The Evolution of Health over the Life Cycle *

Roozbeh Hosseini[†]
University of Georgia
FRB Atlanta

Karen A. Kopecky[‡]
FRB Atlanta
Emory University

Kai Zhao[§]
University of Connecticut

May 24, 2019

Abstract

We construct a unified objective measure of health status: *the frailty index*, defined as the cumulative sum of all adverse health indicators observed for an individual. First, we show that the frailty index has several advantages over self-reported health status, particularly when studying health dynamics. Then we estimate a stochastic process for frailty dynamics over the life cycle. We find that the autocovariance structure of frailty in panel data strongly supports a process that allows the conditional variance of frailty shocks to increase with age. Our frailty measure and dynamic process can be used by researchers to study the evolution of health over the life cycle and its economic implications.

Keywords: health, frailty index, life cycle profiles
JEL Classification numbers: I10, I14, C33

*We would like to thank Jordan Herring for outstanding research assistance and Toni Braun, Eric French, Melinda Pitts, and Dan Waggoner for insightful comments and suggestions. We also benefited from comments received at the 2018 Midwest Macro Meetings, the 2018 SED, the 2018 Michigan Retirement Research Center Workshop and seminar participants at FRB Atlanta and UGA. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

[†]Department of Economics, Terry College of Business, University of Georgia, 620 South Lumpkin Street, Athens, GA 30602

[‡]Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St. NE, Atlanta, GA 30309

[§]Department of Economics, University of Connecticut, 365 Fairfield Way, Storrs, CT 06269-1063

1 Introduction

Recent studies have identified health dynamics and health shocks as major sources of risk over the life cycle. Health has implications for many economic variables including asset accumulation, labor supply, and income and wealth inequality.[1] Most studies use survey responses on individuals' self-assessed health status to measure health. This assessment is by definition subjective and, as we argue, it is often not consistent across different surveys. Moreover, 'self-reported health status' (SRHS from now on) is always a discrete (category) variable. For example, individuals are asked to describe their health status by reporting a number between 1 and 5, with 1 meaning 'excellent', and 5 meaning 'poor' health. Therefore, an individual who reports the number 1 is considered healthier than an individual who reports the number 2. However, this information does not help us understand how much healthier is a 1 relative to a 2.

In this paper we construct a single, continuous variable called a *frailty index* (or *frailty* for short) that can summarize individual health. The frailty index is simply the accumulated sum of all adverse health events that an individual has incurred. Our construction is inspired and based on findings in the gerontology literature.[2] The idea behind the construction of the frailty index is as follows. As individuals age, they accumulate health problems. These health problems can range from symptoms to clinical signs, and laboratory abnormalities to diseases and disabilities. Each health problem is referred to as a *deficit*. Mitnitski et al. (2001) and Mitnitski et al. (2002) have demonstrated that health status can be represented by combining deficits in an index variable, called a frailty index. Mitnitski et al. (2005) and Goggins et al. (2005) find that the frailty index is comparable between databases even when the list of deficits used to construct the index do not coincide. They also find that the frailty index is a better predictor of mortality and institutionalization than age.

Following the guidelines described in Searle et al. (2008), we construct a frailty index for individuals using three different datasets: the Panel Study of Income Dynamics (PSID), the Health and Retirement Study (HRS) and the Medical Expenditure Panel Survey (MEPS). All three datasets contain a rich set of survey questions on various aspects of individual health conditions. In each case, we normalize the frailty index to be a variable between 0 and 1. Therefore, a frailty index of 0.2 means that a person has accumulated 20 percent of all deficits potentially observed.

We start by comparing the frailty index to SRHS. All three datasets that we use collect responses about SRHS by asking individuals to assess their own health using a number between 1 and 5 which correspond to 'excellent', 'very good' , 'good', 'fair', and 'poor' health. We show that the frailty index has several advantages over SRHS, especially when studying health dynamics over the life cycle.

First, SRHS underestimates the average rate of deterioration of objective health (as measured by the frailty index) with age. Specifically, we document that the fraction of good health individuals in the population declines faster with age when health is measured via the frailty index as compared to SRHS. To establish this fact, we identify cutoff points of frailty

[1] See De Nardi et al. (2017), Blundell et al. (2017), O'Donnell et al. (2015), Kopecky and Koreshkova (2014) and De Nardi et al. (2010), among many others.

[2] See Searle et al. (2008); Rockwood and Mitnitski (2007); Rockwood et al. (2007); Mitnitski et al. (2001, 2005); Kulminski et al. (2007a,b); Goggins et al. (2005); Woo et al. (2005), among others.

using the frailty distribution of 25 to 29 year-olds. The cutoff points are chosen to partition the distribution into five bins where the size of each bin is equal to the fraction of individuals in each SRHS category. We then use these fixed cutoffs to assign health status to individuals at older ages. The result is that the fraction of the population in the 'excellent' and 'very good' categories declines much faster with age when these categories are constructed using the frailty index (instead of being self-reported). This finding suggests that older individuals may be overly optimistic in their assessment or reporting of their health status. It is also possible that individuals assess their health by comparing to their peers and, thus, SRHS is relative to a reference point that is declining with age. Consistent with this theory, we document more persistence in health status when health groups are determined using frailty indices as opposed to SRHS.

Second, the frailty index is a more consistent measure of health when comparing across datasets. The distribution of SRHS evolves very differently in MEPS than in PSID and HRS. In contrast, in all three datasets, the dynamics of the frailty distribution are very similar. The frailty dynamics are similar despite the fact that the set of deficit variables that we use to construct the frailty indices is not exactly the same across the three datasets.

Third, the frailty index measures health on a finer scale than SRHS. We exploit the richer variation in frailty as compared to SRHS to document several facts about how cross-sectional dispersion in health evolves with age. We find that dispersion in frailty increases with age and that the frailty distribution is significantly right-skewed. We also document substantial variation in frailty within the 'poor' self-reported health status category. This finding suggests that 'poor' self-reported health status is a weak indicator that an individual is in extremely bad health.

Finally, we demonstrate that, compared to SRHS, frailty is a better predicator of major health-related outcomes. In particular, the frailty index outperforms SRHS in predicting mortality, nursing home entry and Social Security Disability Insurance recipiency. Frailty can also help to account for variation in health outcomes within SRHS groups. In particular, we find a statistically significant positive correlation between frailty and each of these health outcomes within the group of individuals with 'poor' SRHS.

Next, we use the frailty index to measure the evolution of individual health over the life cycle. To this end, we first exploit the long panel dimension of the PSID to document several properties about the dynamics of the cross-sectional frailty distribution. Specifically, we show how the empirical variance and covariances of frailty vary with age. We then explore which types of statistical models of frailty dynamics are consistent with these patterns. In the process, we estimate the models via a GMM estimation that identifies the model parameters by targeting the empirical variance-covariance profile.

