $(t, s), t \neq s$ pairs identifies the entire correlation matrix for the standardized indices, $V_t^j = \frac{x_t^j - E(x_t^j | U_t) + v_t^j}{s_t^j}$. Because v_t^j are independent across time and from states, the latent covariances satisfy

$$cov(x_t^j, x_s^j |\, \mathcal{U}_{ts}) = \hat{\rho}_j^{ts} s_t^j s_s^j.$$

Hence, $cov(x_t^j, x_s^j |\, \mathcal{U}_{ts})$ is identified up to scale of s_t^j, s_s^j, for all $t, s \in \{1, ..., T_i\}, t \neq s$ and $j \in \mathcal{B}$.

(b) **Mixed measure moments**: From the mixed-type measure moments, $cov(z_t^i, z_s^j)$, we have the following probit identity

$$\begin{aligned}M_{t,s}^{ij} \equiv \frac{s_s^j}{\phi(\hat{k}_s^j)} cov(z_t^i, z_s^j |\, \mathcal{U}_{ts}) = cov(x_t^i, x_s^j |\, \mathcal{U}_{ts})\lambda^i \\ \Rightarrow\ cov(x_t^i, x_s^j |\, \mathcal{U}_{ts}) = M_{t,s}^{ij} / \lambda^i.\end{aligned}$$

Hence, the entire conditional cross-covariance block between x_t^i and x_s^j is revealed up to scale, λ^i and s_s^j.

(c) **Continuous measure moments**: For $t \neq s$, we have

$$cov(x_t^i, x_s^i | \mathcal{U}_{ts}) = cov(z_t^i, z_s^i | \mathcal{U}_{ts})/(\lambda^i)^2,$$
$$var(x_t^i | U_t) = \left(var(z_t^i | U_t) - \sigma^2_{v^i}\right)/(\lambda^i)^2,$$

where $var(v_t^i) = \sigma^2_{v^i}$, $\forall t$. Hence, $cov(x_t^i, x_s^i | \mathcal{U}_{ts})$ are known up to scale, $(\lambda^i)^2$, for all $t, s \in \{1, .., T_i\}$ and $i \in \mathcal{C}$.

2. Consider the first three periods. With the covariance structure of observed measures characterized in terms of the covariance structure of latent health components, the model's imposed dynamic structure solves for $(\lambda, s_2, s_3, \Sigma_1)$, where $\lambda = \{\lambda^i\}_{i \in \mathcal{C}}$, $s_2 = \{s_2^j\}_{j \in \mathcal{B}}$, and $s_3 = \{s_3^j\}_{j \in \mathcal{B}}$. For any $t \neq s$, the elements of $\Gamma_{t,s} = cov(X_t, X_s | \mathcal{U}_{st})$ are as follows.

   - If x_t^j and x_s^r are observed by binary measures, then

   $$\Gamma_{ts}(j, r) = \rho_j^{ts} s_t^j s_s^r.$$

   - If x_t^i is observed by a continuous measure and x_s^j is observed by a binary measure, then

   $$\Gamma_{ts}(i, j) = M_{ts}^{ij}/\lambda^i.$$

   - If x_t^i and x_s^l are observed by continuous measures, then

   $$\Gamma_{ts}(i, l) = cov(z_t^i, z_s^l | \mathcal{U}_{ts})/(\lambda^i \lambda^l).$$

   Consider the two cross-lag anchor matrices, Γ_{21} and Γ_{31}, where period 1 serves as the anchor. The previous step expresses Γ_{21} and Γ_{31} as explicit functions of the unknown parameters, $(\lambda, s_2, s_3, \Sigma_1)$, and observables only. Using Equation (15), we have

   $$\Gamma_{2,1} = B^* \Sigma_1$$
   $$\Rightarrow B^* = \Gamma_{21} \Sigma_1^{-1}.$$

Hence, B^* is an explicit matrix function of (λ, s_2, Σ_1),

$$B^*\big(\lambda, s_2, \Sigma_1\big) = \Gamma_{21}\big(\lambda, s_2\big)\Sigma_1^{-1}.$$

Using Equation (16), and substituting yields

$$\Gamma_{31}\big(\lambda, s_3\big) = B^*\big(\lambda, s_2, \Sigma_1\big)\Gamma_{21}\big(\lambda, s_2\big).$$

With four latent states and $|\mathcal{C}| \geq 1$, this gives us 9 scalar equations to solve for at most 3+4+1=8 parameters.[17] Everything else on both sides of the equations is observable or already expressed in terms of the parameters. Thus we identify $(\hat{\lambda}, \hat{s}_2, \hat{s}_3, \hat{\Sigma}_1)$, and subsequently identify B^*.

