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

By Robert Millard* and Steven Stern†

Preliminary Draft [Please do not cite or circulate]

March 25, 2026

Abstract

Health in older adults deteriorates along multiple, interdependent dimensions, yet the empirical literature either aggregates these into a single index or treats health domains as separate processes, obscuring the dynamic pathways through which decline in one domain accelerates decline in others. This paper develops and estimates a structural dynamic model of health formation that explicitly allows for complementarities across three latent health domains: physical functioning, cognitive functioning, and mental health. The model also links these domains to employment, health investments, and economic conditions. We estimate the model using a comprehensive set of health measures from the Health and Retirement Study, recovering the transition processes governing each domain. The estimates reveal strong cross-domain complementarities: physical functioning exhibits substantial persistence and significantly shapes mental health trajectories, while mental health feeds back into both physical and cognitive outcomes. Counterfactual decompositions show that physical health is the primary determinant of survival and health inequality, while mental health plays an important role in late-life labor-market participation. Finally, we map the latent health domains to commonly used measures, such as self-reported health, demonstrating that each domain contributes meaningfully to individuals' self-assessments of their health.

*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.

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, 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, estimates of the effects of health are sensitive to both the measurement of health and the empirical methods used, both of which can vary greatly across empirical studies (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 consider a tractable yet flexible model of health formation. The model captures the complexity of interactions 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 domains. We model the dynamic interdependence among the domains of health to identify both self- and cross-complimentarities in shaping future health outcomes. Moreover, the latent health domains 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 illustrate novel 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 domains 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 domains are categorized as 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 domains 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;

[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).

Capatina and Keane, 2023; Millard, 2025). Second, there is an emerging economic literature documenting the considerable consequences of mental health challenges in the labor market (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.

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 domains 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 domains, which measure the degree of complementarity across domains 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 domains 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 ages. In the current draft, we adopt a flexible Translog production 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]

The estimates reveal strong complementarities across the three health domains and highlight the central role of physical health in the health formation process. Physical health exhibits the greatest persistence and exerts a substantial influence on both mental health and economic outcomes. While employment improves cognitive health, it is associated with declines in physical and mental health, suggesting that continued work at older ages may place strain on overall well-being. Health investments improve outcomes across all domains, with particularly strong effects on mental health. The model also shows that physical health is the most important determinant of employment and mortality, although mental health also plays a meaningful role in labor-market participation.

Counterfactual decompositions and shock experiments further highlight the asymmetric structure of health dynamics. Physical health operates as a central hub: shocks to physical health propagate broadly across other health domains, reduce employment, and increase mortality, while shocks to mental health generate meaningful spillovers primarily through their effects on physical health. In contrast, shocks to cognitive health are more localized and have comparatively smaller downstream effects. These results also show that heterogeneity in physical and mental health is a primary driver of dispersion in overall health and related economic outcomes, while commonly used summary measures such as self-reported health obscure the distinct mechanisms through which different health dimensions affect behavior and inequality at older ages.

Furthermore, our analysis provides insight into how objective health conditions and an individual's self-perceived health relate to latent health domains. 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 in this way 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 will be a weighted average of the estimated effects of physical, cognitive, and mental health.

This paper makes several novel contributions to the related literature. We fit into a large literature aimed

[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.

at better 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.[3] 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 (Butler et al., 1987; Benítez-Silva et al., 2004; French, 2005; Kreider and Pepper, 2007; Meyer and Mok, 2019), objective measures (Bartel and Taubman, 1979; Bound, 1989; Smith, 2004), and their combinations (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; 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 all heterogeneity into a single dimension.[4]

Our model fits among more recent multidimensional frameworks (Conti et al., 2010; Amengual et al., 2021; Wen, 2022; 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 the dimensions of health. 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 domains, 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

[3]For further discussion see Currie and Madrian (1999); O'Donnell et al. (2015); French and Jones (2017), and Blundell et al. (2023)

[4]Univariate measures (e.g., self-reported health) have documented empirical content in that 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).

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 (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 domains and the interdependencies across the domains. 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 domains, our framework mitigates an omitted variable bias present when narrowing ones focus to estimate the effect of one domain 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 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.[5] 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 domains, yielding intuition about which health domains 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 (Jolivet and Postel-Vinay, 2020; Biasi et al., 2021). Our analysis complements this

[5] For instance, see Bartel and Taubman (1979); Kahn (1998); Moran et al. (2011); Stephens et al. (2018).

perspective by investigating the economic environments that aggravate mental health challenges (Adhvaryu et al., 2019; 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 model, estimation, and identification. Section 3 describes the data used for the analysis. Section 4 reviews the main estimation results. Section 5 and 6 conduct counterfactual exercises using the model. Finally, Section 7 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 A of the Appendix. At each age, t, health is represented as a vector of stocks, $x_t = (x_t^p, x_t^c, x_t^m)'$, where $\{p, c, m\}$ corresponds to physical, 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, $x_0 = (x_0^p, x_0^c, x_0^m)'$. 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.[6] 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 domain 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

[6] 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).

oldest age we observe an individual. The dynamic formation of each health domain, x_t^k, is determined by

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

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 $k \in \{p, c, m\}$, x_t^k is determined by $x_{t-1} = (x_{t-1}^p, x_{t-1}^c, x_{t-1}^m)'$, capturing both self- and cross-dependencies in health formation across the latent health domains. We assume that only health stocks from the preceding period contribute to the formation process, suggesting that prior shocks to any health domain impacts only the current period's health through the previous period.[7]

The transition process depends on the previous period's investments into health, which encompass both healthy behaviors, such as exercise, and unhealthy behaviors, such as smoking. We summarize the set of health investment behaviors using a single variable I_t. 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 initial heterogeneity,

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

Additional controls, W_t^I, encompass observable yet unmodeled determinants of health investment decisions.

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

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

where the value of employment depends on the stocks of latent health, x_t, age, and latent heterogeneity, b_0. Furthermore, additional controls, W_t^L, 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

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

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).[8] 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

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

where $u_t^k = (u_t^p, u_t^c, u_t^m)$ are assumed to be mutually independent idiosyncratic shocks to health formation that 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 the omitted variables problem.[9]

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 domains. 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

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

[9] This identification argument follows from the solution to the theoretical life-cycle consumption and labor supply model described in Section A 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.

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 domain, we have a set of measures that are generated exclusively by that domain. For each $k \in \{p, c, m\}$ 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).[10] Typically, data on health outcomes encompass three distinct types of measurements: continuous, categorical, and binary.

1. For continuous measures, we adhere to the prevailing literature and assume a semi-log linear mapping,

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

   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 $z^{k*}_{t,m} = ln(x^k_t) + v^k_{t,m}$, where

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

   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

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

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

[10] 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 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.

Moreover, we have a set of observed measures, z_m^b, $m \in \{1, ..., M^b\}$, which we assume are exclusively generated by latent initial heterogeneity, b_0. These measures are based on individual characteristics prior to model entry, including their 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 domains, 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.

2.2 Implementation

The model described above is written in general form. Below we impose structure on the functional forms of the health formation process and the other behavioral equations in the model.

Initial Conditions

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 $w_0 = (\ln x_0^p,\ \ln x_0^c,\ \ln x_0^m,\ b_0)'$ denote the vector of initial conditions. We approximate the distribution of w_0 using a two component Gaussian mixture. Let F_{w_0} denote the joint distribution of all log latent factors in the model in period $t = 0$. Then,

$$F_{w_0} = \pi\Phi(\omega^1, \Sigma^1) + (1 - \pi)\Phi(\omega^2, \Sigma^2), \tag{6}$$

where $\pi \in (0, 1)$ is a common mixture weight, $\Phi(\omega^j, \Sigma^j)$ is a CDF of a normal distribution with mean vector ω^j and Σ^j is a 4x4 variance-covariance matrix.

The location and scale of latent factors is not identified without further normalization on either F_{w_0} or the measurement system.[11] First, we assume the vector of initial conditions is mean zero, $E(w_0) = 0$. Second, we assume that the diagonal of the variance-covariance matrix of initial endowments is one on along the diagonal. These normalizations anchor the location and scale of the latent variables and eliminate observational equivalence between latent variances and measurement loadings. Off-diagonal elements of the component-specific variance–covariance matrices are left unrestricted and capture cross-domain dependence at baseline.

[11] In dynamic latent factor models, some normalization is required. An alternate approach can be to normalize parameters of the measurement system, as in Cunha et al. (2010).

Health Formation Technology

We consider a flexible translog specification of the health production technology, which allows for nonlinearities, interactions across health domains, and dynamic complementarities in health production. In addition, the translog specification has desirable properties for identifying treatment effects in dynamic factor models of the type considered here (Del Bono et al., 2022). Define the vector $w_t = (\ln x_t^p, \ln x_t^c, \ln x_t^m, b_0, I_t)'$. Then the transition equation for each health domain k is then given by

$$\ln x_{t+1}^k = a_0^k + w_t' \beta^k + w_t' B^k w_t + c_e^k L_t + \nu_t^k \tag{7}$$

where $\beta^k \in \mathbb{R}^4$ and B^k is a symmetric matrix governing second-order terms and interactions.

Employment, Mortality, and Health Investments

We use a nonstructural approximating model for the health investment and employment policy functions. Health investment decisions, I_t, reflect choices made conditional on health. The policy function for health investments is assumed linear in the health domains,

$$\begin{aligned} I_t &= \beta_0^I + \beta_1^I \ln x_t^p + \beta_2^I \ln x_t^c + \beta_3^I \ln x_t^m + \beta_4^I b_0 + \beta_5^I t + \nu_t^I, \qquad (8) \\ \nu_t^I &\sim iid\ \mathcal{N}(0, \sigma_I^2). \end{aligned}$$

The employment decision is a discrete choice that is modeled using a latent variable framework. The latent value of work, L_t^*, is assumed linear in its arguments,

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

The value of employment varies by health status, for instance if health makes it more costly to perform productive tasks, or lowers productivity and subsequent returns to working. We omit additional control variables, W_t^L and W_t^I, in the current draft of the paper. Further, we assume that ν_t^k, ν_t^I, and ν_t^e are mutually independent for all $k \in \{p, c, m\}$, and $t \in \{1, ..., T_i\}$.

In a similar manner, for individuals that have survived to $t-1$, the probability they die in period t

$(d_t = 1)$ is determined by a latent variable, d_t^*, where

$$d_t^* = \beta_0^s + \beta_1^s \ln x_t^p + \beta_2^s \ln x_t^c + \beta_3^s \ln x_t^m + \beta_4^s b_0 + \beta_5^s t + \nu_t^s, \tag{10}$$
$$\epsilon_t^s \sim iid\, \mathcal{N}(0,1),$$
$$d_{it} = \mathbb{1}\{d_t^* \geq 0\}.$$

Finally, we estimate the model using an unbalanced panel, where sample attrition is prevalent. We model sample attrition as endogenous to one's health.[12] 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

$$a_t^* = \beta_0^a + \beta_1^a \ln x_t^p + \beta_2^a \ln x_t^c + \beta_3^a \ln x_t^m + \beta_4^a b_0 + \beta_5^a t + \nu_t^a, \tag{11}$$
$$\epsilon_t^a \sim iid\, \mathcal{N}(0,1),$$
$$a_{it} = \mathbb{1}\{a_t^* \geq 0\}.$$

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

2.3 Identification of Health Formation Technology

Identifying and estimating the technology function is challenging, as both inputs and outputs are only observed through noisy 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. $v_{t,m}^k$ are mean zero for all $k \in \{p, c, m, b\}$, $t \in \{1, .., T_i\}$, and $m \in \{1, ..., M^k\}$. The variance of measurement errors for discrete and categorical measures is normalized to one.

2. $v_{t,m}^k$ is independent of $(x_\tau^p, x_\tau^c, x_\tau^m, b_0)$ for all $t, \tau \in \{1, .., T_i\}$; $m \in \{1, ..., M^k\}$; and $k \in \{p, c, m, b\}$.