The variance of frailty is increasing and slightly convex in age. We start with a stochastic process that has the ability to match this feature of the data. The macro/labor literature on estimating earnings processes has favored models in which the residual consists of an AR(1) process, a transitory shock, and a fixed effect because these processes are easier to embed into structural life cycle models (see, for example, Storesletten et al. (2004)). Drawing on the earnings process estimation literature, we assume a similar model for frailty dynamics.

The autocovariances of frailty are declining in lag length and the rate of their decline is increasing with age. We show that the baseline model can only match the autocovariance structure if we allow for a time-varying conditional variance. To this end, we estimate two

versions of the baseline model. In the *restricted* version, we assume that shocks to frailty have a constant age-invariant variance. Under this view the increasing variance of frailty with age is driven by persistent (perhaps even unit root) shocks. The restricted version cannot simultaneously match both the convex variance profile and declining autocovariance profiles. Thus, our preferred version is the *unrestricted* one which allows for a linear trend in age in the variance of innovations to the AR(1) shocks. This version is consistent with the view that as individuals age, they face higher risk of adverse events, even after controlling for observable characteristics.

We also explore two alternative, richer, model specifications: age-varying persistence and heterogeneous profiles (similar to Guvenen (2009)). We find no significant improvement in fit of the model after adding either of these features. In particular, in the absence of age variation in the variance of AR(1) innovations, neither of the model specifications can simultaneously match both the increasing pattern of variance in frailty and the fact that autocovariances are declining with lags and the rate of decline is increasing with age.

Mitnitski et al. (2006) also estimate a frailty process. Specifically, they estimate a stationary discrete Markov process of frailty using two waves of the Canadian Study of Health and Aging (CSHA). Although their statistical model is elegant and simple, it is essentially an age-invariant autoregressive process. Our estimation results indicate that an age-invariant autoregressive process cannot (simultaneously) match the qualitative features of the life cycle variance profile and the autocovariance profiles. This finding illustrates the value of having a long panel. The autocovariance profiles cannot be obtained from a short panel. However, these moments are extremely informative about underlying data-generating process.

Our paper contributes to the quantitative literature that studies health dynamics over the life cycle and their implications. For instance, De Nardi et al. (2017) estimate a process for health dynamics that allows for history dependence. They then quantitatively evaluate the lifetime consequences of bad health in a structural life-cycle model. Cole et al. (2019) study the effect of labor and health insurance market policies on the the evolution of the cross-sectional health distribution. Capatina (2015) quantifies the impact of health status on labor supply, asset accumulation and welfare. French and Jones (2011) use a structural model of health and labor supply to estimate the effect of health insurance on retirement behavior. All of these studies, like most of the studies in the literature, use SRHS to measure health status.[3] We propose the frailty index as an alternative method and illustrate that it has several attractive relative to SRHS when studying health dynamics.[4]

There are a number of papers in the literature that use objective health condition variables, other than frailty, to measure health status.[5] For instance, Gilleskie et al. (2017) use

[3]One notable exception to the use of SRHS is Dalgaard and Strulik (2014) who, also inspired by the gerontology literature, model health evolution over the lifecycle as a deterministic process of deficit accumulation to study the cross-country link between longevity and income known as the Preston curve. This model has been used by Schünemann et al. (2017a) to study the role of gender-specific preferences in accounting for gender differences in life expectancy and by Schünemann et al. (2017b) to study the impact of deteriorating health on the value of life. Another notable exception is Ozkan (2017) who estimates the health shock process by targeting survival probabilities and medical expenditures.

[4]Though it has been documented in the literature that SRHS is highly correlated with objective measures and is a strong predictor of mortality risk (see, for example, Idler and Benyamini (1997), Van Doorsaler and Gerdtham (2002)), the limitations of SRHS, in particular for life-cycle analysis, still remain.

[5]See Bound (1991), Smith (2004).

body mass to measure health status and study the impact of health status on wages in a life-cycle model. Amengual et al. (2017) construct an objective discrete measure of health using information on Activities of Daily Living (ADL's) and Instrumental Activities of Daily Living (IADL's) in the HRS. They estimate a panel Markov switching model of old-age health dynamics.[6] By using objective indicators of health conditions, these latter studies avoid the disadvantages of subjective self-reported health measures. However, as argued by Blundell et al. (2017), the objective health indicators used in these studies provide an incomplete view of health since they only cover a subset of health conditions. The frailty index, in contrast, serves as a comprehensive summary of an individual's overall health status.

Poterba et al. (2017) also construct an objective health measure for HRS respondents using a similar set of variables to ours and principle component analysis. Our constructed frailty measure is similar in the sense that it is a summary statistic that captures the variations in a collection of indicator variables. The advantage of equally-weighting the deficit variables when constructing the frailty index is that, aside from its simplicity, it directly corresponds to the notion of deficit accumulations.[7]

This paper is also related to the literature on estimating earnings and medical expenditure processes.[8] Our statistical analysis of the underlying stochastic processes for the frailty index draw heavily from this literature, which has favored simpler models so that the estimated processes can be easily incorporated into quantitative life-cycle models. Following this tradition, we model the frailty residual as an AR(1) process plus a transitory shock. Recently, several papers have documented that the conditional distribution of persistent labor income shocks is non-stationary and time or age-dependent including Baker and Solon (2003), Blundell et al. (2015), De Nardi et al. (2018), Guvenen et al. (2015), Karahan and Ozkan (2013), and Meghir and Pistaferri (2004). For instance, Karahan and Ozkan (2013) assume that the conditional distribution of the persistent component of earnings is age-dependent, i.e. both the persistence and the variance of the innovations to the persistent shock vary with age. Karahan and Ozkan (2013) and De Nardi et al. (2018) show that these features of earnings are important for understanding the impact of earnings shocks on consumption and how the cross-sectional dispersion in consumption varies with age in the data. They also show that this age-variation in the persistence and the variance matter for the welfare costs of earnings risk. Fella et al. (2017) explore methods for discretizing non-stationary processes that are applicable to discretizing the frailty process proposed in this paper.

The rest of the paper is organized as follows. In Section 2 we present the frailty index, discuss its construction, and compare it to SRHS. In Section 3, we present and estimate a dynamic stochastic process for frailty over the life cycle. In this section we also present results from estimating the baseline model on subsamples that vary by gender and education.

[6]It is worth noting that Amengual et al. (2017) argue that their discrete measure has an advantage over a continuous measure as the latter cannot be included in structural models. We would argue that it is in fact a disadvantage as it is less flexible than a continuous measure like ours. One can always discretize a continuous process but not so obvious how to go the other way.

[7]In the Appendix we show that the properties of the dynamics of the cross-sectional health distribution we document are very similar whether the measure of health is frailty or a health index constructed using the first principal component as weights as in Poterba et al. (2017).

[8]See, for example, Storesletten et al. (2004) and Guvenen (2009) for the estimation of earnings processes, and Hubbard et al. (1995) and French and Jones (2004) for estimating medical expenses processes. See also Jung and Tran (2014) who document facts about medical expenditures over life cycle using MEPS data.

The last part of the section compares the baseline estimation results to those obtained from estimating alternative statistical models. Section 4 concludes.