Intuitively, we compute B^* from period 1-2 cross-covariances of measures. Then $\Gamma_{31} = B^*\Gamma_{21}$ insists the same B^* must also propagate period 1-3 via period 2. Equating both ways to do the propagation pins down the scalar parameters.

3. Given $(\hat{\lambda}, \hat{s}_2, \hat{s}_3, \hat{\Sigma}_1)$, we can recover Σ_2, Σ_3 and Σ_u. First the off-diagonal elements of Σ_2 are

$$\begin{aligned}
cov(x_2^j, x_2^r|U_2) &= \hat{\rho}_{jr}^{22}\hat{s}_2^j\hat{s}_2^r \\
cov(x_2^i, x_2^j|U_2) &= M_{22}^{ij}/\hat{\lambda}^i \\
cov(x_2^i, x_2^l|U_2) &= cov(z_2^i, z_2^l|U_2)/\hat{\lambda}^i\hat{\lambda}^l
\end{aligned}$$

For the diagonal elements of Σ_2 that correspond to discrete measures,

$$var(x_2^j|U_2) = (\hat{s}_2^j)^2 - 1.$$

for the diagonal elements of Σ_2 that correspond to continuous measures,

$$\Gamma_{32} = \hat{B}^*\Sigma_2,$$

[17] Each continuous measure is associated with one parameter, λ^i, and each discrete measure is associated with has two parameters, (s_2^j, s_3^j). Moreover, there are three correlation coefficients to be identified in Σ_1. Hence, with at least one measure of each type, we have at most 8 parameters to identify with nine equations.

where the second column/row from

$$cov(x_3^i, x_2^l|\, \mathcal{U}_{23}) = (\hat{B}^*\Sigma_2)_{i,l}$$
$$= cov(z_3^i, z_2^l|\, \mathcal{U}_{23})/\hat{\lambda}^i\hat{\lambda}^l.$$

Given $\hat{\Sigma}_2$, we recover Σ_u using relation (17),

$$\hat{\Sigma}_u = \hat{\Sigma}_2 - \hat{B}^*\hat{\Sigma}_1\hat{B}^{*\prime}.$$

Similarly, we recover Σ_3 from

$$\hat{\Sigma}_3 = \hat{B}^*\hat{\Sigma}_2\hat{B}^{*\prime} + \hat{\Sigma}_u.$$

Continuing in this fashion identifies Σ_t for all $t > 3$.

4. Next, we separately identify the components of B^* (B, C_1, and D_1) and Σ_u (Σ_e and Σ_ν) using the serial covariances between X_t and W_s. Using equation (12) we can then write these moments as functions of the parameters to identify. In the first two periods we have,

$$cov(W_2, X_1|\, \mathcal{U}_{12}) = D_1(B + C_1D_1)\Sigma_1 = D_1B^\star\Sigma_1, \tag{18}$$
$$cov(X_2, W_1|\, \mathcal{U}_{12}) = (B + C_1D_1)\Sigma_1D_1' + C_1\Sigma_\nu = B^\star\Sigma_1D_1' + C_1\Sigma_\nu, \tag{19}$$
$$var(W_1|\, U_1) = D_1\Sigma_1D_1' + \Sigma_\nu. \tag{20}$$

Given $(\hat{\lambda}, \hat{s}_2, \hat{s}_3)$, these moments pin down the remaining model parameters in sequence.[18]

$$\hat{D}_1 = cov(W_2, X_1|\, \mathcal{U}_{12})(\hat{B}^*\hat{\Sigma}_1)^{-1},$$
$$\hat{\Sigma}_\nu = var(W_1|U_1) - \hat{D}_1\hat{\Sigma}_1\hat{D}_1',$$
$$\hat{C}_1 = \Big(cov(X_2, W_1|\, \mathcal{U}_{12}) - \hat{B}^*\hat{\Sigma}_1\hat{D}_1'\Big)\hat{\Sigma}_\nu^{-1}.$$

Then, we can identify B (separate from C_1 and D_1) as $B = \hat{B}^* - \hat{C}_1\hat{D}_1$. Further we can identify Σ_e

[18]Note that given, $(\hat{\lambda}, \hat{s}_2, \hat{s}_3)$, $cov(W_2, X_1|\, \mathcal{U}_{12})$, $cov(X_2, W_1|\, \mathcal{U}_{12})$, $var(W_1|\, \mathcal{U}_{12})$ can be written in terms of data only.