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

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

The parameters to identify involve the parameters of the measurement system $\{\mu_m^k, \lambda_m^k, \tau_{j,m}^k, \sigma_{v_m^k}\}_{k\in\{p,c,mh,b\},m\in 1,\ldots,M_k}$, the distributional parameters of initial conditions $\{\omega^1, \Sigma^1, \omega^2, \Sigma^2, \pi\}$, the parameters of the transmission process $\{a^k, \beta^k, B^k, c_e^k, \sigma_{v^k}^2\}_{k\in\{p,m,c\}}$ and parameters on the nonstructural equations for employment, mortality, attrition, and healthy behaviors, $\{\beta_j^I, \beta_j^L, \beta_j^s, \beta_j^a\}_{j=1,\ldots,5}$. Identification of the model's parameters is described in two steps. First, we identify the distribution of the initial endowments and parameters of the measurement system. Given these, the second step recovers the remaining model parameters.

Initial Distributional Parameters and Measurement System

In the first step, we identify the joint distribution of initial latent endowments and the parameters of the measurement system using baseline observations. Since measurement parameters are assumed age-invariant, the first observed period contains sufficient information for both objects.

Consider a single latent factor, x_0^k, $k \in \{p, c, m, b\}$, that follows a two-component Gaussian mixture

$$\ln x_0^k \sim \pi\mathcal{N}(\omega_k^1, s_k^1) + (1-\pi)\mathcal{N}(\omega_k^2, s_k^2)$$

with mean and variance normalized to zero and one, respectively, leaving three free mixture parameters (ω_k^1, s_k^1, π). Suppose x_0^k is measured by $M \geq 3$ binary indicators generated exclusively by that factor. This sub-problem is to identify M threshold parameters of the binary measures and three distributional parameters, giving $M+3$ free parameters. Under conditional independence of measurement errors, all dependence across the M measures arises from the shared latent variable. First, the set of marginal probabilities, $Pr(z_{0,m}^k = 1)$ provide M moments linking the thresholds and mixture parameters, while the pairwise joint probabilities, $Pr(z_{0,m}^k = 1, z_{0,l}^k = 1)$, $m \neq l$, supply an additional $M(M-1)/2$ moments to identify the within-component factor variance, as the correlation between any two binary indicators within a component is determined solely by the signal-to-noise ratio, $s_k^j/(s_k^j + 1)$.[13] With three free mixture parameters after normalization and M threshold parameters, the system is generically overidentified for $M \geq 3$, and the mixture distribution and thresholds are identified up to label switching, which is resolved by an ordering restriction on component means. Identification of finite Gaussian mixtures from conditionally independent discrete proxies in this setting is established by Allman et al. (2009).

In our application, some latent factors are measured using a combination of binary, categorical,

[13]The within-component correlation between any two binary indicators is determined entirely by how much of their variation is driven by the shared latent factor relative to idiosyncratic measurement error. The stronger the common factor relative to noise, the more tightly correlated any two indicators will be within a component. Since measurement errors are independent across indicators by assumption, this pairwise co-movement has no other source, and so the pairwise probabilities directly pin down the variance of the latent factor within each mixture component.

and continuous indicators. Categorical measures are treated as ordered threshold crossings and can be decomposed into binary comparisons, so the same identification logic applies. Once the mixture distribution of the latent factor is identified using discrete measures, parameters of continuous measures (locations, loadings, and error variances) are identified from means, variances, and covariances of the observed outcomes.

Finally, cross-domain co-movement in observed measures identifies the off-diagonal elements of the component covariance matrices Σ^j. Since measurement errors are assumed independent across domains conditional on w_0, any cross-domain correlation in observed proxies is attributable to dependence between latent factors. With a sufficiently rich baseline measurement system, the full joint distribution of initial endowments is therefore identified. Further detail is provided in Appendix B.

Identification of the Transition Process and Remaining Model Parameters

Conditional on the parameters of the measurement system and the joint distribution of initial latent endowments identified in Step 1, the model implies a well-defined joint distribution of latent health stocks and heterogeneity in every period. Because the measurement system is known and measurement errors are independent of the latent states, the evolution of the observed measures over time reflects only the transition dynamics of the latent factors and the stochastic shocks governing those transitions.

To exploit this structure, we summarize the information in the observed binary measures using principal components constructed separately for each latent health domain $k \in \{p, c, m\}$ and for the time-invariant heterogeneity, b_0, using only measures that load exclusively on each domain. Given the identified measurement thresholds and PCA weights, the first principal component, PC_t^k, provides a quasi-continuous auxiliary statistic whose conditional expectation is a strictly increasing function of the corresponding latent factor

$$E(PC_t^k | x_t^k) = h(x_t^k).$$

Although the mapping from the latent factor to the index is generally nonlinear due to thresholding in the measurement system, the index preserves the ordering and dependence of the latent state and therefore serves as a monotone proxy for $\ln x_t^k$.

The transition process is identified using an auxiliary regression that characterizes the conditional mean of PC_{t+1}^k as a function of current PCA indices and state variables, estimated jointly across domains. The joint specification with pairwise interaction terms is essential: cross-domain parameters, which govern the complementarities at the center of our analysis, are identified only from the co-movement of indices across domains. Estimation of this auxiliary model yields reduced-form coefficients $\beta^A(\theta)$ that depend on the structural parameters governing the latent transition process. Structural parameters such as location, persistence, and innovation variance affect distinct observable features of the PCA indices; their mean, serial correlation, and dispersion, so that different parameter values generate different auxiliary coefficients.

Identification follows if the mapping from structural parameters to auxiliary coefficients is injective

$$\beta^A(\theta_1) = \beta^A(\theta_2) \Rightarrow \theta_1 = \theta_2$$

In practice, this condition can be verified numerically by checking the Jacobian $\partial\beta^A(\theta)/\partial\theta$ has full rank, following the indirect-inference framework of Gourieroux et al. (1993).

Finally, conditional on the identified latent dynamics, the remaining behavioral equations (health investment, employment, mortality, and attrition) are identified from their conditional relationships with the PCA indices and observed covariates, since the indices are monotone transformations of the latent health factors. Further details of this identification argument can be found in Appendix Section B.

2.4 Estimation

We estimate the model parameters using a two-step procedure. In the first step, we recover the distributional parameters of the initial latent endowments and the parameters of the measurement system using data from the first observed period. We employ simulated maximum likelihood to estimate the mixture distribution governing w_0 and the parameters of the measurement equations. This step anchors the scale and dependence structure of the latent factors and fully characterizes the mapping from latent states to observed measures. Details of the likelihood function and the simulation procedure are provided in the Appendix B.

In the second step, conditional on the parameters identified in the first step, we use minimum-distance moment matching to identify the remaining model parameters. The key moments are empirical analogues of the model's production functions where the latent factors are replaced by domain-specific indices that are obtained from principal components applied to blocks of measures that load exclusively on a single latent factor. Substituting these indices into the transition equations yields reduced-form relationships whose conditional means and variances are functions of the underlying production-function parameters and shock variances. Intuitively, the indices act as sufficient statistics for the latent factors: because they preserve the ordering and dependence structure of the latent states, their conditional moments capture the same information about the transition dynamics as the latent factors themselves.

In practice, the moments we match correspond to the OLS estimates of Equation (7), where the latent variables are replaced by the corresponding principal-component indices. The same indices are also used to construct empirical analogues of the behavioral equations governing investment, employment, mortality, and attrition. For investments, we match the OLS estimates of Equation (31), again replacing latent health variables with the principal-component indices. For employment, mortality, and attrition, we match OLS estimates from linear probability models that use the principal-component indices in place of the latent health domains.

Finally, we include moments corresponding to the age specific means of each observed health measure, employment, investments, mortality and attrition. These moments help ensure the model reproduces the relevant marginal distributions in the data. Further details on estimation can be found in Appendix section C.

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 2,163 individuals (13,456 total observations).

Table 1 shows the number of observations 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 attrition and right censoring.

Next, we describe the various sets of health measures used to recover the distribution of the latent health domains.

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

Table 1: Sample Counts

Age Group	Observed	Dead	Attrition	Right Censoring
1	2,163	0	0	0
2	2,160	3	0	0
3	1,963	11	127	62
4	1,680	31	220	232
5	1,422	41	317	383
6	1,166	58	393	546
7	932	72	474	685
8	720	94	539	810
9	520	103	608	932
10	373	112	645	1033
11	259	119	678	1107
12	98	126	716	1223
Totals:	13,456	770	4,717	7,013

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

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 seven 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)).[14] 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

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

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. Additionally, they self-report their memory on a five-point scale (scale is from poor to excellent). 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 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.

The HRS groups functional limitations into three areas of physical limitation, based on similarity of the nature of limitation. First, mobility limitations include: walking several blocks, walking one block, walking across a room, climbing several flights of stairs, and climbing one flight of stairs. Second, large muscle limitation includes four tasks: sitting for two hours, getting up from a chair, stooping, kneeling or crouching, and pushing or pulling a large object.

Employment and Health Investments

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.[15]

[15] 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.

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. We use this to construct a binary indicator for the extensive margin of drinking, and another indicator relating to excessive drinking, which is defined as drinking more than 4 alcoholic drinks per day when the respondent drinks. Last, we have an indicator variable of whether the respondent currently smokes cigarettes.

We use these four measures to construct a composite measure of health investments. We extract the first principal component of these four measures and treat this as an observable measure of overall healthy investments.

Measures of Latent Heterogeneity

Respondents enter the sample at age 55, implying 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 education and 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 the 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 the factor analytic framework to recover distribution 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 domains. 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.[16] In our analysis, we use respondents' self-rated health, measured on a five-point categorical scale where "5" indicates excellent health.

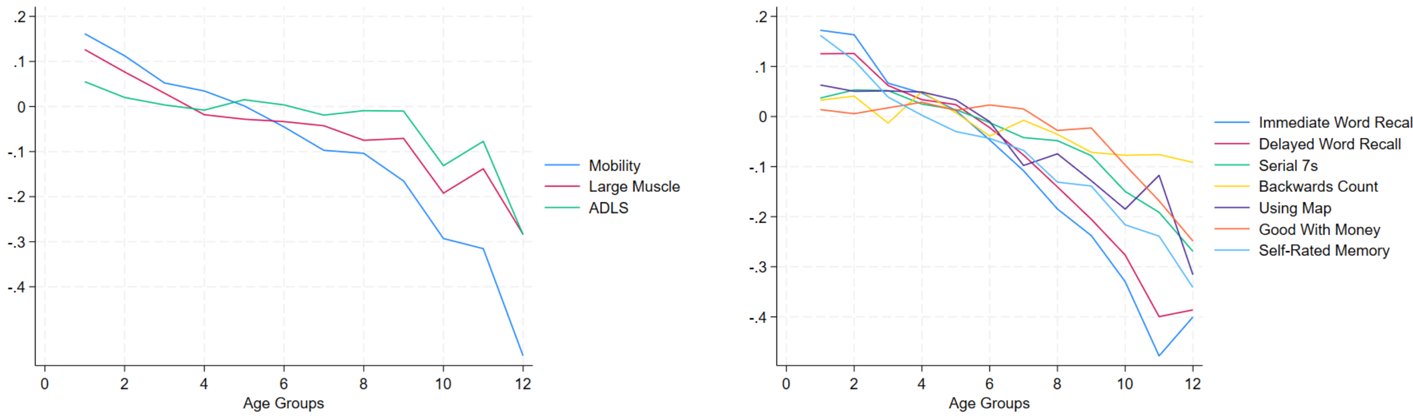
Second, we incorporate a set of objective health indicators, capturing whether the respondent has ever been diagnosed with high blood pressure, diabetes, cancer, lung disease, heart disease, stroke, psychiatric problems, or arthritis. Linking these diagnoses to the latent health domains reveals how specific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

3.1 Descriptive Statistics

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) and cognitive health measures (right). To facilitate comparison, each measure is standardized and signed so that higher values indicate "better" health.