2 Frailty Index

As individuals age they develop an increasing number of health problems, functional impairments, and abnormalities. Some of these conditions are rather mild (e.g., reduced vision) while others are serious (e.g., cancer). However, as the number of these conditions rises, the person's body becomes more frail and vulnerable to adverse outcomes. We refer to each of these individual conditions as a *deficit*. In their pioneering work, Mitnitski et al. (2001) and Mitnitski et al. (2002) have demonstrated that the health status of individuals can be represented by combining deficits that an individual accumulates into an index variable, called the *frailty index*. The index is constructed as the ratio of deficits a person has accumulated to the total number of deficits considered. For example, if 30 deficits were considered and 3 were present for a person, that person is assigned a frailty index of 0.1.

Despite its simplicity Mitnitski et al. (2004) and Mitnitski et al. (2005) (among others) have found that having a higher frailty index is associated with a higher likelihood of an adverse health outcome, such as death or institutionalization.[9] Moreover, these findings have been shown to be robust with respect to the choice of dataset that is used to construct the index and the number of potential deficits that are considered.[10] In other words, it does not matter if study A considered 30 deficits from the set X of deficits and study B considered 40 deficits from set Y. The frailty index constructed using each dataset grows at roughly 3% per year, predicts mortality better than age (Mitnitski et al. (2005) and Goggins et al. (2005)), and hardly anyone in the sample accumulates more than 2/3 of total deficits considered. These findings suggest that the frailty index is a good and robust proxy for health status.

2.1 Frailty Index Construction

Motivated by previous studies, we construct frailty indices for samples of individuals in three different datasets: the Panel Study of Income Dynamics (PSID), the Health and Retirement Study (HRS) and the Medical Expenditure Panel Survey (MEPS). The construction of the indices mostly follows the guidelines laid out in Searle et al. (2008), and uses sets of variables similar to those used to create a frailty index in Yang and Lee (2009).

All three datasets contain a rich set of survey questions on various aspects of individual health conditions. We include the following broad categories of variables in our calculations:[11]

- Restricted activity, difficulty in Activities of Daily Living (ADL) and Instrumental ADL (IADL): such as difficulty eating, dressing, walking across room, etc.
- Cognitive impairment: such as immediate word recall, backwards, counting, etc.

[9] See also Searle et al. (2008); Rockwood and Mitnitski (2007); Rockwood et al. (2007); Mitnitski et al. (2001, 2005); Kulminski et al. (2007a,b); Goggins et al. (2005); Woo et al. (2005).

[10] Especially when at least 30 conditions are included, see Kulminski et al. (2007a).

[11] See the Appendix for a complete list of variables used in each dataset.

- Medical diagnosis/measurement: such as high blood pressure, diabetes, heart disease, cancer, high BMI, etc.

We conduct most of our analysis using PSID data. However, for the purpose of comparison and to demonstrate consistency and robustness of our findings we also repeat the analysis in HRS (only available for older individuals) and MEPS (only two frailty observations per individual). Below we briefly describe these datasets and our samples.

2.1.1 Panel Study of Income Dynamics

The PSID is a longitudinal panel survey of U.S. families that was started in 1968. Its disability and health-related questions were expanded in 2003 to include questions on specific medical conditions, ADL's and IADL's. We rely on these questions to construct individuals' frailty indices. For this reason we restrict our sample to the 2003 to 2015 period. The PSID is biennial over this period. We also restrict the sample to household heads and their spouses who are at least 25 years of age. Our sample consists of 84,884 observations of 18,524 individuals (8,738 men and 9,786 women).

Table 23 in the Appendix lists the 27 variables we used to construct the frailty index for PSID respondents. The index is constructed by summing the variables in the first column of the table using their values which are assigned according to the rules in the second column. Then dividing this sum by the total number of variables observed for the individual in the year.

The second column of Table 1 shows summary statistics on the frailty index in the PSID sample. The table shows that mean frailty is higher among the sample of women versus men and higher in older age groups versus younger age groups. It also shows that the distribution of frailty is right-skewed and that increases in frailty are three times more common than decreases.

2.1.2 Health and Retirement Survey

The HRS is a biennial longitudinal survey of Americans over age 50. We use the HRS waves spanning the period 1998 to 2014. Our sample consists of 205,711 observations of 36,032 individuals (15,860 men and 20,172 women). Table 24 in the Appendix lists the 36 variables we used to construct the respondents' frailty index values. The index is constructed in the same way as for PSID respondents. The advantage of HRS over PSID is that it contains a larger number of deficit variables. Specifically, the HRS includes information about cognitive impairment which is not included in PSID. The disadvantage, however, is that, aside from spouses of respondents, it does not survey individuals under the age of 51.

The third column of Table 1 shows summary statistics on the frailty index in the HRS sample. The table shows similar patterns for the frailty distribution in HRS as in PSID. There are two main differences. First, the HRS is a sample of older individuals so mean frailty is higher. Second, both positive and negative changes in frailty across waves are much more common in the HRS than in the PSID. There are two important differences between the HRS and the PSID that help to explain this second difference. First, some of the deficit variables in the HRS, namely the cognitive variables, take on values other than 0 and 1 and naturally fluctuate because they are test scores. Second, the denominator of the frailty index

	PSID	HRS	MEPS
Mean	0.11	0.21	0.11
Mean by demographics groups			
males	0.10	0.19	0.09
females	0.12	0.22	0.11
ages 25-49	0.08	NA	0.06
ages 50-74	0.14	0.19	0.14
ages 75+	0.25	0.28	0.24
Standard deviation	0.11	0.16	0.14
Min	0.00	0.00	0.00
5th percentile	0.00	0.04	0.00
50th percentile	0.07	0.17	0.04
95th percentile	0.33	0.53	0.45
Max	0.92	0.97	0.98
Movement in frailty across waves			
fraction positive change	0.30	0.58	0.41
fraction negative change	0.09	0.40	0.34
Effect of 1 additional deficit	+0.037	+0.028	+0.037

Table 1: Frailty index summary statistics in the PSID, HRS, and MEPS samples. The second and third rows from the bottom are the fraction of all consecutive observations of frailty that show positive and negative change. One minus the sum of these numbers is the fraction in which there is no change in frailty. The bottom row is the effect of accumulating one additional deficit on an individual's frailty index value. It is equal to one over the total number deficits observable in the dataset.

is not constant over time within individuals in HRS due to occasional missing observations. Fluctuations in frailty driven by these factors tend to be very small. If we only count changes in frailty that are greater than or equal in magnitude to the effect of incurring one additional deficit than positive changes account for 37% of movements and negative changes account for 21%.

2.1.3 Medical Expenditure Panel Survey

The MEPS consists of a collection of rotating two-year panels. We use MEPS data from the 2000 to 2016 period. Our sample consists of respondents aged 25 to 84 years. We do not include individuals aged 85 years or older because, starting in 2001, MEPS top codes age at 85. The base sample contains 345,022 observations on 191,165 individuals (88,389 men and 102,776 women). Table 25 in the Appendix lists the 27 variables we used to construct respondents' frailty index values. The index is constructed in the same way as for PSID and HRS respondents. One advantage of the MEPS sample is its large number of observations. However, because MEPS is a two-year rotating panel, it only has two frailty observations, at most, per individual.