using

$$\Sigma_2 = B^*\Sigma_1 B^{*\prime} + \Sigma_u$$
$$= B^*\Sigma_1 B^{*\prime} + C\Sigma_\nu C_1' + \Sigma_e$$
$$\Rightarrow \hat{\Sigma}_e = \hat{\Sigma}_2 - \hat{B}^*\hat{\Sigma}_1\hat{B}^{*\prime} + \hat{C}_1\hat{\Sigma}_\nu\hat{C}_1'$$

5. To identify the contemporaneous effects of control variables, U_t, (C_2, D_2). First, distinguish those that do and do not vary over time. For those that deterministically vary over time (i.e., age), we can back out it's coefficient from how the average path of the latent state shifts with age, relative to what dynamics B, C_1 would predict absent age. A similar argument holds for the W_t equation.

   For time-fixed control vars, their effects are identified off cross-sectional convariation across time. For instance, suppose b_0 is observed, then we take covariances of the state/input equations with b_0.

$$cov(X_t, b_0) = Bcov(X_{t-1}, b_0) + C_1 cov(W_{t-1}, b_0) + C_2^{(b_0)} Var(b_0) \tag{21}$$
$$cov(W_t, b_0) = D_1 cov(X_t, b_0) + D_2^{(b_0)} var(b_0) \tag{22}$$

   using exogeneity of b_0 to (e_t, ν_t). With B, C_1, D_1 already identified in earlier steps, these linear relationships pin down $C_2^{(b_0)}$ and $D_2^{(b_0)}$ from observed $\{cov(X_t, b_0), cov(W_t, b_0)\}$.

   Now, if b_0 is latent (measured with its own noisy indicators), we use the dedicated "b_0" block. Normalize b_0 in period 1, as in the main text. Then use co-variation of the b_0 measures with the x_0 measures to get initial correlations. And use co-variation with x_t and with W_t to recover contemporaneous effects.

6. Lastly, we recover the intercepts from Equation (12). Once $\hat{\lambda}^i$ and $\hat{\tau}^i$ are known, we compute $E(X_t|U)$ from the expected value of measurements. Then we have,

$$E(X_t|U_t) = a + BE(X_t|U_t) + C_1 E(W_{t-1}|U_t) + C_2 U_t \tag{23}$$

   and can solve linearly for a.

B Estimation Details

The likelihood function is constructed from the joint distribution of initial heterogeneity, the latent health components, and the observed measures of health inputs and employment over t. Define the parameter vector as

$$\theta = (\rho, \{a^k, C^k, \sigma_{\nu^k}\}_{k \in \{p,c,mh\}}, \theta^m, \beta^L, \{\beta^{I_j}\}_{j=1,\ldots,K_I}, \beta^s, \beta^a), \tag{24}$$

which collects all relevant measurement and transition parameters, where

$$\begin{aligned}
a^k &= (a_0^k, ..., a_4^k)', \\
C^k &= (c_1^k, ..., c_{K_I}^k, c_e^k)', \\
\beta^L &= (\beta_1^L, ..., \beta_6^L)', \\
\beta^{I_j} &= (\beta_1^{I_j}, ..., \beta_5^{I_j})', \\
\beta^s &= (\beta_0^s, ..., \beta_5^s)', \\
\beta^a &= (\beta_0^a, ..., \beta_5^a)', \\
\theta^m &= \{\mu_m^k, \lambda_m^k, \tau_{j,m}^k, \sigma_{v_m^k}\}_{k \in \{p,c,mh\}, m \in 1,\ldots,M_k}
\end{aligned}$$

Let $z_t = (z_{t,1}^p, ..., z_{M^p}^p, z_{t,1}^{mh}, ..., z_{M^{mh}}^{mh}, z_{t,1}^c, ..., z_{M^c}^c)$ denote the vector of all health measures observed at time t. Moreover, we observe employment status, L_t, and the vector of health investments, I_t. The joint density of observed health measures at time t, given latent health components, $x_t = (x_t^p, x_t^{mh}, x_t^c)$, and θ is

$$f(z_t | x_t^k, \theta) = \prod_{k \in \{p,mh,c\}} \prod_{m=1}^{M_k} f_{z_{t,m}^k}(z_{t,m}^k | x_t^k, \theta) \tag{25}$$

where $f_{z_{t,m}^k}(z_{t,m}^k | x_t^k, \theta)$ is the density of measure $z_{t,m}^k$ for $m = 1, ..., M^k, k \in \{p, mh, c\}$. The joint density of the latent health vector, given last periods latent health vector, intial heterogeneity, previous employment status, health investments, and θ is $f_{x_t}(x_t | x_{t-1}, b_0, L_{t-1}, I_t, \theta)$.