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

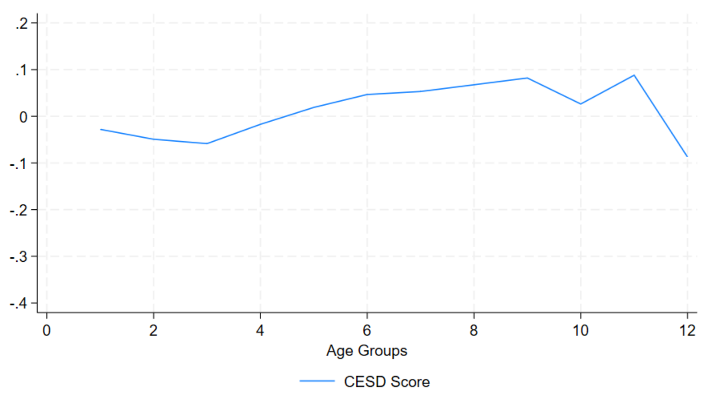
Figure 1: Mean of Standardized Health Measures by Age



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..

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 age trend in the overall CESD score, constructed by aggregating responses across all CESD questions to form the CESD index. The figure reports the index using the same standardization and sign convention as in the previous health measures. In contrast to the age profiles for physical and cognitive health, mental health follows a concave pattern whereby it improves at earlier ages and then gradually declines after approximately ages 67–68.

The improvement in mental health prior to retirement is consistent with evidence that the life-cycle profile of mental health is U-shaped (Dijk and Mierau, 2023). However, the subsequent decline may reflect the accumulation of health deficits at older ages. These patterns suggest that mental health follows a distinct life-cycle trajectory relative to physical and cognitive health, raising the question of how changes in mental health interact with other health domains in shaping economic behavior at older ages

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

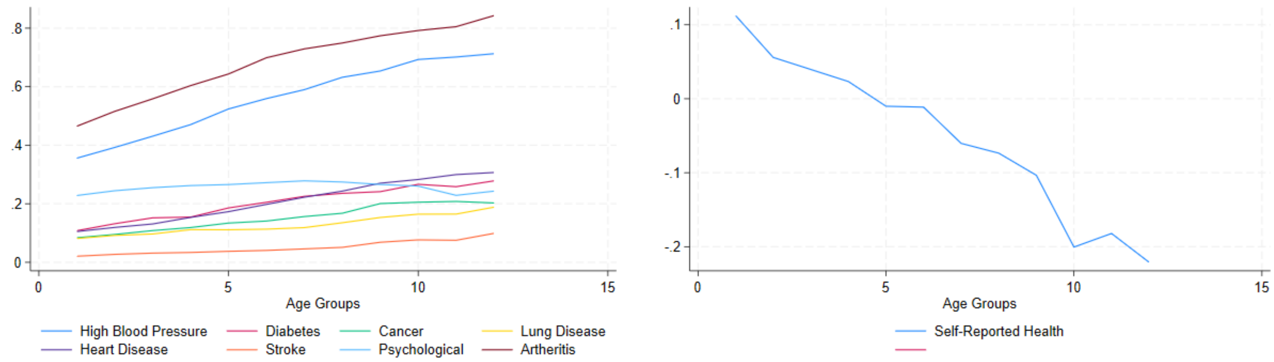
Note: ages are grouped into two-year bins. That is, age group 1 is 55-56 years old, age group 2 is 57-58 years old, etc..

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 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

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

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.

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

This section presents the parameter estimates from the empirical model and discusses their implications for health dynamics and economic behavior. The fit of the simulated moments relative to their data counterpart is reported in Appendix Section B.

Table 3 presents the estimated covariance matrix of the initial endowments, x_0^c, x_0^m, x_0^p, and b_0. The endowments of the three health domains are positively correlated, with the strongest association between

Table 3: Model Estimates: Covariance Matrix of Initial Endowments

	$\ln x_0^c$	$\ln x_0^m$	$\ln x_0^p$	b_0
$\ln x_0^c$	1.000	0.432	0.377	0.865
$\ln x_0^m$		1.000	0.554	0.472
$\ln x_0^p$			1.000	0.434
b_0				1.000

Note: Standard errors presented in parentheses below point estimates.

physical and mental health. Moreover, initial heterogeneity positively correlates with each health domain, indicating that individuals with higher endowments of b_0 tend to exhibit better health outcomes. These correlations indicate that favorable initial conditions tend to span multiple health domains, consistent with the presence of common underlying determinants of health at older ages.

Table 4: Conditional Mean of Latent Heterogeneity, b_0

	Education		Mother's Education		SES
None	-1.202	None	-0.628	Poor	-0.839
Above Bachelors	1.182	College +	1.688	Rich	1.192
	Self Reported Health in Childhood		**Ever-Smoked**		**Parents Drink/ Drug**
Poor	-1.881	Yes	-0.496	Yes	-0.933
Excellent	0.418	No	0.505	No	0.257

Table 4 presents the mean of b_0 conditional on selected realizations of the measures used to recover its distribution, providing intuition about the types of characteristics captured by this variable. Initial heterogeneity is measured using retrospective information on individuals and their parents. The statistics indicate that b_0 is positively correlated with both the individual's and their mother's education, as well as with childhood socioeconomic status and self-reported health during childhood. In contrast, 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. Taken together, these correlations suggest that b_0 captures persistent heterogeneity related to early-life human capital and health environments.

The results for initial endowments 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 reports the estimates of the health formation process. The columns correspond to cognitive, mental, and physical health, respectively. The first row reports the intercept, indicating that cognitive and physical health decline with age, while the rate of decline in mental health is significantly smaller. Rows 2-4 capture both own-lag and cross-lag effects. The diagonal elements show that each health domain exhibits substantial own-persistent, with the strongest autocorrelation observed in physical health (0.858). The cross-lag estimates reveal an asymmetric pattern of interdependence: physical health exerts a positive and significant effect on subsequent mental health (0.084), while the reverse and cognitive cross-effects are smaller in magnitude. The positive coefficient on $(\ln x_t^p)^2$ in the mental health equation (0.038) indicates that this effect is nonlinear.

Latent heterogeneity b_0 positively affects the formation of cognitive (0.038) and mental health (0.070), indicating that individuals with favorable early-life endowments experience a slower rate of decline in these domains. The effect on physical health formation is small and imprecisely estimated (-0.009), suggesting that the long-run influence of early-life factors operates primarily through cognitive and mental rather than physical pathways.

The interaction terms reveal three economically distinct forms of complementarity. First, physical and mental health are mutually reinforcing in their own formation, with positive cross-domain interactions implying that decline in one domain accelerates decline in the other. Second, mental health formation exhibits convexity in both own and physical health stocks, meaning individuals with higher health levels accumulate further advantages, which is a mechanism that generates divergence in mental health trajectories over time. Third, early-life endowments partially substitute for current health stocks in sustaining mental well-being, suggesting that the long-run protective effect of favorable initial conditions operates through reducing sensitivity to current health shocks rather than through direct level effects alone.

Next, consider the effect of lagged employment on health formation. Working in the previous period improves subsequent cognitive health, consistent with evidence showing adverse health effects associated with labor-market exit (Fitzpatrick and Moore, 2018; Black et al., 2018). The negative effects on mental (-0.030)

Table 5: Model Estimates: Health Formation Process

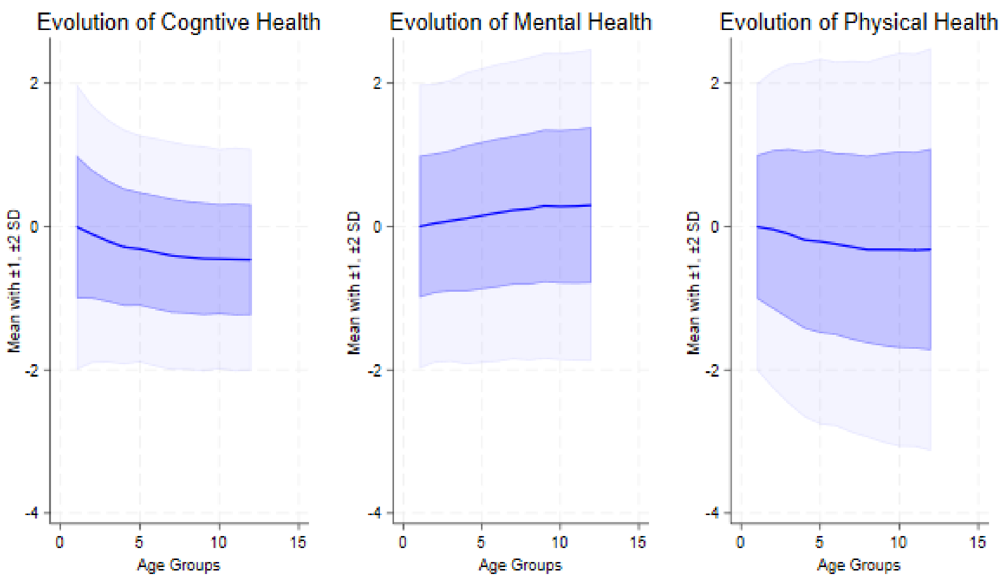
	(1) $\ln x_{t+1}^c$	(2) $\ln x_{t+1}^m$	(3) $\ln x_{t+1}^p$
Intercept	-0.113	-0.003	-0.116
	(0.001)	(0.000)	(0.001)
$\ln x_t^c$	0.748	-0.071	0.068
	(0.001)	(0.003)	(0.007)
$\ln x_t^m$	0.026	0.671	0.056
	(0.002)	(0.006)	(0.006)
$\ln x_t^p$	0.024	0.084	0.858
	(0.002)	(0.007)	(0.004)
b_0	0.038	0.070	-0.009
	(0.002)	(0.003)	(0.006)
$(\ln x_t^c)^2$	-0.034	0.001	-0.001
	(0.001)	(0.000)	(0.000)
$(\ln x_t^m)^2$	0.002	0.035	0.020
	(0.000)	(0.001)	(0.002)
$(\ln x_t^p)^2$	0.000	0.038	0.004
	(0.000)	(0.001)	(0.000)
$(b_0)^2$	0.005	0.000	-0.001
	(0.000)	(0.000)	(0.000)
L_t	0.023	-0.030	-0.011
	(0.001)	(0.002)	(0.001)
$\ln x_t^c \times \ln x_t^m$	-0.001	-0.027	-0.003
	(0.000)	(0.001)	(0.000)
$\ln x_t^c \times \ln x_t^p$	-0.001	-0.005	0.010
	(0.000)	(0.001)	(0.003)
$\ln x_t^m \times \ln x_t^p$	0.002	-0.005	0.020
	(0.000)	(0.001)	(0.003)
$b_0 \times \ln x_t^c$	0.001	0.003	0.000
	(0.000)	(0.000)	(0.000)
$b_0 \times \ln x_t^m$	-0.003	-0.018	-0.001
	(0.000)	(0.002)	(0.000)
$b_0 \times \ln x_t^p$	0.002	-0.018	0.002
	(0.000)	(0.003)	(0.000)
I_t	0.011	0.047	0.001
	(0.002)	(0.002)	(0.000)
$I_t \times \ln x_t^c$	-0.007	0.033	0.027
	(0.001)	(0.003)	(0.003)
$I_t \times \ln x_t^m$	-0.001	-0.023	-0.005
	(0.000)	(0.003)	(0.000)
$I_t \times \ln x_t^p$	0.004	-0.014	-0.022
	(0.000)	(0.003)	(0.002)
$I_t \times b_0$	-0.019	0.007	0.000
	(0.001)	(0.001)	(0.000)
σ_v^2	0.416	0.691	0.665
	(0.004)	(0.001)	(0.004)

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

and physical health (-0.011) suggest that continued work at older ages carries costs along these dimensions, though we note that the joint estimation accounts for the endogeneity of employment with respect to current health status. Health investments improve outcomes across all domains, with the largest effect on mental health (0.047), consistent with the role of lifestyle behaviors in sustaining mental well-being.

Figure 5: Simulated Path in Latent Health Domains



Note: Figures plot mean path in simulated latent health factors over time. The figures also show one and two standard deviations in the shaded regions.