The fourth column of Table 1 shows summary statistics on the frailty index in the MEPS

sample. The statistics show that the frailty distributions in the MEPS and the PSID have the same mean and similar properties in general. The only large difference between the PSID and the MEPS statistics is that, like the HRS, the MEPS has more changes in frailty across waves. This occurs in the MEPS sample for the same reasons as in the HRS sample. Thus, a more comparable way to compute the changes in the MEPS data is to only count changes in frailty that, in magnitude, are greater than or equal to the effect of incurring one additional deficit. Using this metric to identify changes, positive changes only account for 20% of movements and negative changes only account for 12%.

2.2 Frailty vs Self-reported Health Status

One of the most commonly used measures of health status is self-reported health status (henceforth, SRHS). In all the surveys we use (PSID, HRS and MEPS) respondents are asked to assess their own health by reporting its category as either 'excellent', 'very good', 'good', 'fair' or 'poor'. In this section we briefly compare the frailty index with SRHS. We point out a few advantages of the frailty index relative to SRHS. These advantages make the frailty index an attractive choice for quantitative or statistical analysis, particularly, when studying the dynamics of health status over the life cycle.

The frailty index is by construction an objective measure of health that is easily comparable across different surveys (in the same way that medical expenditure or labor earnings is comparable). Like SRHS, the frailty index is a ranking of individuals (higher frailty means poorer health). However, in contrast to SRHS, the magnitude of the difference in the frailty index between two individuals is informative about how much healthier one is relative to the other.[12] Another desirable feature of the frailty index is that it can be treated as, or approximated by, a continuous variable. This feature is particularly useful in statistical analysis or economic modeling.

These qualitative features are not the only advantages of the frailty index over SRHS. The frailty index also gives a more accurate picture of how an individual's health evolves with age. To make this point concrete, we compare and contrast how the frailty index and SRHS evolve over the life cycle. In each case we illustrate the main point using our constructed frailty index and the survey responses on SRHS in the PSID. In the Appendix we show that the same conclusions hold if instead we use the HRS or MEPS.

2.2.1 Evolution of health status over the life cycle

We start by comparing the evolution of the frailty distribution with the evolution of the SRHS distribution over the life cycle. To facilitate the comparison between the frailty index (a continuous variable) and SRHS (a category variable), we partition individuals within each 5 year age group into five frailty categories. We label these categories 'excellent', 'very good', 'good', 'fair', and 'poor'. The cutoff values of frailty that determine which category is assigned are age-independent and determined such that the distribution of individuals across frailty categories and SRHS categories is the same for the 25-29 year-old age group. For example, the fraction of 25-29 year-olds with SRHS of 'excellent' is 28%. We set the

[12] For example, a person with frailty index of 0.2 has accumulated twice as much deficits as a person with frailty index of 0.1.

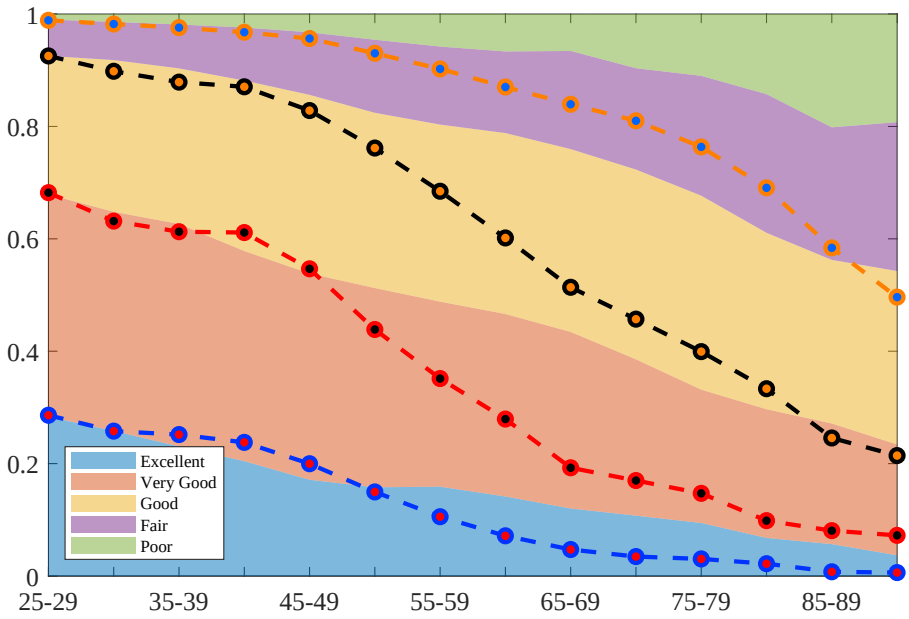


Figure 1: Distribution of health status by age. The colored areas show the fraction of individuals by SRHS at each age. The dashed lines show the fraction of individuals by frailty category at each age. Source: authors' calculation using PSID.

cutoff value for 'excellent' frailty such that 28% of 25-29 year-olds are also in the 'excellent' frailty category. The resulting cutoff value of frailty is 0.04. At each age, individuals with a frailty index value less than 0.04 are assigned to the frailty category 'excellent'. [13] Next, we find the cutoff value for the 68th percentile (68% of 25-29 year-olds have a SRHS of 'excellent' or 'very good'). This frailty cutoff is 0.07. In each age group, anyone whose frailty is larger than 0.04 but smaller than 0.07 is assigned to the frailty category 'very good' and anyone whose frailty value is exactly 0.07 is randomly assigned to either 'very good' or 'good'. The other two cutoffs are chosen accordingly at the 93rd and 99th percentiles and determine the assignment of the remaining individuals to the 'good', 'fair' and 'poor' frailty categories. Using this procedure the frailty categories and SRHS categories of the 25-29 year-old age group are perfectly aligned (by construction).

The shaded areas in Figure 1 show how the distribution of SRHS evolves with age. For each age group, the height of each shaded area is the fraction of individuals in the corresponding SRHS category. As expected, the fraction of individuals with 'excellent' or 'very good' SRHS falls with age (going from 68% for age group 25-29 to less than 25% for age group 90-94). At the same time, the fraction of individuals with 'fair' or 'poor' SRHS increases with age (going from 7% for age group 25-29 to 46% for age group 90-94). There is also a small increase in the share in the middle group, i.e., those with SRHS of 'good' (from 25% to 31%).

The dashed lines in the figure show how the distribution of frailty evolves with age when

[13] Individuals with a frailty index value that is equal to 0.04 are randomly assigned to either the frailty category 'excellent' or the frailty category 'very good' such that, on average, 28% of 25-29 year-olds end up in the frailty category 'excellent'.