Similarly, the joint density of observed measures at $t = 0$ is

$$f(z_0 | x_0^k, \theta) = \prod_{m=1}^{M_b} f_{z_{0,m}^k}(z_{0,m}^b | b_0, \theta) \prod_{k \in \{p,mh,c\}} \prod_{m=1}^{M_k} f_{z_{0,m}^k}(z_{0,m}^k | x_0^k, \theta), \tag{26}$$

where $f_{z^k_{0,m}}(z^b_{0,m}|b_0,\theta)$ is the density of initial latent heterogeneity measure $z^b_{0,m}$ for $m = 1, ..., M^b_0$. The joint density of initial endowments, $w_0 = (x^p_0, x^{mh}_0, x^c_0, b_0)$ given θ is $f_{w_0}(w_0)$.

The likelihood contribution of observed control variables, L_t and I_t are

$$
\begin{aligned}
f_L(L_t|x_t, b_0, t, \theta) &= P(L_t = 1|x_t, b_0, t, \theta)^{\mathbb{1}\{L_t=1\}} P(L_t = 0|x_t, b_0, t, \theta)^{\mathbb{1}\{L_t=0\}} \\
f_{I_j}(I_{jt}|x_t, b_0, \theta) &= P(I_{jt} = 1|x_t, b_0, t, \theta)^{\mathbb{1}\{I_{jt}=1\}} P(I_{jt} = 0|x_t, b_0, t, \theta)^{\mathbb{1}\{I_{jt}=0\}}, \text{ for all } j \in K_I
\end{aligned}
$$

The likelihood contribution for individual i, given parameter values θ can be expressed recursively as

$$
\begin{aligned}
L_i(\theta) = \int \cdots \int f_{w_0}(w_0) f(z_0|x^k_0, \theta) f(L_0|w_0, \theta) f(I_0|w_0, \theta) \prod_{t=1}^{T_i} \Bigg[& f_{x_t}(x_t|x_{t-1}, b_0, L_{t-1}, I_t, \theta) \\
& \times \Big\{ H(d_t = 1|x_t, t, b_0, \theta) \Big\}^{d_t(1-a_t)} \\
& \times \Big\{ P(a_t = 1|x_t, b_0, t, \theta) \Big\}^{a_t(1-d_t)} \\
& \times \Big\{ \big(1 - P(a_i = 1|x_t, b_0, t, \theta)\big)\big(1 - H(d_t = 1|x_t, b_0, t, \theta)\big) \\
& \times f(z_t|x^k_t, \theta) f_L(L_t|t, x_t, b_0, t, \theta) \prod_{j=1}^{K_I} \big[f_{I_j}(I_{jt}|x_t, b_0, t, \theta) \big] \Big\}^{(1-d_t)(1-a_t)} \Bigg] db_0 dx_0, ..., dx_t,
\end{aligned}
\tag{27}
$$

where $H(d_t = 1|x_t, b_0, t, \theta)$ is the probability of death in t, $P(a_i = 1|x_t, b_0, t, \theta)$ is the probability of attrition in t, and T_i is the period where individual i is observed to die or attrit from the sample. Initial heterogeneity and the health components are latent and need to be integrated out of the likelihood. The integrals are taken over the latent initial heterogeneity, b_0 and latent health vector x_t for all t.

C Details on Simulated Maximum-Likelihood Estimation (SML)

In our model the observed-data likelihood involves integrating over the entire sequence of latent health stocks $\{x_t^1, x_t^2, x_t^3,\}_{t=1\ldots T}$, and initial heterogeneity, b_0. We approximate this high-dimensional integral by employing Monte Carlo draws from the joint distribution of latent states as dictated by the model. The distribution of $\{x_t^p, x_t^c, x_t^{mh}, b_0\}_{t=0\ldots T}$ is jointly normal given the assumptions described in the main text. The following describes the steps to simulate a time series and calculate the likelihood for a single individual.