Figure 5 plots the mean paths of the health domains using data simulated from the estimated model. The simulated paths for cognitive and physical health follow age trends similar to those observed in the raw health measures discussed in Section 3.1. Likewise, the model reproduces the positive age trend in mental health observed in the aggregated CESD score. In terms of dispersion, physical and mental health become more dispersed over time, whereas cognitive health becomes less dispersed. This findings suggests that heterogeneity in physical and mental health accumulates with age, while cognitive health trajectories become more compressed over time

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

Despite an 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

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

	Employment	Investment	Mortality	Attrition
Intercept	1.052	0.000	-2.758	-1.955
	(0.004)	(0.000)	(0.022)	(0.005)
$\ln x_t^c$	0.035	0.059	-0.018	-0.015
	(0.005)	(0.007)	(0.002)	(0.002)
$\ln x_t^m$	0.096	0.028	-0.017	0.010
	(0.007)	(0.002)	(0.003)	(0.002)
$\ln x_t^p$	0.279	0.112	-0.162	0.013
	(0.008)	(0.005)	(0.017)	(0.001)
b_0	0.064	0.188	0.005	0.029
	(0.009)	(0.007)	(0.000)	(0.006)
t	-0.270	0.009	0.087	0.089
	(0.001)	(0.001)	(0.001)	(0.001)

Note: Standard errors presented in parentheses below point estimates.

utilized to evaluate these effects. Table 6 estimates of the behavioral equations governing employment, health investment, mortality, and attrition. Across all equations, the latent health domains enter with the expected signs: better health is associated with higher employment, greater health investment, lower mortality, and lower attrition.

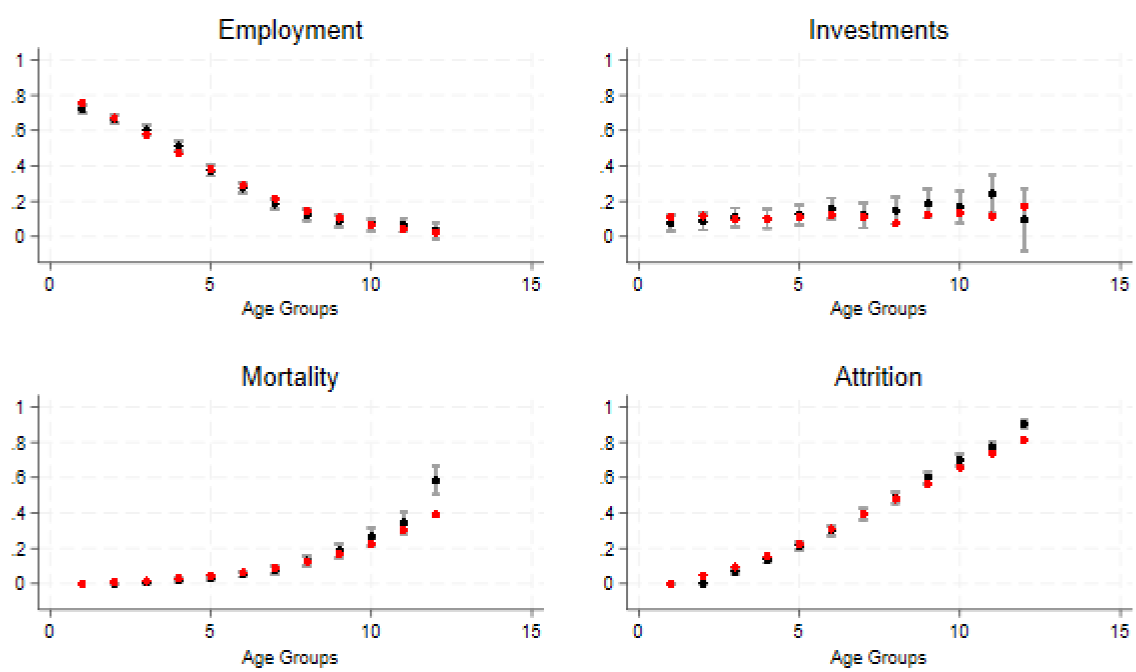
Physical health is the dominant predictor of employment (0.279), with mental health contributing a sizable additional effect (0.096) and cognitive health a more modest one (0.035). The relative magnitudes suggest that physical capacity is the primary constraint on labor supply near retirement, but that mental well-being constitutes a quantitatively important secondary channel. This finding is consistent with Blundell et al. (2023), who document modest cognitive effects on employment once physical health is controlled for. These findings lend support to the use of physical health measures as the primary indicators for analyzing the effects of health on employment. However, mental health plays a sizable role as well. Moreover, Table 5 indicates important complementary between one's physical and mental health. Ignoring the relationship between health domains can result in potentially biased estimates of health's employment effects.

In the health investment equation, b_0 has the largest coefficient (0.188), substantially exceeding the effects of any current health domain. This suggests that investment behavior is more strongly shaped by persistent early-life endowments, such as capturing education, socioeconomic background, and health habits, than by current health status. Physical health nonetheless has the largest effect among the current health stocks (0.112), consistent with physical capacity enabling health-promoting behaviors such as exercise.

In the mortality equation, physical health again dominates, with a coefficient of -0.162 that is roughly nine times larger than those on cognitive (-0.018) and mental health (-0.017). Mortality risk increases significantly with age (0.087), consistent with the well-documented exponential rise in mortality rates in this age range. The small positive coefficient on b_0 (0.005) reflects a conditional association, that holding current health stocks fixed, individuals with higher early life endowments face a marginally higher mortality risk.

Finally, the attrition equation indicates that individuals with better cognitive health are less likely to attrit from the panel (-0.015), while mental health has a small positive association with attrition (0.010). Physical health has a near-zero effect. The positive time trend (0.089) confirms that attrition increases with age in this sample, as expected. These patterns inform the interpretation of the panel estimates and motivate the joint modeling of attrition alongside the health transition process.

Figure 6: Model Fit of Behavioral Functions



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 6 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. Overall, we see that the model fits the overall age-trajectories of the observed data fairly well. The simulated data starts to miss mortality and attrition at older ages, which may be related to increased right-censorship lowering sample counts and moment precision at older ages.

Objective Health Measures and Self-Reported Health

Table 7: Objective Health Conditions

	High Blood Pressure	Diabetes	Cancer	Lung Disease
$\ln x^c$	-0.0575	-0.1266	0.0153	-0.0468
$\ln x^m$	-0.0199	-0.0457	-0.0001	-0.0738
$\ln x^p$	-0.0828	-0.2733	-0.0544	-0.4048
b_0	0.0365	0.0336	-0.0085	-0.0009
Age	0.0234	0.0083	0.0584	0.0116
Intercept	-0.1781	-0.9693	-1.0494	-1.4468
	Heart Problem	**Stroke**	**Psychological**	**Arthritis**
$\ln x^c$	-0.0021	-0.1888	-0.0434	0.0140
$\ln x^m$	-0.0487	-0.0602	-0.2316	-0.0073
$\ln x^p$	-0.2335	-0.1045	-0.1552	-0.2535
b_0	0.0036	0.0363	0.0214	0.0127
Age	0.0618	0.0602	0.0084	0.0628
Intercept	-1.1897	-2.1170	-0.6396	-0.2122

Note: Bootstrap standard errors are not available at this time.

The final set of results examine how 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, are generated by the latent health vector. The estimated parameters reveal the extent to which each latent health domain 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 domain determines an individual's self-perception of their overall health.

Table 7 presents marginal effects of the latent health domains on the probability of eight doctor-diagnosed conditions. The probability of each condition increases with age, with particularly steep age gradients for cancer (0.058), heart problems (0.062), stroke (0.060), and arthritis (0.063). Physical health is the primary determinant of most conditions, with the largest effects on lung disease (-0.405), diabetes (-0.273), arthritis (-0.254), and heart problems (-0.234), consistent with functional physical capacity being closely linked to chronic disease burden. Two exceptions are noteworthy: cognitive health is the dominant predictor of stroke (-0.189 vs. -0.105 for physical health), and mental health is the primary determinant of psychological conditions (-0.232 vs. -0.155 for physical health), confirming that the latent domains capture distinct clinical constructs. Cancer shows weak associations with all latent domains, likely reflecting the heterogeneity of cancer types, which vary substantially in their physiological pathways, and the large idiosyncratic component in cancer onset that the latent health stocks do not capture.

Table 8: Results: Self-Reported Health

	Self-Reported Health
$\ln x^c$	0.2484
$\ln x^m$	0.1689
$\ln x^p$	0.4382
b_0	-0.2571
Age	0.0108
Intercept	2.6849
σ_v	0.4873

Last, in Table 8, we examine the relationship between latent health domains 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 domains significantly influence self-reported health, indicating that individuals with varying perceptions of latent health domains may still self-report a comparable overall rating of their health.

Taken together, these findings demonstrate that self-reported health is a weighted aggregate of distinct latent dimensions rather than a sufficient statistic for any single domain. While self-reported health is widely used in applied research and correlates strongly with mortality and future health outcomes (French, 2005; Pijoan-Mas and Ríos-Rull, 2014; Capatina, 2015; De Nardi et al., 2024), its aggregated nature obscures the domain-specific mechanisms through which health shapes economic behavior.

5 Counterfactual Exercise: Decomposition

This section uses the estimated model to decompose the determinants of health formation and related behavioral outcomes. The structural parameters are grouped into seven channels: initial heterogeneity, the three latent health domains (cognitive, mental, and physical health), idiosyncratic health shocks, health investments, and employment. Each channel is shut down separately while holding all other model features fixed. For example, when eliminating the cognitive-health channel, we set $x_t^c = 0$ for all t, and re-simulate the model without altering any other parameters.

Table 9: Percentage Change in Health Components Relative to Baseline

	x^c	x^m	x^p	d_t	L_t	I_t	Self-Reported Health	Self-Reported Health Variance
$t = 5$								
b_0	-4.22	14.96	0.07	0.28	-0.35	-0.75	0.35	7.14
x^c	-100	-28.54	-5.10	-3.03	0.89	18.39	3.33	-5.78
x^m	2.45	-100	29.93	0.83	-4.03	-16.56	-2.47	-10.01
x^p	-1.36	-47.11	-100	-22.59	4.22	29.68	4.68	-38.68
ϵ	-3.13	-88.80	20.70	-6.06	-4.28	-11.08	-1.74	-35.30
I	-4.22	-9.47	-0.79	-0.55	-0.11	0.65	0.10	-0.22
L	10.07	36.92	-7.34	-0.55	0.89	2.37	0.38	0.36
$t = 10$								
b_0	-6.26	11.58	-1.70	0.36	-1.75	8.04	0.34	1.65
x^c	-100	-39.41	-15.98	-3.76	3.34	28.32	5.78	-7.20
x^m	6.65	-100	38.35	4.38	-19.87	-28.97	-5.19	-12.67
x^p	-5.22	-35.52	-100	-20.48	-5.25	44.49	8.21	-42.09
ϵ	-1.69	-109.86	41.40	-1.34	-33.70	-30.37	-5.22	-48.84
I	-3.69	-3.30	-2.81	-0.45	1.43	1.21	0.41	0.94
L	3.40	14.79	-5.03	-0.27	2.54	4.49	0.47	1.54

Notes: Entries report percentage changes relative to the baseline simulation when the corresponding channel is removed.

Table 9 reports the resulting percentage changes relative to the baseline simulation using the estimated parameters. Outcomes are evaluated five years after model entry (ages 70–71) and ten years after entry (ages 75–76). Columns correspond to the different health domains and behavioral outcomes.

Cognitive health is relatively insensitive to most channels. Aside from its own persistence, the largest determinant is initial heterogeneity, which raises cognitive health by about 15 percent at $t = 5$. Eliminating the mental-health channel also slightly lowers cognitive outcomes (2–7 percent), suggesting that the gradual rise in mental health exert some upward pressure on cognition. Other factors have comparatively small effects.

Mental health, by contrast, is highly sensitive to interactions with other health domains. Removing the cognitive-health channel substantially improves mental health by about 30 percent at $t = 5$ and 38 percent at $t = 10$. Eliminating physical health produces even larger improvements, roughly 47 percent and 36 percent at t=5 and t=10, respectively. These results reflect the strong spillovers across health domains: declines in cognitive or physical functioning significantly worsen mental health in the baseline model. Idiosyncratic shocks, which have the largest variance for mental health, also play an important role, particularly at longer horizons.