Transition Probabilities (%)					
Self Reported Health Status					
	'excellent'	'very good'	'good'	'fair'	'poor'
'excellent'	56.5	32.3	9.1	1.6	0.5
'very good'	13.9	57.2	25.0	3.3	0.6
'good'	4.2	24.7	54.7	14.1	2.3
'fair'	1.6	7.0	28.9	48.9	13.6
'poor'	0.8	1.7	9.0	29.7	58.9
Health Status by Frailty Index					
	'excellent'	'very good'	'good'	'fair'	'poor'
'excellent'	66.3	28.7	4.5	0.4	0.2
'very good'	9.0	60.1	28.1	2.4	0.4
'good'	0.3	14.4	62.7	20.9	1.6
'fair'	0.0	0.4	13.4	69.3	16.8
'poor'	0.0	0.2	0.7	13.2	86.0

Table 2: Transition probabilities across health status levels. Top panel: SRHS. Bottom panel: Frailty by frailty categories. Source: authors' calculation using PSID data.

individuals are assigned to frailty categories using the method described above. As we see, the overall pattern is similar to that of SRHS. The important difference, however, is that the decline in 'excellent'/'very good' shares and rise in 'fair'/'poor' shares happens more rapidly with age when health is measured by frailty instead of SRHS. Up to the 45-49 age group both measures give very similar distributions of health status. More than half of individuals in the 45-49 age group have 'excellent' or 'very good' health according to both SRHS and frailty. However, there is a departure for the older age groups. By ages 70 to 74, only 17% of individuals have a frailty index low enough to fall into the 'excellent' or 'very good' frailty category. However, 39% report a SRHS of 'excellent' or 'very good'. At the same time 54% of individuals in the 70-74 age group have a frailty index higher than the cut off for 'fair' or 'poor' health, while only 28% of them report SRHS of 'fair' or 'poor'.

Note that the dashed lines are constructed using fixed frailty cutoffs. For example, all individuals in all age groups who are assigned to either the 'excellent' or 'very good' frailty categories (represented by the red dashed line), have a frailty index of less than 0.07. The fact that, after age 49, the fraction of these individuals declines faster than the share of individuals who report SRHS of 'excellent' or 'very good' indicates that in older age groups many individuals may be more optimistic about their health relative to what is implied by objective measures. Another possible explanation is that as individuals age they adjust the reference point they use when assessing their health.[14] Regardless of the explanation

[14]A third possibility is that individuals have private knowledge of their health that is not captured by the frailty index. The fact that SRHS still has a statistically significant effect on health outcomes even after controlling for frailty supports this view. (See Tables 3 through 5 in Section 2.2.4.) However, it is unlikely that individuals' private knowledge systematically points to better health status than that inferred from their frailty index. In fact, the regression results in the tables suggest the opposite, namely, that when individuals have private information about their health it is private information that their health is worse than what it inferred from their frailty index.

for this discrepancy, health status appears to depreciate much more rapidly with age when measured by the frailty index as opposed to SRHS. We interpret these patterns as evidence that SRHS underestimates the decline in observable health. This conclusion is not specific to PSID data. As we demonstrate in the Appendix, one arrives at the same conclusions by comparing the frailty index distributions by age with the SRHS distributions by age in the HRS and MEPS.

2.2.2 Persistence of health status

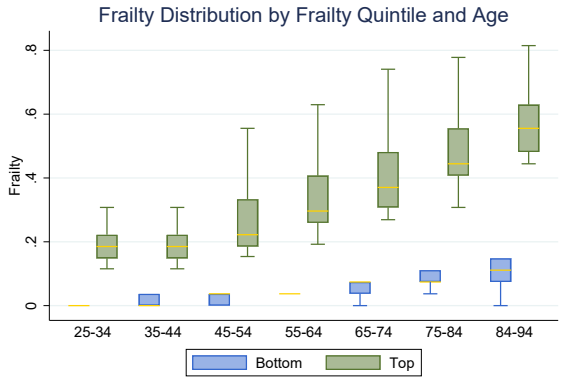
Next, we compare the persistence of SRHS with that of frailty. To do this we compute the conditional probabilities of transitioning between SRHS categories and the conditional probabilities of transitioning between frailty index categories across sample periods. Since we observe respondents every two years, we calculate two year transition probabilities. The top panel of Table 2 shows the transition probabilities between different SRHS categories. Conditional on initially being in the SRHS category labeling each row, each number in the row shows the probability of being in the SRHS category labeling the column next period. For instance, conditional on reporting a SHRS of 'excellent', 56.5% of individuals report a SRHS of 'excellent' two years later, 32.3% report a SRHS of 'very good', 9.1% report a SRHS of 'good', and so on. Similarly, the bottom panel of the table shows the transition probabilities between different frailty categories. Recall that we define what it means to be in each frailty category by choosing frailty cutoffs so that these categories coincide with SRHS categories for the age group 25 to 29. Thus the bottom panel of the table is essentially showing the transition probabilities across these cutoff points.

The table shows that the frailty index is more persistent than SRHS. Notice that the diagonal values are all higher for the frailty index relative to SRHS. For example, individuals with frailty category 'excellent' have a 66.3% chance of maintaining this status while individuals with 'excellent' SRHS have only a 56.5% chance. The difference in persistence is largest at the poor health end of the spectrum. Once an individual's frailty index is high enough that he is assigned to the 'poor' frailty category the probability he is there two years later is 86%. In contrast, individuals who report a SRHS status of 'poor' have only 59% chance of reporting poor health two years later.

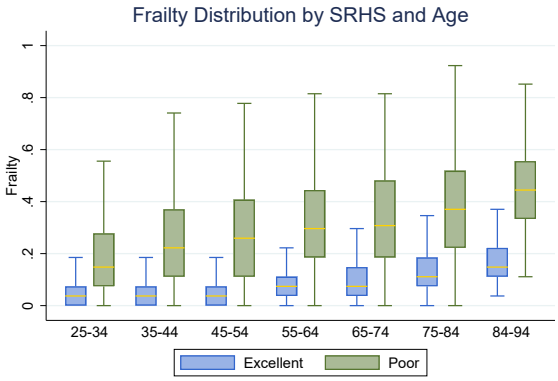
2.2.3 Dispersion in health status by age

The frailty index, by construction, measures health status on a finer scale than SRHS.[15] Thus, measuring health by the frailty index, allows us to study the evolution of the health distribution with age in more detail. The summary statistics in Table 1 indicate that the overall distribution of frailty is right-skewed. This is also the case for the distribution of frailty within smaller age groups. The left panel of Figure 2 shows box and whisker plots of the top (green) and bottom (blue) frailty quintiles within 10-year age groups. As the plot demonstrates, within each age group, there is more variation in frailty among individuals in the top quintiles (the most unhealthy quintiles) than in the bottom quintiles. The plot also shows that dispersion in frailty increases significantly between ages 25 and 74. Notice that

[15] Although it is not exactly a continuous variable, for many practical purposes in can be treated as continuous.



(a) Box and whisker plots of the top (green) and bottom (blue) frailty quintiles by 10 year age groups.