1. *Simulate Latent Paths:* For a given guess of parameters θ, we simulate R independent draws of $\tilde{b}_0$ and paths of the latent states $\{\tilde{x}_t^p, \tilde{x}_t^c, \tilde{x}_t^{mh},\}_{t=0\ldots T}$ for r = 1, ..., R.
    - Initial draw: For each replication r= 1, ..., R, we draw $w_0^{(r)} = (\tilde{x}_0^p, \tilde{x}_0^c, \tilde{x}_0^{mh}, \tilde{b}_0)' \sim p(w_0|\theta)$, where $w_0 = (x_0^p, x_0^c, x_0^{mh}, b_0)'$, and evaluate the density fo the initial observable, L_0 and z_0.
    - Forward recursion: for t = 1, ..., T, we update the latent state $w_t^{(r)} = (\tilde{x}_t^p, \tilde{x}_t^c, \tilde{x}_t^{mh}, \tilde{b}_0)$ using the transition equation, (6), observed data, and a simulated draw of $\nu_t^{(r)}$. We compute the conditional probability of the realized data, L_t, I_t, z_t, given the simulated state, $w_t^{(r)}$.
2. *Evaluate the Likelihood Contribution of a Path:* For each replication r, compute the likelihood of observing the measures by (27), which we denote as $L^{(r)}(\theta)$.
3. *Approximate the Likelihood:* The true likelihood is then approximated by averaging over the R simulated paths,

$$\hat{L}_{obs}(\theta) \approx \frac{1}{R}\sum_{r=1}^{R} L^{(r)}(\theta). \tag{28}$$

Then for this individual, the log likelihood contribution is $l_{SML}(\theta) = \hat{L}_{obs}(\theta)$. We maximize the approximated log-likelihood with respect to parameters θ using a combination of Nealder-Meade and the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm.

D Motivating Framework for the Technology of Health Formation

This section presents a dynamic model to illustrate the many ways individuals' choices and circumstances affect, and in turn are affected by their health. We adapt the stylized model presented in Blundell et al. (2023), expanding the choice set and augmenting the health production technology to depend on individuals decisions. This exercise aids in guiding the empirical analysis and identification strategy for estimating the structural parameters of the health formation process.

We model an individual's decision process before their formal retirement age, $T+1$. The terminal condition is post-retirement, where no choices are made and the health stock, among relevant state variables, is taken as given. At each age, $t \in \{0, ..., T\}$, individuals choose how much to work and how much of their income to save, to consume of a regular consumption good, and to consume of a separate "unhealthy" good. The unhealthy consumption good, such as smoking or drinking, gives positive utility but is a detrimental input to health formation. Choices for individual i are made to maximize the expected discounted value of their current and future utility. Health is a stock which affects individuals choices through impacting current utility, as well as dynamically by determining future health stocks via the health formation process. In what follows, we consider the problem of a single cohort, so that time and age are interchangeable.

At each age t, individual's receive utility from consumption and leisure. Utility, given labor supply, Y_{it}, and stock of health capital H_{it}, is

$$U(C_{it}, I_{it}; Y_{it}, H_{it}, \xi_i, \zeta_{it}) = \frac{(C_{it} + \alpha I_{it})^{1-\gamma}}{1-\gamma} - v(Y_{it}, H_{it}, \xi_i, \zeta_{it}), \tag{29}$$

where C_{it} is normal good consumption and I_{it} is unhealthy good consumption. Unhealthy goods are detrimental to health production, so in order to rationalize interior solutions there is an additional preference parameter associated with to unhealthy good consumption, α, as in (Strulik, 2022). The utility cost of working is additively separable from consumption, which simplifies the solution to the problem and is commonly made in the related literature (Blundell et al., 2023; De Nardi et al., 2024). The utility cost of work, $v()$, depends on the stock of health, reflecting that working is more costly in periods of bad health. The utility costs of work depends on unobserved tastes heterogeneity, ξ_i, and idiosyncratic cost shock, ζ_{it}.

The individual's decision problem is subject to several dynamic constraints. First, the budget constraint summarizes available resources for consumption and savings:

$$A_{it+1} = (1 + r_t)(A_{it} + Y_{it}W_{it}(1 - \tau_t(Y_{it}W_{it}, H_{it})) + b_t(H_{it}) - C_{it} - p_t^I I_{it}) \tag{30}$$

The budget is determined by the assets available at the start of the period, A_{it}, potential earnings, W_{it}, and benefits that may be available for individuals in bad health, such as disability insurance, $b_t(H_t)$. If working, $Y_{it} = 1$, employment income is taxed according to the function $\tau()$, which depends on health, capturing health related tax credits that may be available. The consumption of regular goods, c_{it} is in real terms, and p_t^I reflects differences in the real cost of unhealthy goods. The interest rate, r_t, determines the per-period return to savings.