Physical health is most strongly affected by mental health. Shutting down the mental-health channel reduces physical health by roughly 30–38 percent, indicating that improvements in mental health help sustain physical functioning over time. Cognitive health also contributes to physical outcomes, though the magnitude of this effect is smaller. The roles of investments and employment are comparatively modest.

The fourth column reports changes in mortality. Physical health is the dominant determinant of survival: removing the physical-health channel reduces survival by about 23 percent at t=5 and 20 percent at t=10. The fifth column shows that physical health is also an important determinant of employment. In addition, eliminating the mental-health channel substantially lowers employment, with employment falling by roughly 4 percent at t=5 and nearly 20 percent at t=10. This pattern suggests that focusing solely on physical health, as is common in much of the literature, may miss an important channel through which health affects economic behaviour, highlighting the value of modelling mental and physical health jointly.

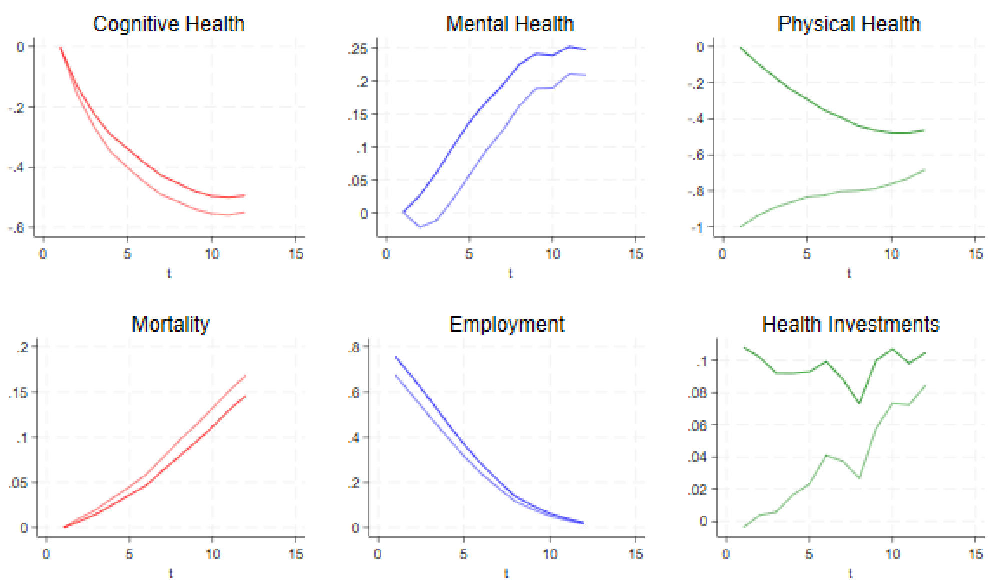
The final two columns report the effects on self-reported health and its variance, the latter often interpreted as a measure of health inequality. Removing the physical-health channel raises average self-reported health by roughly 5 percent at t=5 and 8 percent at t=10, reflecting the removal of a domain that declines sharply with age. However, it also produces the largest reduction in health inequality, as the variance of self-reported health falls by about 39 percent after five years and 45 percent after ten years. This pattern reflects the fact that physical health both deteriorates substantially with age and exhibits large cross-individual variation, making it a primary source of dispersion in overall health.

Eliminating the mental-health channel also reduces inequality, though to a lesser extent. Taken together, these results suggest that heterogeneity in both physical and mental health is a major driver of dispersion in overall health status among older adults. At the same time, the reliance on self-reported health as a summary measure obscures the distinct sources of this variation, making it difficult to distinguish whether inequality arises from differences in physical or mental health and, consequently, to identify the mechanisms through which health disparities translate into economic and welfare consequences..

6 Experiments Using the Model: Shocks to Health Domains

In this final section, we use the estimated model to simulate how a shock to each latent health domain propagates through the system and shapes health formation, employment and investment decisions, and mortality. Figures 7–9 report the simulated effects of a one standard deviation shock to physical, mental, and cognitive health, respectively. Each figure shows the dynamic responses of the three health domains along with mortality, employment, and health investment, allowing us to visualize how shocks diffuse across dimensions and over time.

Figure 7: Effects of a -1 SD Shock to Physical Health at $t = 1$

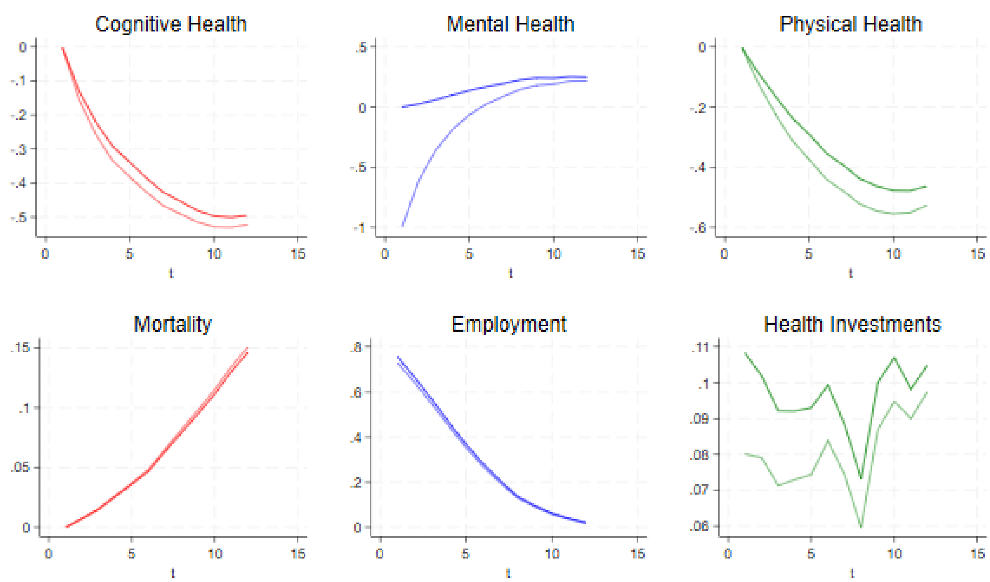


Note: Solid line is baseline, dashed line is counterfactual scenario

A negative shock to physical health generates the strongest cross-domain spillovers. Mental health declines immediately and exhibits persistent losses, failing to return to its pre-shock path. Cognitive health also deteriorates, though more gradually. While physical health partially recovers, it remains persistently below baseline. These health losses translate into economically meaningful consequences: employment declines steadily, mortality rises, and health investments adjust in response to the worsened health trajectory. These results underscore the central role of physical health in the system, as shocks originating in this domain propagate broadly across other health dimensions and translate into sustained economic consequences.

A shock to mental health spreads both downward into physical health and cognitive health. Physical health declines progressively following the mental health shock, highlighting the importance of psychological well-being for somatic outcomes. Cognitive health is also adversely affected. Unlike physical shocks, however, mental health itself largely converges back toward baseline over time. Even so, the indirect effects on physical health persist, generating sustained impacts on employment and mortality. Thus, even transitory disruptions to mental health can produce long-run economic effects through their downstream impact on physical health.

Figure 8: Effects of a -1 SD Shock to Mental Health at $t = 1$

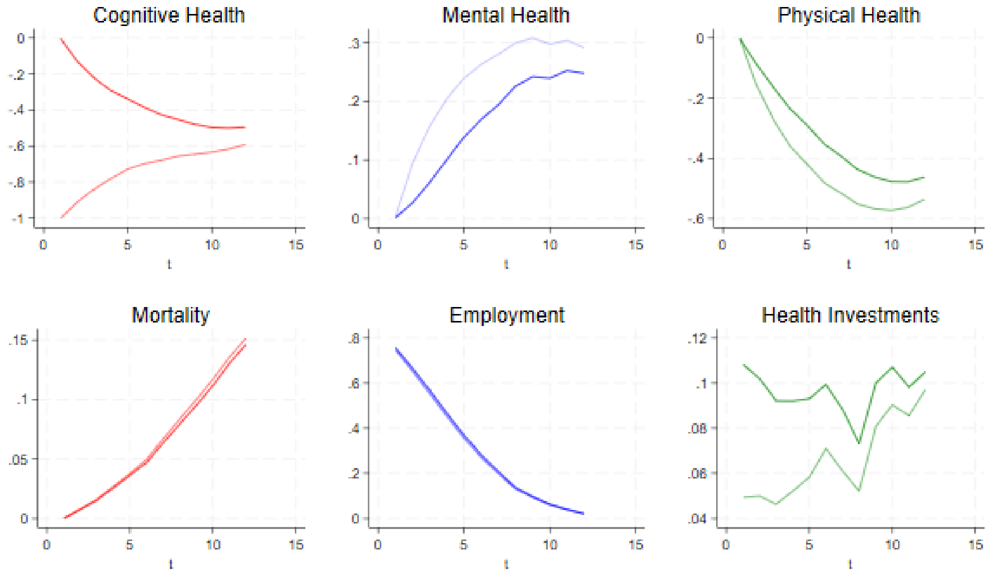


Note: Solid line is baseline, dashed line is counterfactual scenario

A shock to cognitive health primarily affects cognitive trajectories and, to a lesser extent, mental health, but it has little to no effect on physical health. Cognitive health only partially recovers, leaving a persistent deficit. The limited spillover to physical health implies more contained downstream effects on mortality and employment relative to a physical health shock. Nonetheless, the persistent cognitive losses still reduce employment and modestly increase mortality.

Taken together, these counterfactuals reveal a clear asymmetry in the health production process. Physical health appears to play a central, system-wide role: shocks in this domain propagate broadly and durably

Figure 9: Effects of a -1 SD Shock to Cognitive Health at $t = 1$



Note: Solid line is baseline, dashed line is counterfactual scenario

across other health dimensions and economic outcomes. Mental health occupies an intermediate position, with meaningful spillovers, particularly into physical health, while cognitive health shocks are comparatively more localized. These dynamics underscore the importance of accounting for cross-domain complementarities when evaluating health policies or insurance design.

7 Conclusion

This research investigates the formation and dynamics of health in older adults, focusing on the interplay among physical, cognitive, and mental health. We develop and estimate a model using data from the Health and Retirement Study to examine how these health dimensions evolve over time and influence economic behavior, including employment and health investment decisions. Our findings suggest important complementarities across health domains. Notably, physical functioning exhibits strong persistence and has a substantial influence on mental health. The analysis also examines how the latent health domains relate to commonly used health measures and how well they predict outcomes such as employment.

Overall, the results indicate that physical health is the most important determinant of employment, health investments, and mortality. However, the joint analysis reveals a nontrivial role for mental health, and to a lesser extent cognitive health. In particular, the estimated health production function highlights strong linkages between physical and mental health. Accounting for multiple health domains and their interdependence is therefore crucial for understanding the mechanisms through which health shapes economic behavior and outcomes at older ages. Each latent health domain also contributes to individuals' self-assessment of their health, providing insight into what this commonly used measure of self-reported health captures in empirical studies.

At its current stage, this paper lays the groundwork for a deeper understanding of health formation and its role in shaping health-related economic inequality at older ages. Ongoing development of this paper and the broader research agenda will proceed along three main directions. First, we will relax the independence assumptions on unobservables and assess the robustness of the results to alternative specifications of the health formation process. Second, we plan to incorporate additional features related to the demand for health, such as medical expenditures, to better understand the drivers of health formation. Finally, we aim to embed the estimated health formation technology within the theoretical model presented in Section 2. Doing so will structurally link health dynamics with economic behavior and enable counterfactual policy analysis, 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.

Allman, E. S., Matias, C., and Rhodes, J. A. (2009). Identifiability of parameters in latent structure models with many observed variables.

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., 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.

Gourieroux, C., Monfort, A., and Renault, E. (1993). Indirect inference. *Journal of applied econometrics*, 8(S1):S85–S118.

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.