(b) Box and whisker plots of frailty for those who report 'excellent' (blue) and 'poor' (green) SRHS by 10 year age groups.

Figure 2: Dispersion in health by age. Source: authors' calculation using PSID data.

across this age range, not only does mean frailty of the top quintile increase faster than the bottom, but dispersion in frailty within the top quintile also increases.

The right panel of Figure 2 shows box and whisker plots of frailty within the 'poor' (green) and 'excellent' (blue) SRHS categories. The plots are also constructed by 10-year age groups. Even though SRHS and frailty are positively correlated at each age, there is substantial variation in frailty within each SRHS category. Moreover, the variation in frailty is larger among individuals who report 'poor' SHRS versus those that report 'excellent' SRHS. Comparing across the two figures, one can see that mean frailty of individuals in 'excellent' SRHS and in the bottom quintile of the frailty distribution evolve similarly with age. This is not the case when comparing 'poor' SRHS to the top quintile of the frailty distribution. Individuals in 'poor' SRHS have similar levels of frailty, on average, to those in the top quintile of the frailty distribution when young. However, they are significantly less frail at older ages. In this sense, the 'poor' SRHS category is a poor identifier of individuals in the bottom quintile of the health distribution at older ages.

2.2.4 Predicting health outcomes

In previous sections we argued that the frailty index is more suitable than SRHS for tracking the dynamics of health status over the life cycle. But how do they compare in predicting future health outcomes? To answer this question we use HRS data to run three groups of probit regressions. Each group of regressions uses a different health outcome (mortality, nursing home entry, and becoming a social security disability insurance recipient) as the dependent variable.[16]

The first group is a series of probit regressions with mortality as the dependent vari-

[16]We also ran the social security disability insurance regressions in PSID. The results are essentially the same as those found using the HRS. The sample size of elderly people in the PSID is substantially smaller than in the HRS. For this reason, we do not run the mortality and nursing home entry regressions in PSID.

Table 3: Probit regression for mortality at age t

	Panel A. Everyone						**Panel B.** Poor health in $t-1$	
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)
frailty$_{t-1}$			4.096***	3.213***	3.443***	2.278***	0.780***	0.820***
			(0.110)	(0.122)	(0.121)	(0.132)	(0.167)	(0.181)
frailty$^2_{t-1}$			-2.383***	-1.676***	-1.881***	-1.055***	0.677**	0.516*
			(0.152)	(0.164)	(0.159)	(0.171)	(0.209)	(0.223)
very good$_{t-1}$	0.151***	0.097***			0.045	0.040		
	(0.023)	(0.026)			(0.024)	(0.026)		
good$_{t-1}$	0.405***	0.308***			0.150***	0.164***		
	(0.022)	(0.025)			(0.023)	(0.026)		
fair$_{t-1}$	0.698***	0.577***			0.226***	0.298***		
	(0.022)	(0.025)			(0.025)	(0.027)		
poor$_{t-1}$	1.004***	0.918***			0.282***	0.463***		
	(0.024)	(0.027)			(0.028)	(0.030)		
Controls	NO	YES	NO	YES	NO	YES	NO	YES
Observations	167,851	167,851	167,851	167,851	167,851	167,851	49,105	49,105
Pseudo R^2	0.049	0.180	0.088	0.191	0.090	0.196	0.024	0.130

Notes: Panel includes everyone in the sample while panel B only includes those with self reported health status of 'poor'. Controls are gender, education, marital status and quadratic in age. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

able. The results of these regressions are reported in Table 3. Panel A in Table 3 shows the regression results when all individuals are included in the sample. In column (1) the explanatory variables are a set of dummies that indicate whether lagged SRHS is 'very good', 'good', 'fair' or 'poor'.[17] Column (2) shows a similar regression that also includes a set of controls (gender, education, marital status and a quadratic polynomial in age). Columns (3) and (4) show the regression results when SRHS is replaced by a quadratic polynomial in lagged frailty. Finally, columns (5) and (6) show results from regressions that include both a quadratic in lagged frailty and dummies indicating SRHS.

The bottom row of the table reports the *pseudo R-squared* for each regression. It is calculated as one minus the ratio of the *full-model* log-likelihood to the *intercept-only* log-likelihood, or

$$\text{pseudo } R^2 = 1 - \frac{LL\,(\text{Full model})}{LL\,(\text{Intercept only model})}.$$

For each regression, the full model log-likelihood is calculated using all the regressors while the intercept-only log likelihood is calculated using only the intercept (constant) term.[18] We use this pseudo R-squared as a measure of explained variation in the dependent variable.

The pseudo R-squared's in columns (3) and (4) are higher than those in columns (1) and (2), indicating that frailty does better than SRHS in predicting mortality. Although columns (5) and (6) demonstrate that SRHS still has independent predictive power, comparing pseudo R-squared's across columns shows that its additional impact is relatively small.

In Panel B of Table 3 we run the same regressions as columns (3) and (4) in Panel A, but restrict the sample to those with SRHS of 'poor'. Notice that the frailty index is predictive of mortality even within this subsample of 'poor' SRHS individuals. Thus, the variation in

[17] Since we include a constant term in the regression, one of the SRHS categories is redundant. Therefore, we drop the 'excellent' category.

[18] See McFadden (1974) for more details.

Table 4: Probit regression for entry into nursing home at age t

	Panel A. Everyone						**Panel B.** Poor health in $t-1$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
frailty$_{t-1}$			4.588***	3.458***	5.019***	3.374***	1.604***	1.125***
			(0.212)	(0.245)	(0.232)	(0.262)	(0.298)	(0.341)
frailty$^2_{t-1}$			-2.710***	-1.497***	-3.007***	-1.522***	0.103	0.667
			(0.278)	(0.311)	(0.292)	(0.322)	(0.361)	(0.403)
very good$_{t-1}$	0.130**	0.077			-0.030	-0.011		
	(0.042)	(0.050)			(0.045)	(0.052)		
good$_{t-1}$	0.298***	0.198***			-0.085	-0.027		
	(0.040)	(0.048)			(0.045)	(0.051)		
fair$_{t-1}$	0.535***	0.421***			-0.151**	0.001		
	(0.040)	(0.048)			(0.047)	(0.054)		
poor$_{t-1}$	0.800***	0.742***			-0.196***	0.088		
	(0.043)	(0.051)			(0.052)	(0.058)		
Controls	NO	YES	NO	YES	NO	YES	NO	YES
Observations	149,230	149,230	149,230	149,230	149,230	149,230	43,478	43,478
Pseudo R^2	0.035	0.222	0.120	0.261	0.121	0.262	0.046	0.197

Notes: Panel includes everyone in the sample while panel B only includes those with self reported health status of 'poor'. Controls are gender, education, marital status and quadratic in age. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

frailty within SRHS groups that is documented in the right panel of Figure 2 is positively correlated with mortality risk. Note that, by construction, zero variation in mortality is explained by SRHS within this subsample.