When working, the potential earnings of individual i at age t are

$$W_{it} = \omega(t, H_{it}, \phi_i, v_{it}). \tag{31}$$

Potential earnings combine the price and supply of labor and depend on age, approximating the life-cycle profile of productive human capital accumulation, and vary with the stock of health, representing a disruption of translating productive human capital into output for those in bad health. Earnings depend on an individual fixed effect, ϕ_i, which can be interpreted as heterogeneous productive ability, and v_{it} is an idiosyncratic transitory earnings shock.

Technology of health formation is described by,

$$H_{it} = h(t, H_{it-1}, Y_{it-1}, I_{it}, \psi_i, \epsilon_{it}). \tag{32}$$

The formation of health is influenced by individuals' prior work decisions. At older ages, physically demanding tasks or sustained exposure to stress in the workplace can negatively affect health (Strulik, 2022; Jolivet and Postel-Vinay, 2020). Conversely, continued labor force participation may have beneficial effects, such as providing a sense of purpose or preserving cognitive and physical functioning, which is consistent with evidence that retirement can lead to health deterioration or increased mortality risk (Black et al., 2018). Lifestyle choices, such as consumption of unhealthy goods like drinking alcohol and smoking cigarettes, also enter directly into the health production function.

Heterogeneity in health formation is captured by an individual-specific parameter, ψ_i, reflecting differences in health productivity. Because health evolves dynamically, early-life endowments and prior choices have long-term consequences for current health. This heterogeneity, which can be thought of as a "type" in the health formation process, accounts for differences in health outcomes arising from factors such as initial health endowments, genetic predispositions, past health shocks, and life-course decisions, including education and work history (Borella et al., 2024; De Nardi et al., 2024). Finally, health formation is subject to unobserved transitory shocks, ϵ_{it}, which introduce additional variation over time.

Finally, one's stock of health determines their expected lifespan. We define the probability that an individual who is alive at t survives to t+1 as

$$S(t, H_t). \tag{33}$$

The individual's survival probability captures a key tension in the decision problem: working more today provides more income for consumption today and savings in retirement, but at a utility cost of working and a costs to future health (affects leisure value in retirement and length of planning horizon).

D.1 Structure of Unobserved Components

The model allows for unobserved heterogeneity in health, earnings, and the preferences for work, (ϕ_i, ψ_i, ξ_i). We allow for arbitrary correlation between these three dimensions. For example, individuals of lower socioeconomic status when growing up may have been relatively deprived in investments that foster good health formation and human capital. Similarly, education may simultaneously impact health productivity, earnings, and preferences for work. These are distinct from transitory shocks to health, earnings, and preferences $(\zeta_{it}, v_{it}, \epsilon_{it})$, which are assumed serially uncorrelated, mutually independent, and independent of unobserved heterogeneity components.

D.2 Individual's Problem

The solution to the individual's problem at age t can be represented in the Bellman formulation. For the set of state variables, $\Omega_{it} = (t, H_{it}, A_{it}, \xi_i, \phi_i, \xi_i, \zeta_{it})$, the individual's choices satisfy,

$$\begin{aligned} V_t(\Omega_{it}) &= \max_{C_{it}, I_{it}, Y_{it}} \left\{ U(C_{it}, I_{it}, Y_{it}; H_{it}, \xi_i, \zeta_{it}) + \beta S(t, H_{it}) \mathbb{E} V_{t+1}(\Omega_{it+1}) \right\} \\ \text{subject to } A_{it+1} &= (1 + r_t)(A_{it} + Y_{it} W_{it}(1 - \tau_t(Y_{it} W_{it}, H_{it})) + b_t(H_{it}) - C_{it} - p_t^I I_{it}) \\ H_{it} &= h(t, H_{i,t-1}, u_{it}, Y_{it}, \psi_i, \epsilon_{it}) \\ W_{it} &= \omega(t, H_{it}, \phi_i, v_{it}). \end{aligned}$$

Consider an interior solution for consumption. Conditional on $Y_{it} = y$, the first order conditions are used to derive the optimal policy functions normal good and unhealthy good consumption. The first order conditions are

$$\frac{\partial}{\partial C}: U_C - \beta(1+r_t)S(t,H_t)\mathbb{E}V_{A,t+1} = 0 \tag{34}$$

$$\frac{\partial}{\partial I}: U_I - \beta(1+r_t)S(t,H_t)\mathbb{E}(p_t^I V_{A,t+1} + V_{H,t+1}H_I) = 0, \tag{35}$$

where $V_{x,t+1}$ denotes the partial derivative of V_{t+1} with respect to x. From equation (34), we obtain the usual relation between marginal utility of consumption and the discounted marginal value of assets, where discounting accounts for mortality risk. Then, combining the first order conditions we obtain

$$U_I = p_t^I U_c + \beta S(t,H_t)\mathbb{E}V_{H,t+1}H_I. \tag{36}$$

The optimal level of unhealthy good consumption, I^*, is such that the marginal utility of consuming unhealthy good equates marginal utility of consumption priced at the unhealthy good plus a term capturing the effect of the unhealthy good on future health times the marginal value of higher health in the future.