A 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

$$U(C_{it}, I_{it}; Y_{it}, H_{it}, \xi_i, \zeta_{it}) = \frac{(C_{it} + \alpha I_{it})^{1-\gamma}}{1-\gamma} - v(Y_{it}, H_{it}, \xi_i, \zeta_{it}), \tag{12}$$

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:

$$A_{it+1} = (1 + r_t)(A_{it} + Y_{it}W_{it}(1 - \tau_t(Y_{it}W_{it}, H_{it})) + b_t(H_{it}) - C_{it} - p_t^I I_{it}) \tag{13}$$

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

$$W_{it} = w(t, H_{it}, \phi_i, v_{it}). \tag{14}$$

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,

$$H_{it} = h(t, H_{it-1}, Y_{it-1}, I_{it}, \psi_i, \epsilon_{it}). \tag{15}$$

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

$$S(t, H_t). \tag{16}$$

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).

A.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.

A.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,

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

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

$$\frac{\partial}{\partial C}: U_C - \beta(1+r_t)S(t,H_t)\mathbb{E}V_{A,t+1} = 0 \tag{17}$$

$$\frac{\partial}{\partial I}: U_I - \beta(1+r_t)S(t,H_t)\mathbb{E}(p_t^I V_{A,t+1} + V_{H,t+1}H_I) = 0, \tag{18}$$

where $V_{x,t+1}$ denotes the partial derivative of V_{t+1} with respect to x. From equation (17), 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

$$U_I = p_t^I U_c + \beta S(t,H_t)\mathbb{E}V_{H,t+1}H_I. \tag{19}$$

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

$$Y_{it} = \mathbb{1}\Bigg[\max_{C_t,I_t}\Big\{U(C_{it},I_{it};Y_{it}=1,H_{it},\xi_i,\zeta_{it}) + \beta S(t,H_{it})\mathbb{E}V_{t+1}(\Omega_{it+1}^1)\Big\} \tag{20}$$

$$- \max_{C_t,I_t}\Big\{U(C_{it},I_{it};Y_{it}=0,H_{it},\xi_i,\zeta_{it}) + \beta S(t,H_{it})\mathbb{E}V_{t+1}(\Omega_{it+1}^0)\Big\} \geq 0\Bigg]. \tag{21}$$

Denote (C_t^y, I_t^y) as solutions to equations (17) and (18). 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

$$Y_{it}^* = \mathbb{1}(Y_t^l > 0). \tag{22}$$

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

$$I^*_{it} = I(H_{it}, W_{it}, A_{it}, t, \xi_i, \phi_i, \psi_i, \zeta_{it}|Y_{it}, \theta) \tag{23}$$

$$Y^*_{it} = Y(H_{it}, W_{it}, A_{it}, t, \xi_i, \phi_i, \psi_i, \zeta_{it}|\theta), \tag{24}$$

where θ is the set of all parameters in equations (12) - (16). This expression is useful to illustrate the determinants of health formation. Substituting equations (23) and (24) into equation (25) yields

$$H_{it+1} = H(t, H_{it}, I^*_{it}, Y^*_{it}, \psi_i, \epsilon_{it+1}), \tag{25}$$

which clearly illustrates the endogeneity present when estimating (25) 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 domains, for instance using a CES technology as in Cunha et al. (2010). The health domains 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.

B Further Details on Identification

B.1 Identification of Initial Distributional Parameters and the Measurement System

This section establishes identification of the measurement system and the joint distribution of initial latent endowments. Throughout, m indexes observed measures, k indexes latent factors, j indexes mixture components, and t indexes time periods. Because all measurement parameters are assumed to be age-invariant, observations in the initial period ($t = 0$) are sufficient to identify both the measurement parameters and the distribution of initial latent states. We therefore suppress the time index for the remainder of this section.

We proceed in three steps. First, using binary proxies for each latent factor, we identify the marginal mixture distributions and the binary thresholds via marginal and pairwise probabilities. Second, using identified mixture parameters and continuous proxies, we recover factor loadings and measurement error variances. Third, we recover the off-diagonal elements of the component covariance matrices from cross-domain proxy co-movement.

Let $w_0 = (\ln x_0^p, \ln x_0^c, \ln x_0^m, b_0)'$ denote the vector of initial latent factors. The joint distribution of w_0 is approximated by a two-component Gaussian mixture:

$$F_{w_0} = \pi\Phi(\omega^1, \Sigma^1) + (1 - \pi)\Phi(\omega^2, \Sigma^2),$$

where ω^j and Σ^j are the mean vector and covariance matrix of component j. Location and scale are normalized by $E(w_0) = 0$ and $\text{diag}(\text{Var}(w_0)) = 1$. These restrictions anchor the latent scale and eliminate observational equivalence between measurement loadings and the variance of latent factors. The off-diagonal elements of Σ^j are unrestricted and capture dependence across initial endowments.

B.1.1 Identification Using Binary Measures

Each latent factor is observed through multiple proxies. Under the conditional independence restrictions on measurement errors, all co-movement across proxies within a domain is attributable to the shared latent factor. To start, we establish identification marginally within each domain.

Fix a latent factor x^k, which follows a two-component Gaussian mixture:

$$\ln x^k \sim \pi N(\omega_k^1, s_k^1) + (1 - \pi)N(\omega_k^2, s_k^2).$$

The normalizations $E(\ln x^k) = 0$ and $\text{Var}(\ln x^k) = 1$ imply that ω_k^2 and s_k^2 are functions of (ω_k^1, s_k^1, π), leaving three free mixture parameters per factor.

Suppose x^k is measured by $M \geq 3$ binary proxies,

$$z_m^k = 1[\ln x^k + v_m^k \geq \tau_m^k],$$

where $v_m^k \sim N(0,1)$. The observed measures provide two sets of identifying moments.

- **(M1) Marginal probabilities.** For each m, $\Pr(z_m^k = 1)$ is a function of τ_m^k, ω_k^1, s_k^1, and π. The M marginal probabilities provide M equations in $M+3$ unknowns (the M thresholds and three mixture parameters).

- **(M2) Pairwise probabilities.** For $j \neq l$, the joint probability $\Pr(z_j^k = 1, z_l^k = 1)$ is given by a mixture of bivariate normal CDFs, where the within-component correlation between the two binary indicators equals
$$\frac{s_k^2}{s_k^2 + 1}.$$

This follows because the conditional correlation of z_j^k and z_l^k given component membership is entirely determined by the factor variance s_k^2 relative to the unit measurement error variance. The $M(M-1)/2$ pairwise probabilities provide additional identifying equations.

Together, (M1) and (M2) yield $M + M(M-1)/2$ moment conditions for the M thresholds and three mixture parameters. For $M \geq 3$, this system is generically overidentified, and the mixture distribution and thresholds are identified up to label switching. A standard ordering restriction on component means (e.g., $\omega_k^1 < \omega_k^2$) resolves the labeling indeterminacy. The generic identification of finite Gaussian mixtures from such moment systems is established in the literature on nonparametric identification of mixture models; see, e.g., Allman et al. (2009).

B.1.2 Identification with Mixed Binary and Continuous Measures

Consider a latent factor measured by two continuous measures and $M \geq 3$ binary indicators. The three mixture parameters, M binary thresholds, and measurement parameters for the continuous measures must be identified. As established in Section 2.1, the $M \geq 3$ binary measures are sufficient to identify the mixture distribution and all thresholds. We now recover the continuous measurement parameters $(\mu_m^k, \lambda_m^k, \sigma_{v,km}^2)$.

The means of the continuous measures identify the intercepts:

$$E(z_m^k) = \mu_m^k + \lambda_m^k E(\ln x_0^k) = \mu_m^k,$$

since $E(x_0^k) = 0$ by normalization.

The covariance between the two continuous measures identifies the product of loadings:

$$\text{Cov}(z_1^k, z_2^k) = \lambda_1^k \lambda_2^k \text{Var}(\ln x^k) = \lambda_1^k \lambda_2^k,$$

where the last equality uses $\text{Var}(\ln x^k) = 1$.

The joint distribution of (z_1^k, z_2^k) conditional on component j is bivariate normal with component-specific means and variances that are functions of $(\mu_m^k, \lambda_m^k, \omega_k^j, s_k^j)$. Provided that the two mixture components are separated in the latent factor ($\omega_k^1 \neq \omega_k^2$, which holds under a non-degeneracy condition that the two components have distinct means), the loadings are identified by:

$$\lambda_1^k = \frac{\omega_{z1}^1 - \omega_{z1}^2}{\omega_k^1 - \omega_k^2}, \qquad \lambda_2^k = \frac{\omega_{z2}^1 - \omega_{z2}^2}{\omega_k^1 - \omega_k^2},$$

where ω_{zm}^j denotes the component-j mean of continuous measure m, which is identified from the mixture. Finally, the measurement error variance is recovered from

$$\text{Var}(z_m^k) = (\lambda_m^k)^2 + \sigma_{v,km}^2.$$

B.1.3 Identification of Cross-Domain Covariances

It remains to identify the off-diagonal elements of Σ^j, which capture the dependence among initial health stocks and latent heterogeneity. Consider two distinct latent factors, x^k and x^l ($k \neq l$), each observed by at least one binary or continuous proxy. Under the conditional independence assumption, measurement errors are independent across domains given w_0. Hence all cross-domain co-movement in the data is attributable to the correlation of (x^k, x^l) within each mixture component.

Formally, for a binary proxy of x^k and a binary proxy of x^l, the cross-domain pairwise probability satisfies a bivariate normal mixture in which the within-component correlation is a function of the off-diagonal element ρ_{kl}^j. Given identified marginal mixture distributions and thresholds (from Steps 1 and 2), the cross-domain pairwise probabilities provide moment equations in ρ_{kl}^j. With sufficient cross-domain proxies, these equations identify all off-diagonal elements of Σ^j.

B.2 Identification of the Transition Process and Remaining Model Parameters

Taking as given the measurement parameters and joint distribution of initial latent endowments identified in Step 1, we establish identification of the health formation technology and the remaining behavioral equations. In line with the estimation strategy, the identification strategy uses an auxiliary model based on principal components of the observed binary measures. Because the latent states are not directly observed, we work with a mapping from structural parameters to the moments of an observable auxiliary statistic, following the logic of indirect inference.

B.2.1 The PCA Index as an Auxiliary Statistic

Let $x_t = (x_t^p, x_t^c, x_t^m)'$ denote the latent health domains and let b_0 denote time-invariant heterogeneity. For each health domain $k \in \{p, c, m\}$, we observe a vector of measures $z_t^k = (z_{t,1}^k, \dots, z_{t,M_k}^k)'$ that are exclusively generated by x_t^k through the measurement system.

Fix a latent domain x_t^k and suppose it is measured by M binary proxies

$$z_{t,m}^k = \mathbb{1}\left[\ln x_t^k + v_{t,m}^k \geq \tau_m^k\right], \qquad v_{t,m}^k \sim N(0,1).$$

Conditional on x^k, each observed indicator satisfies

$$P(z_{t,m}^k = 1 \mid x_t^k) = \Phi(\ln x_t^k - \tau_m^k),$$

where Φ is the Standard Normal CDF. This conditional probability is a strictly increasing function of $\ln x_t^k$. Under the conditional independence of measurement errors, all co-movement across binary indicators is driven by x_t^k.

Let

$$PC_t^k = \sum_m g_m^k z_{t,m}^k$$

denote the first principal component of the observed binary vector. Because PCA is applied to nonlinear threshold variables rather than to linear measurements of $\ln x_t^k$, the first principal component does not recover the latent factor exactly. Instead, its conditional expectation satisfies

$$E(PC_t^k \mid x_t^k) = \sum_{m=1}^{M} g_m^k \Phi(\ln x_t^k - \tau_m^k) \equiv h(x^k),$$

where $h(\cdot)$ is a known, strictly increasing function once the PCA weights $\{g_m^k\}$ and the identified thresholds $\{\tau_m^k\}$ are fixed. Thus PC_t^k is a monotone but nonlinear proxy for $\ln x_t^k$. We use PC_t^k as an auxiliary statistic for the latent state, not as a direct estimate of it.