Next, we look at the relationship between health status and the probability of entering a nursing home. We repeat the previous exercise but replace mortality with nursing home entry as the dependent variable. This means that we restrict the sample to those individuals who are not in a nursing home. The dependent variable is one at age t, if they enter a nursing home at age t and zero otherwise. Table 4 reports the regression results. We observe a similar pattern as above. Frailty is better than SRHS at explaining variations in nursing home entry (as measured by the pseudo R-squared) and continues to have predictive power even when we only consider individuals with SRHS of 'poor'. Moreover, SRHS has close to no impact on predictive power when frailty is also included in the regression. This can be seen by comparing the pseudo R-squared's across columns (4) and (6), or by observing Column (6) in Panel A. Notice that when frailty is included in the regression, SRHS is no longer statistically significant.

Finally, we examine the relationship between health status and becoming a Social Security Disability Insurance (SSDI) beneficiary. To this end we restrict our HRS sample to those younger than 66 years old and not receiving SSDI. The dependent variable is one at age t if they become a SSDI beneficiary and zero otherwise. Table 5 shows the regression results. Once again, frailty explains a larger fraction of variations in the dependent variable (as measured by pseudo R-squared). Also, frailty explains some variations within the sample with common SRHS of 'poor'. While adding SRHS to the frailty regressions does increase predictive power, Columns (5) and (6) in Panel A show that when both frailty and SRHS are included, only 'fair' and 'poor' SRHS are statistically significant.

To summarize, frailty is a strong predictor of health outcomes. Tables 3, 4, and 5 show that it performs better than SRHS at predicting mortality, nursing home entry, and SSDI recipiency. The tables also shows that it can account for variation in these outcomes even

Table 5: Probit regression for going on Social Security Disability Insurance at age t

	Panel A. Everyone						**Panel B.** Poor health in $t-1$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
frailty_{t-1}			7.937***	7.886***	6.456***	6.549***	5.375***	5.573***
			(0.268)	(0.277)	(0.293)	(0.301)	(0.391)	(0.400)
frailty^2_{t-1}			-5.571***	-5.628***	-4.820***	-4.953***	-3.350***	-3.602***
			(0.395)	(0.404)	(0.415)	(0.423)	(0.525)	(0.534)
very good_{t-1}	0.087	0.082			-0.081	-0.071		
	(0.051)	(0.052)			(0.054)	(0.055)		
good_{t-1}	0.473***	0.438***			0.052	0.042		
	(0.047)	(0.048)			(0.052)	(0.053)		
fair_{t-1}	1.060***	0.994***			0.348***	0.324***		
	(0.046)	(0.048)			(0.054)	(0.055)		
poor_{t-1}	1.722***	1.635***			0.647***	0.609***		
	(0.050)	(0.051)			(0.060)	(0.061)		
Controls	NO	YES	NO	YES	NO	YES	NO	YES
Observations	69,438	69,438	69,438	69,438	69,438	69,438	14,450	14,450
Pseudo R^2	0.162	0.181	0.222	0.239	0.239	0.254	0.108	0.123

Notes: Panel includes everyone in the sample while panel B only includes those with self reported health status of 'poor'. Controls are gender, education, marital status and quadratic in age. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

within a sample of individuals with common SRHS (of 'poor').

3 Estimation of Frailty Process

Our goal in this section is to propose and estimate a stochastic process for frailty over the life cycle. The statistical model we propose is designed to be as parsimonious as possible while still being flexible enough to capture the main qualitative properties of frailty dynamics. For instance, the model allows for innovations to frailty to be persistent. The model also allows the variance of frailty to increase with age. Both features are consistent with the findings in Section 2. We first present the model and describe the estimation procedure. Then we describe the empirical moments used to estimate the model and present the estimation results. In the second part of the section we present results from estimating the model separately for different demographic subgroups of the population. The last part of the section shows estimation result from modified versions of the baseline model.

3.1 Baseline Statistical Model

Our statistical model is very similar to ones used to estimate the earning process (see Guvenen (2009), Karahan and Ozkan (2013), and Storesletten et al. (2004) among many others). In particular, we assume that the frailty index f_{it} for individual i at age t is the sum of a deterministic component whose effect is common to all individuals and a residual that is individual-specific:

$$f_{it} = X'_{it}\beta + R_{it}, \tag{1}$$

where X_{it} is a set of covariates including age, age-squared, gender, marital status and education. The set of covariates also includes a full set of year dummies. The residual consists

of two components and is given by

$$R_{it} = \alpha_i + z_{it} + u_{it}. \tag{2}$$

The first variable, α_i, is individual-specific and allows us to capture ex-ante heterogeneity in individuals' initial frailty levels. We assume that α_i is randomly distributed across individuals with mean zero and variance σ^2_α.

The second component captures the dynamics in frailty as individuals go through various random health events over their life cycles. This component is the sum of an AR(1) process and a white noise shock u_{it}. Thus

$$z_{it} = \rho z_{it-1} + \varepsilon_{it}, \tag{3}$$

where $z_{i,0} = 0$.[19] The shocks ε_{it} and u_{it} are assumed to be independent of each other and over time, and independent of α_i. We assume that u_{it} has mean zero and variance σ^2_u and that ε_{it} has mean zero but that its variance is age-dependent. Specifically, we assume that the variance of ε_{it} is given by

$$\sigma^2_{\varepsilon,t} = \delta_{\varepsilon,1} t + \delta_{\varepsilon,0}, \tag{4}$$

where $\delta_{\varepsilon,0}$ is the initial variance level and $\delta_{\varepsilon,1}$ is the rate at which the variance changes with age.

The white noise shock u_{it} captures both measurement error and acute health events such as a temporary inability to walk due to a broken leg. The persistence ρ and the variances of the innovations to the persistent process $\sigma^2_{\varepsilon,t}$ determine individuals' exposure to persistent health shocks. As we discuss in detail below, allowing the variance of the persistent shocks to be age dependent is crucial for matching the qualitative properties of frailty dynamics.

We estimate the model in two stages. First, we estimate β using OLS and compute the residuals R_{it}. Second, we estimate the parameters of the stochastic component, equation (2), using a minimum distance estimator. The procedure minimizes the distance between the variances and covariances of the residuals R_{it}, and their empirical counterparts. This is the GMM estimator proposed by Chamberlain (1984). We estimate the model at the annual frequency. However, since our data is biennial, we can only compute empirical covariances between current frailty residuals and lagged values of frailty residuals that are multiples of 2. To deal with this discrepancy, the minimization procedure simulates the annual model and uses the simulated data to construct model counterparts to the biennial empirical covariances.

Table 6 provides the results from the first stage of the estimation run on our main PSID sample. Frailty is increasing with age and decreasing in years of schooling. Being male decreases frailty by 0.0137 units. To put this number in perspective recall that having one additional deficit increases the frailty index by 0.037 units. Thus, on average, being male reduces the number of deficits an individual has by 0.37 deficits. Being married has a larger

[19] Note that t represents age and not time which means we are assuming that the stochastic component of frailty can vary with age but is time-invariant. The variance of frailty increases with both age and time in both the PSID and HRS samples. However, the increase with age is much more dramatic. Therefore, we chose a specification with an age-dependent but time-invariant stochastic component.