Denote Ω_{t+1}^y as next periods state variables conditional on $Y_{it} = y$, for $y \in \{0,1\}$. Then the decision rule for employment is

$$Y_{it} = \mathbb{1}\Bigg[\max_{C_t,I_t}\Big\{U(C_{it},I_{it};Y_{it}=1,H_{it},\xi_i,\zeta_{it}) + \beta S(t,H_{it})\mathbb{E}V_{t+1}(\Omega_{it+1}^1)\Big\} \tag{37}$$

$$- \max_{C_t,I_t}\Big\{U(C_{it},I_{it};Y_{it}=0,H_{it},\xi_i,\zeta_{it}) + \beta S(t,H_{it})\mathbb{E}V_{t+1}(\Omega_{it+1}^0)\Big\} \geq 0\Bigg]. \tag{38}$$

Denote (C_t^y, I_t^y) as solutions to equations (34) and (35). Then we can express this using a variable capturing the latent value working,

$$Y_t^* = \frac{(C_{it}^1 + \alpha I_{it}^1)^{1-\gamma} - (C_{it}^0 + \alpha I_{it}^0)^{1-\gamma}}{1-\gamma} - v(Y_{it},H_{it},\xi_i,\zeta_{it}) + \beta S(t,H_{it})\Big(\mathbb{E}V_{t+1}(\Omega_{it+1}^1) - \mathbb{E}V_{t+1}(\Omega_{it+1}^0)\Big),$$

and then optimal policy for employment is

$$Y_{it}^* = \mathbb{1}(Y_t^l > 0). \tag{39}$$

We can express the policy functions characterizing unhealthy good demand and labor supply as

$$I_{it}^{*} = I(H_{it}, W_{it}, A_{it}, t, \xi_i, \phi_i, \psi_i, \zeta_{it}|Y_{it}, \theta) \tag{40}$$

$$Y_{it}^{*} = Y(H_{it}, W_{it}, A_{it}, t, \xi_i, \phi_i, \psi_i, \zeta_{it}|\theta), \tag{41}$$

where θ is the set of all parameters in equations (29) - (33). This expression is useful to illustrate the determinants of health formation. Substituting equations (40) and (41) into equation (42) yields

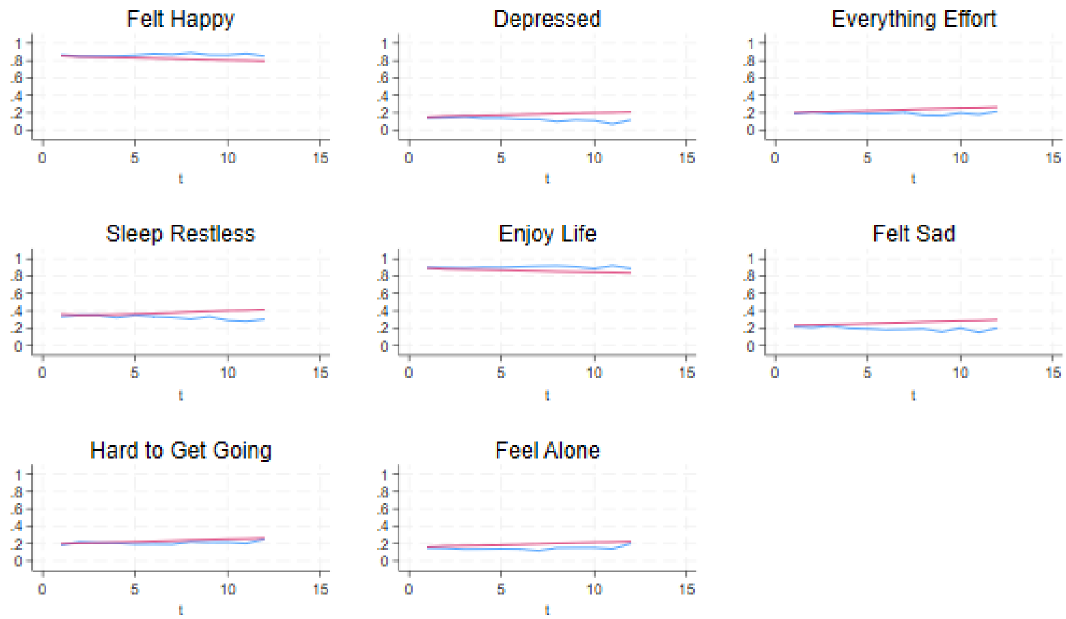
$$H_{it+1} = H(t, H_{it}, I_{it}^{*}, Y_{it}^{*}, \psi_i, \epsilon_{it+1}), \tag{42}$$