When measures are continuous, the mapping from the principal component to the respective latent factor is clearly one-to-one and invective,

$$PC_t^k = \sum_m g_m^k z_{t,m}^k = \sum_m g_m^k(\mu_m^k + \lambda_m^k \ln x_t^k + v_{t,m}^k).$$

We focus this section on a case more similar to our setting, in which we have discrete measures exclusively for many latent health domains. For identification using continuous linear measures, see Cunha et al. (2010); Agostinelli and Wiswall (2025).

B.2.2 Identification of Transition Parameters via Indirect Inference

The structural parameters are identified through the unique reduced-form implications they generate for the PCA-based auxiliary moments. Cross-domain interaction parameters are identified provided the auxiliary regression is specified jointly, which includes all PC indices and their pairwise interactions as regressors, and the known measurement rescalings are accounted for when mapping auxiliary coefficients back to structural parameters.

For each domain k, define the auxiliary transition equation

$$PC_{t+1}^k = \tilde{a}_0^k + \tilde{w}_t^\top \tilde{\beta}^k + \tfrac{1}{2}\tilde{w}_t^\top \tilde{B}^k \tilde{w}_t + \tilde{c}_e^k e_t + \tilde{\nu}_{t+1}^k,$$

where

$$\tilde{w}_t = (PC_t^p,\, PC_t^c,\, PC_t^m,\, PC_0^b,\, I_t)'$$

collects the principal component indices and observed state variables (I_t), e_t denotes the observed employment status, and $\tilde{\nu}_{t+1}^k$ is a mean-zero innovation independent of $\tilde{w}_t$.

Estimating this auxiliary model by OLS yields a reduced-form parameter vector $\beta^A(\theta)$, which is a function of the structural parameters θ. The structural parameters are identified if the auxiliary mapping $\theta \to \beta^A(\theta)$ is injective over the admissible parameter space:

$$\beta^A(\theta_1) = \beta^A(\theta_2) \Rightarrow \theta_1 = \theta_2.$$

The economic content of this condition is as follows. The structural parameter a_0^k shifts the unconditional location of the latent process, altering the average probability of positive responses across indicators and hence the mean of PC_t^k. The own-lag coefficient in β^k changes the persistence of x_t^k, which maps directly into the autocorrelation of the binary indicators and the serial correlation of PC_t^k. The innovation variance $\sigma_{\nu^k}^2$

affects the cross-sectional dispersion and intertemporal movement of the latent process, which are reflected in the variance and transition moments of PC_t^k. Because these three features, location, persistence, and dispersion, affect distinct moments of the observable auxiliary statistic, perturbations in the structural parameters generate distinguishable changes in $\beta^A(\theta)$.

B.2.3 Identification of Behavioral Equations

Conditional on the identified latent dynamics, the remaining model parameters governing health investment, employment, mortality, and attrition are identified from the conditional distributions of observed outcomes given the current state vector. Because each outcome depends on the current latent factors only through

$$w_t = (\ln x_t^p,\, \ln x_t^c,\, \ln x_t^m,\, b_0,\, I_t)',$$

and because the PCA index $h(x_t^k)$ is a monotone transformation of each latent factor with known parameters, the coefficients in the reduced-form regressions of observed outcomes on

$$(PC_t^p,\, PC_t^c,\, PC_t^m,\, PC_0^b)$$

bear a one-to-one relationship to the underlying structural parameters, up to the rescaling implied by the measurement system.

C Estimation Details

C.1 Step 1: Distribution of Initial Endowments and Measurement Parameters

Recall the distribution of initial conditions is

$$F_{w_0} = \pi\Phi(\omega^1, \Sigma^1) + (1-\pi)\Phi(\omega^2, \Sigma^2), \tag{26}$$

$$\text{where} \quad \omega^j = \begin{bmatrix} \omega_1^j \\ \omega_2^j \\ \omega_3^j \\ \omega_4^j \end{bmatrix} \quad \text{and} \quad \Sigma^j = \begin{bmatrix} s_1^j & \rho_{12}^j & \rho_{13}^j & \rho_{1b}^j \\ \rho_{12}^j & s_2^j & \rho_{23}^j & \rho_{2b}^j \\ \rho_{13}^j & \rho_{23}^j & s_3^j & \rho_{3b}^j \\ \rho_{1b}^j & \rho_{2b}^j & \rho_{3b}^j & s_4^j \end{bmatrix}.$$

Let

$$\rho^j = (\rho_{12}^j, \rho_{13}^j, \rho_{1b}^j, \rho_{23}^j, \rho_{2b}^j, \rho_{3b}^j) \quad \text{and} \quad s^j = (s_1^j, s_2^j, s_3^j, s_4^j), \quad \text{and} \quad \omega^j = (\omega_1^j, \omega_2^j, \omega_3^j, \omega_4^j).$$

We denote the vector of parameters to estimate in the initial step by

$$\theta_0 = \{\{\omega^j, \rho^j, \sigma^j\}_{j=1,2}, \{\mu_m^k, \lambda_m^k, \tau_{j,m}^k, \sigma_{v_m^k}\}_{k\in\{p,c,mh,b\}, m\in 1,\dots,M_k}\}.$$

We denote the set of observed measures in period 0 as

$$z_0 = (z_{0,1}^p, \dots, z_{0,M^p}^p, z_{0,1}^m, \dots, z_{0,M^m}^m, z_{0,1}^c, \dots, z_{0,M^c}^c, z_1^b, \dots, z_{M^b}^b).$$

The likelihood function is constructed from the joint distribution of initial heterogeneity and the latent health domains, w_0. As w_0 is latent, we write the joint density of observed measures at $t=0$ conditional on w_0 as

$$f(z_0|x_0^k, \theta_0) = \prod_{j=1}^{M_b} f_{z_{0,j}^k}(z_{0,j}^b|b_0, \theta_0) \prod_{k\in\{p,m,c\}} \prod_{j=1}^{M_k} f_{z_{0,j}^k}(z_{0,j}^k|x_0^k, \theta_0), \tag{27}$$

where $f_{z_{0,j}^k}(z_{0,j}^b|b_0, \theta_0)$ is the density of initial latent heterogeneity measure $z_{0,j}^b$ for $j = 1, \dots, M_0^b$. The joint density of initial endowments, $w_0 = (x_0^p, x_0^m, x_0^c, b_0)$ given θ_0 is $f_{w_0}(w_0)$.

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

$$L_i(\theta_0) = \int \cdots \int f_{w_0}(w_0) f(z_0 | x_0^k, \theta) db_0 dx_0 \tag{28}$$

Initial heterogeneity and the health domains 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 $x_0 = (x_0^p, x_0^c, x_0^{mh})$.

We use simulated Maximum-Likelihood Estimation (SML) to maximize the likelihood function and solve for $\hat{\theta}_0$. In our model the observed-data likelihood involves integrating over the entire sequence of latent health stocks $\{x_0^1, x_0^2, x_0^3, \}$, and initial heterogeneity, b_0. We approximate this integral by employing Monte Carlo draws from the joint distribution of latent states as dictated by the model. The following describes the steps to simulate a time series and calculate the likelihood for a single individual.

1. *Simulate Latent variables:* For a given guess of parameters $\tilde{\theta}_0$, we simulate R independent draws of $\tilde{b}_0$ and latent states $\{\tilde{x}_0^p, \tilde{x}_0^c, \tilde{x}_0^{mh}, \}$ for r = 1, ..., R. For each replication r= 1, ..., R, we draw $w_0^{(r)} = (\tilde{x}_0^p, \tilde{x}_0^c, \tilde{x}_0^{mh}, \tilde{b}_0)' \sim F_{w_0}(\tilde{\theta}_0)$, where $w_0 = (x_0^p, x_0^c, x_0^{mh}, b_0)'$ compute the conditional probability of the realized data given the simulated state, $w_0^{(r)}$.

2. *Evaluate the Likelihood Contribution of a Path:* For each replication r, compute the likelihood of observing the measures by (28), which we denote as $L_i^{(r)}(\tilde{\theta}_0)$.

3. *Approximate the Likelihood:* The likelihood contribution of individual i is then approximated by averaging over the R simulated paths,

$$\hat{L}_i(\tilde{\theta}_0) \approx \frac{1}{R} \sum_{r=1}^{R} L_i^{(r)}(\tilde{\theta}_0). \tag{29}$$

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.

C.2 Step 2: Transition, mortality, employment, investments

Fixing the parameters of the measurement system and joint distribution of initial endowments, we use minimum distance moment matching to pin down the remaining model parameters. To construct the empirical moments used in estimation, we first extract principal components corresponding to each latent health domain using the observed health measures in the data. Specifically, we perform principal component analysis (PCA) separately on groups of measures that load exclusively on a single latent health domain. These correspond to cognitive health, mental health, physical health, and a time-invariant heterogeneity component. Categorical measures are first converted into binary indicators equal to one when the corresponding category is selected. Applying PCA to these binary measures over all time periods yields a set of principal components for each individual i and period t, denoted

$$PC_{it} = \{PC^c_{it}, PC^m_{it}, PC^p_{it}, PC^b_i\}$$

where PC^b_i captures persistent heterogeneity and therefore does not vary over time.

These principal components serve as empirical proxies for the latent health factors and allow us to construct moments that correspond to the model's transition equations. In particular, define $\tilde{w}_t = \{PC^c_{it}, PC^m_{it}, PC^p_{it}, PC^b_i, I_t\}$, then we estimate reduced-form analogues of the health production functions of the form

$$PC^k_{t+1} = a^k_0 + \tilde{w}_t'\beta^k + \tfrac{1}{2}\tilde{w}_t' B^k \tilde{w}_t + c^k_e L_t + \nu^k_t. \tag{30}$$

The coefficients from these regressions, together with the variance of the residuals, form the key moments used to identify the parameters governing the latent health transition process.

A key feature of the procedure is that the mapping from observed measures to principal components is fixed using the data. PCA is performed once using the observed outcomes, producing a vector of loadings, centering constants, and scaling constants. These quantities define a linear transformation from the observed measures z^{k*}_{it} to the principal components, where z^{k*}_{it} refers to the set of measures such that categorical measures are discretized. When simulating the model, the same transformation is applied to the simulated measures so that the simulated indices are constructed using the identical mapping:

$$PC^k_{it} = \sum_{m=1}^{M^{k*}} v_m \frac{z^{k*}_{im} - \mu_m}{\sigma_m},$$

where M^{k*} is the number of measures for health domain k, after converting categorical to binary. That is, I fix the mapping from the data, and simple calculate the principal components from the simulated model by

$$PC^{k,sim}_{it} = \sum_{m=1}^{M^{k*}} v_m \frac{\bar{z}^{k*,sim}_i - \mu_m}{\sigma_m}$$

where $\bar{y}_{ik}^{sim}$ is the predicted probability that the simulated binary measure $y_{ik}^{sim} = 1$

In addition to the health transition equations, we construct moments from behavioral relationships linking health to investment and labor market outcomes. Using the principal component indices as proxies for the latent health domains, we estimate the investment equation

$$I_t = \beta_0^I + \beta_1^I PC_t^p + \beta_2^I PC_t^c + \beta_3^I PC_t^{mh} + \beta_4^I PC^b + \beta_5^I t + \nu_t^I, \tag{31}$$
$$\nu_t^I \sim iid\, \mathcal{N}(0, \sigma_I^2),$$

where $\nu_t^I \sim N(0, \sigma_{\nu^I})$. The coefficients from this regression, along with the variance of the residual, are included among the moments matched in estimation. We also estimate linear probability models describing employment, mortality, and attrition.

Finally, we match unconditional moments of key outcomes across time. These include the period-specific means of each health measure, as well as the average levels of employment, health investment, mortality, and attrition in each period. Matching these moments ensures that the model reproduces both the dynamic relationships between health domains and the empirical distributions of the key outcomes observed in the data.