which clearly illustrates the endogeneity present when estimating (42) using observed data. The unobserved input to health formation, ψ_i, is correlated with observed optimal employment and demand for unhealthy goods. The correlation works through ψ_i directly and through its covariance with individual preference heterogeneity, ξ_i, and heterogeneity in ability, ϕ_i. Hence, controlling for ψ_i is needed to obtain unbiased estimates of the health transition process.

The model implements a uni-dimensional notion of health as the relevant state variable for the individual's decisions process. This approach is commonplace in the related literature, using self-reported health (French, 2005) or combining various measures into a single index (Bound et al., 2010). However, the model can readily accommodate multidimensional health by specifying H as an aggregator function for an underlying set of health components, for instance using a CES technology as in Cunha et al. (2010). The health components each evolve dynamically according to their own technology process, given state variables, and the nature of endogeneity arises in a similar way to the stylized model above. The next section considers the estimation of such a health formation process.

E Fit of the Measurement System

Figure 6: Fit of Measures of Mental Health

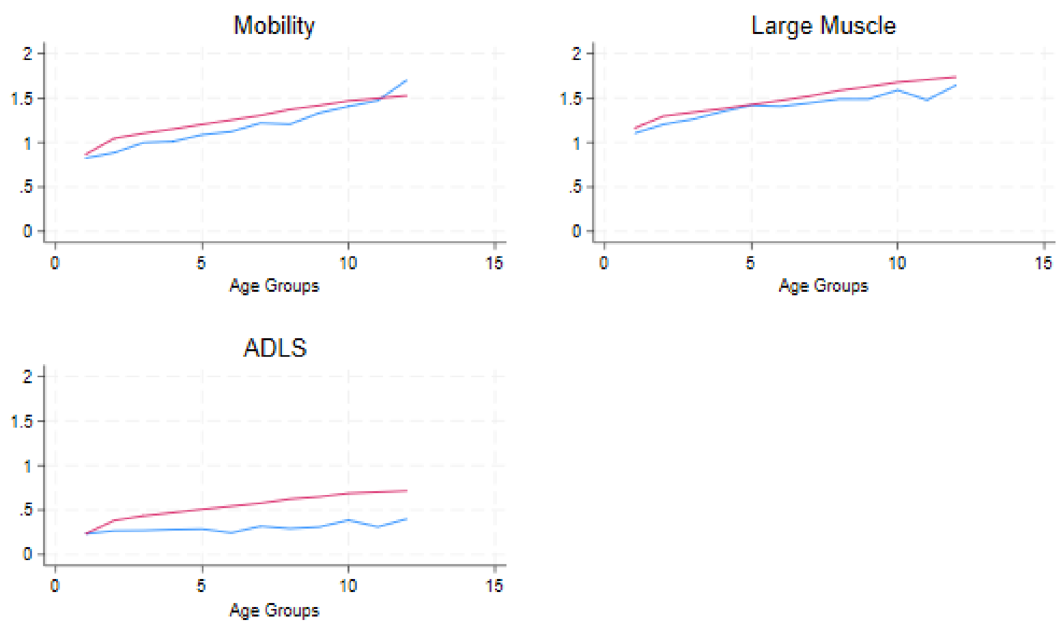


Note: ages are gouped into two-year bins. That is, age group 1 is 55-56 years old, age group 2 is 57-58 years old, etc. Red lines correspond to predictions from the simulated model and blue lines correspond to the observed data.

Figure 7: Fit of Measures of Cognitive Health

Note: ages are gouped into two-year bins. That is, age group 1 is 55-56 years old, age group 2 is 57-58 years old, etc. Red lines correspond to predictions from the simulated model and blue lines correspond to the observed data.

Figure 8: Fit of Measures of Physical Health



Note: ages are gouped into two-year bins. That is, age group 1 is 55-56 years old, age group 2 is 57-58 years old, etc. Red lines correspond to predictions from the simulated model and blue lines correspond to the observed data.