C.3 Smoothed Simulation and Construction of Principal Components

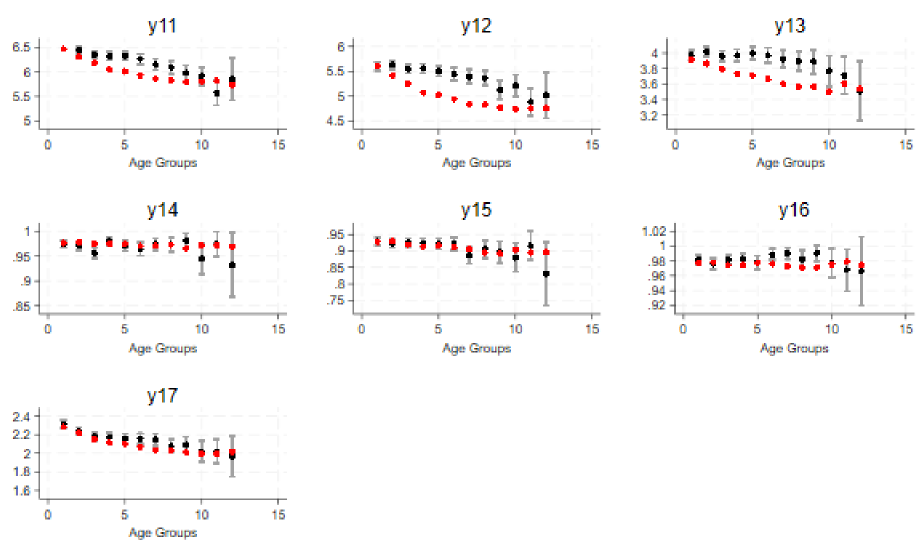
A practical complication arises because the health measures used to construct the principal components are discrete, while the simulated model produces latent states and discrete realizations only after applying threshold rules. If the simulated principal components were constructed directly from simulated binary outcomes, the resulting objective function would generally be non-smooth in the structural parameters. Small changes in the parameter vector may change whether a simulated latent index crosses a threshold, producing discontinuous changes in the simulated moments. To avoid this problem, I replace simulated discrete outcomes with their simulated conditional probabilities and construct the simulated principal components from these smoothed objects.

C.4 Fit of Moments

Table 10: Model Fit: Health Formation Process - Cognitive

	Sim	Data	SE	95 CI
$\ln x_t^c$	0.568	0.565	0.008	[(0.548, 0.581)]
$\ln x_t^m$	0.066	0.052	0.012	[(0.027, 0.076)]
$\ln x_t^p$	0.066	0.067	0.010	[(0.048, 0.086)]
b_0	0.084	0.097	0.008	[(0.081, 0.114)]
$(\ln x_t^c)^2$	-0.015	-0.017	0.004	[(-0.025, -0.01)]
$(\ln x_t^m)^2$	0.010	0.004	0.003	[(-0.002, 0.01)]
$(\ln x_t^p)^2$	0.006	0.002	0.003	[(-0.004, 0.007)]
$(b_0)^2$	-0.009	0.003	0.004	[(-0.006, 0.012)]
L_{t-1}	0.149	0.131	0.023	[(0.087, 0.175)]
$\ln x_t^c \times \ln x_t^m$	-0.007	-0.004	0.004	[(-0.012, 0.004)]
$\ln x_t^c \times \ln x_t^p$	-0.003	-0.002	0.004	[(-0.01, 0.006)]
$\ln x_t^m \times \ln x_t^p$	0.004	0.002	0.003	[(-0.005, 0.008)]
$b_0 \times \ln x_t^c$	-0.006	0.001	0.005	[(-0.01, 0.011)]
$b_0 \times \ln x_t^m$	-0.003	-0.001	0.004	[(-0.009, 0.006)]
$b_0 \times \ln x_t^p$	0.001	0.003	0.004	[(-0.005, 0.011)]
I_t	0.056	0.059	0.013	[(0.034, 0.084)]
$I_t \times \ln x_t^c$	-0.013	-0.008	0.009	[(-0.026, 0.01)]
$I_t \times \ln x_t^m$	0.001	-0.001	0.007	[(-0.014, 0.013)]
$I_t \times \ln x_t^p$	0.005	0.010	0.008	[(-0.005, 0.025)]
$I_t \times b_0$	-0.018	-0.008	0.009	[(-0.025, 0.008)]
Intercept	-0.212	-0.107	0.024	[(-0.154, -0.059)]
σ_v^2	1.268	1.163	0.008	[(1.148, 1.178)]

Figure 10: Fit of Measures of Cognitive Health

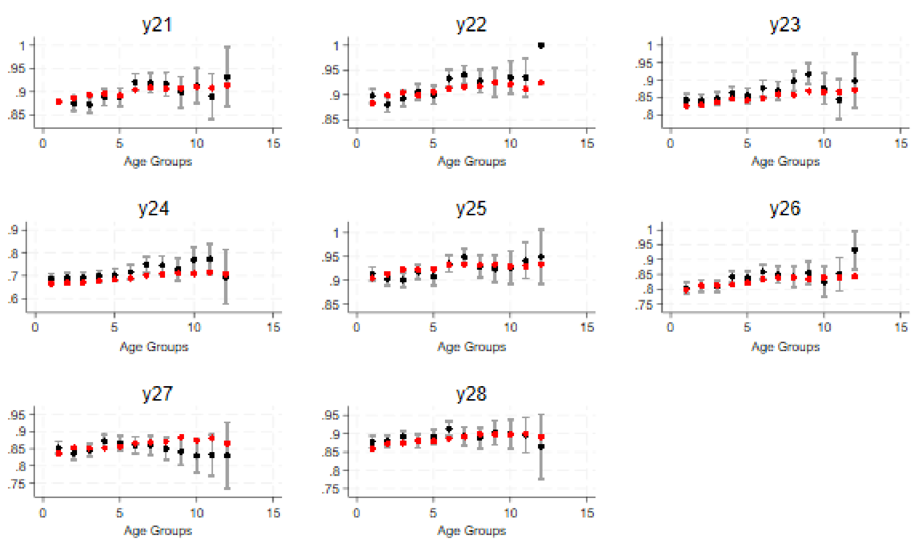


Note: ages are grouped into two-year bins. That is, age group 1 is 55-56 years old, age group 2 is 57-58 years old, etc. Red lines correspond to predictions from the simulated model and blue lines correspond to the observed data.

Table 11: Model Fit: Health Formation Process - Mental Health

	Sim	Data	SE	95 CI
$\ln x_t^c$	-0.010	0.079	0.011	[0.057, 0.102]
$\ln x_t^m$	0.512	0.483	0.017	[0.45, 0.516]
$\ln x_t^p$	0.107	0.156	0.013	[0.13, 0.182]
b_0	0.047	0.060	0.012	[0.038, 0.083]
$(\ln x_t^c)^2$	0.001	0.001	0.005	[-0.009, 0.012]
$(\ln x_t^m)^2$	0.025	0.016	0.004	[0.008, 0.024]
$(\ln x_t^p)^2$	0.020	0.014	0.004	[0.007, 0.022]
$(b_0)^2$	-0.002	-0.001	0.006	[-0.013, 0.011]
L_{t-1}	0.081	-0.053	0.031	[-0.113, 0.008]
$\ln x_t^c \times \ln x_t^m$	-0.008	-0.028	0.006	[-0.039, -0.017]
$\ln x_t^c \times \ln x_t^p$	0.002	-0.003	0.006	[-0.014, 0.008]
$\ln x_t^m \times \ln x_t^p$	-0.006	-0.007	0.004	[-0.016, 0.001]
$b_0 \times \ln x_t^c$	-0.001	0.006	0.007	[-0.008, 0.02]
$b_0 \times \ln x_t^m$	-0.017	-0.014	0.005	[-0.024, -0.003]
$b_0 \times \ln x_t^p$	-0.010	-0.011	0.006	[-0.022, 0]
I_t	0.079	0.058	0.018	[0.024, 0.093]
$I_t \times \ln x_t^c$	0.018	0.028	0.013	[0.003, 0.053]
$I_t \times \ln x_t^m$	-0.030	-0.019	0.010	[-0.037, 0]
$I_t \times \ln x_t^p$	-0.005	-0.013	0.011	[-0.034, 0.008]
$I_t \times b_0$	0.001	-0.040	0.012	[-0.063, -0.017]
Intercept	-0.055	-0.025	0.033	[-0.09, 0.04]
σ_v^2	1.268	1.589	0.011	[1.568, 1.609]

Figure 11: Fit of Measures of Mental Health

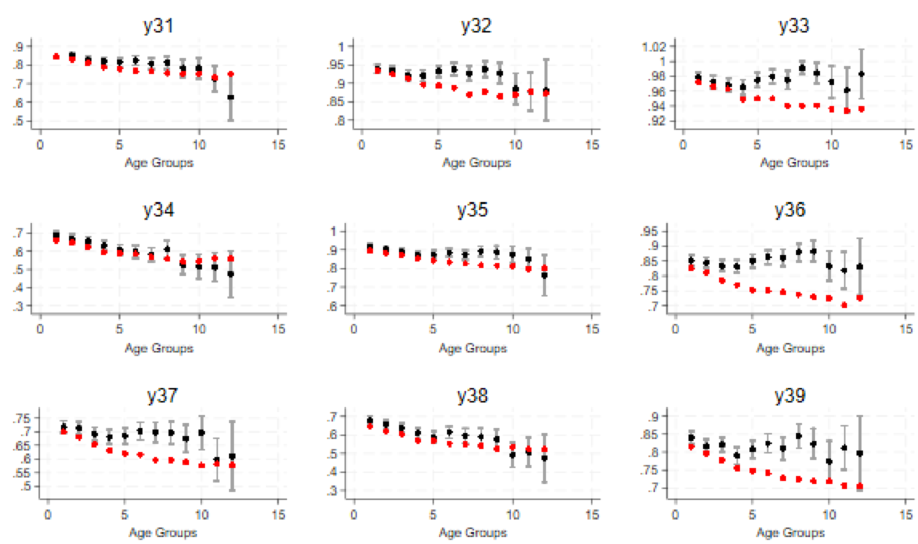


Note: ages are grouped into two-year bins. That is, age group 1 is 55-56 years old, age group 2 is 57-58 years old, etc. Red lines correspond to predictions from the simulated model and blue lines correspond to the observed data.

Table 12: Model Fit: Health Formation Process - Physical

	Sim	Data	SE	95 CI
$\ln x_t^c$	0.048	0.056	0.009	[0.039, 0.073]
$\ln x_t^m$	0.067	0.092	0.013	[0.066, 0.118]
$\ln x_t^p$	0.736	0.780	0.010	[0.76, 0.8]
b_0	0.018	0.032	0.009	[0.015, 0.05]
$(\ln x_t^c)^2$	0.000	-0.001	0.004	[-0.009, 0.007]
$(\ln x_t^m)^2$	0.010	0.012	0.003	[0.006, 0.018]
$(\ln x_t^p)^2$	0.013	0.005	0.003	[-0.001, 0.01]
$(b_0)^2$	0.000	0.001	0.005	[-0.008, 0.01]
L_{t-1}	0.156	0.078	0.024	[0.031, 0.125]
$\ln x_t^c \times \ln x_t^m$	-0.003	-0.015	0.004	[-0.023, -0.006]
$\ln x_t^c \times \ln x_t^p$	0.000	0.004	0.004	[-0.005, 0.013]
$\ln x_t^m \times \ln x_t^p$	0.008	0.010	0.003	[0.003, 0.016]
$b_0 \times \ln x_t^c$	-0.002	-0.001	0.006	[-0.012, 0.01]
$b_0 \times \ln x_t^m$	0.001	0.000	0.004	[-0.009, 0.008]
$b_0 \times \ln x_t^p$	-0.002	0.004	0.004	[-0.005, 0.013]
I_t	0.029	0.087	0.014	[0.061, 0.114]
$I_t \times \ln x_t^c$	0.010	0.024	0.010	[0.005, 0.043]
$I_t \times \ln x_t^m$	0.001	-0.006	0.007	[-0.02, 0.009]
$I_t \times \ln x_t^p$	-0.015	-0.036	0.008	[-0.052, -0.019]
$I_t \times b_0$	0.006	0.000	0.009	[-0.018, 0.018]
Intercept	-0.135	-0.179	0.026	[-0.229, -0.128]
σ_v^2	1.268	1.234	0.008	[1.218, 1.25]

Figure 12: Fit of Measures of Physical Health



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.