Variable	Coefficient (×100)	Std. Err. (×100)
Age	-0.028	(0.013)
Age^2	0.003	(0.0001)
Years of School	-0.630	(0.014)
Male	-1.371	(0.069)
Married	-3.157	(0.077)
Const.	13.035	(0.379)
Year dummies included		
	$N = 81,664$, $R^2 = 0.250$	

Table 6: OLS regression results for frailty using PSID sample for ages 25–95.

effect on frailty, reducing it by 0.0316 units or 0.85 deficits. The negative effect of marriage on frailty is consistent with the Guner et al. (2017) finding that marriage has positive effect on health.[20]

Figure 3 presents the empirical variances and covariances of the frailty residuals that are targeted in the second stage of the estimation. As is commonly done in the literature on earning dynamics, we use empirical moments that have been adjusted for cohort effects.[21] The left panel shows the cross-sectional variances of the frailty residuals, R_{it}, by age. The panel shows both the raw variances and the cohort-adjusted ones. To construct the variances we group individuals into 2-year, non-overlapping, age groups (25–26 year-olds, 27–28 year-olds, and so on). The raw variance profile is the means of the squared residuals of each age group. To obtain the cohort-adjusted variance profile, we regress the raw variances on a full set of age and cohort dummies to obtain cohort-adjusted squared residuals. To maintain the same level of inequality after cohort effects are removed, the cohort-adjusted variances are rescaled such that the adjusted variance at age 35 is the same as the raw variance at age 35.

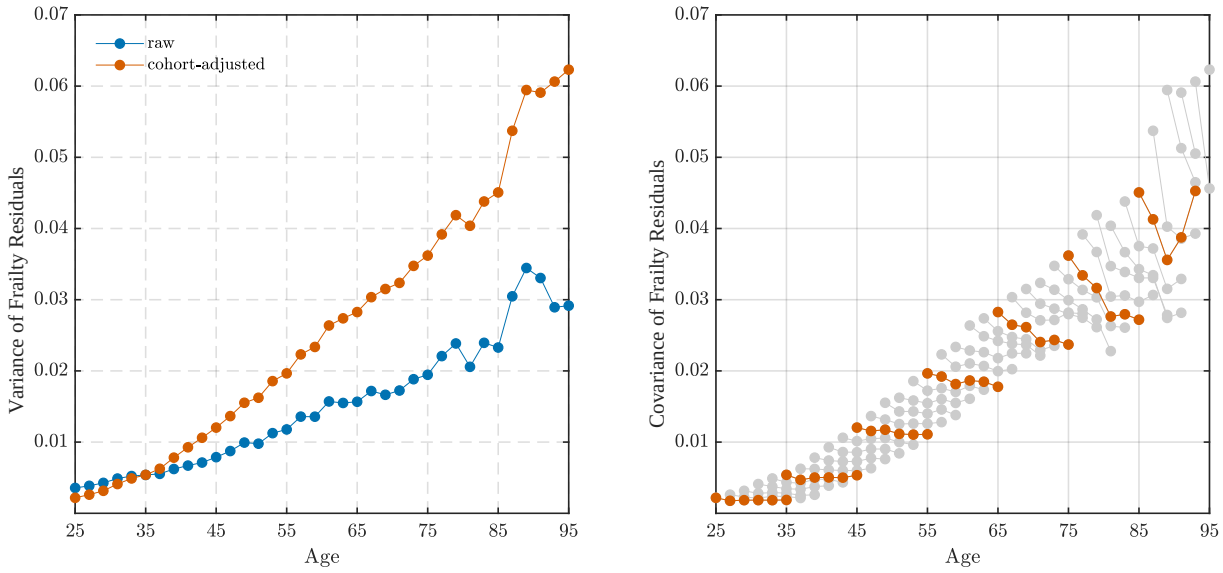
The right panel of Figure 3 shows the entire empirical variance-covariance matrix after adjusting for cohort effects. To get the cohort-adjusted covariances we regress individual-specific moments on cohort and age dummies separately for each age group. We then compute cohort-adjusted individual-specific moments using the residuals and age effects rescaled in the same manner as we rescaled the variances. The cohort-adjusted covariances are the means of these moments for each age group.[22] The first point in each line in the figure is the variance of that age group's frailty residual R_{it} at age t, the next point is the covariance between R_{it} and R_{it-2} followed by the covariance between R_{it} and R_{it-4} and so on.

Several properties of the dynamics of frailty over the life cycle can be observed by studying Figure 3. First, as individuals age, the cross-sectional variance of frailty increases. Second, the rate at which the variance increases with age is slightly higher for older individuals. In other words, the variance age profile is slightly convex in age. Third, the covariance between frailty at age t and frailty at age $t-k$ is declining in the lag length k. Fourth, the rate at

[20] Although they find that the positive effect is primarily due to selection, they also find that protective effects play an important role in older ages. A summary of the literature on the effect of marriage on health is provided by Wood et al. (2009).

[21] See Deaton and Paxson (1994), Guvenen (2009) and Storesletten et al. (2004).

[22] Additional details on the construction of the cohort-adjusted variance-covariance matrix can be found in the Appendix.



(a) (b)

Figure 3: Raw and cohort-adjusted variances (left) and cohort-adjusted variances and covariances (right) of the residuals, R_{it}, by age in the PSID.

which the autocovariances decline with lag length is increasing in age (i.e., the autocovariance profiles become steeper at older ages).

One potential concern is that the increasing rate of decline of the autocovariance profiles with age is due to sample attrition. Highly frail individuals are more likely to die and, as a result, are less likely to contribute to higher-order autocovariances. Since this selectivity bias becomes more severe as the lag length increases it puts downward pressure on the autocovariance structure. While the mortality rates, and hence attrition rates, of working-age individuals are fairly low, retirees are more likely to both be highly frail and to die. Hence the effect of this selectivity bias on the autocovariance structure of retirees is of particular concern. Figure 4 plots the cohort-adjusted variance-covariance matrix of the frailty residuals for a modified version of the sample. The modified sample excludes individuals who exit the baseline sample due to death. The baseline variance-covariance matrix is also plotted for purposes of comparison. Notice that the higher order covariances computed using the modified sample due in fact tend to be larger than those for the baseline after age 75. However, the differences are small and the steepening autocovariance pattern observed under the baseline sample remains. This suggest that the steepening pattern of autocovariances is not due to attrition. In Section C we show that our estimation results are robust to using this alternative set of empirical moments as targets.

Under the statistical model presented in equations (2)-(4), for each individual i, the cross-sectional variance of R_{it} and its covariance with R_{it+k} are given by

$$var(R_{it}) = \sigma_\alpha^2 + var(z_{it}) + \sigma_u^2, \tag{5}$$

$$cov(R_{it}, R_{it+k}) = \sigma_\alpha^2 + \rho^k var(z_{it}), \tag{6}$$

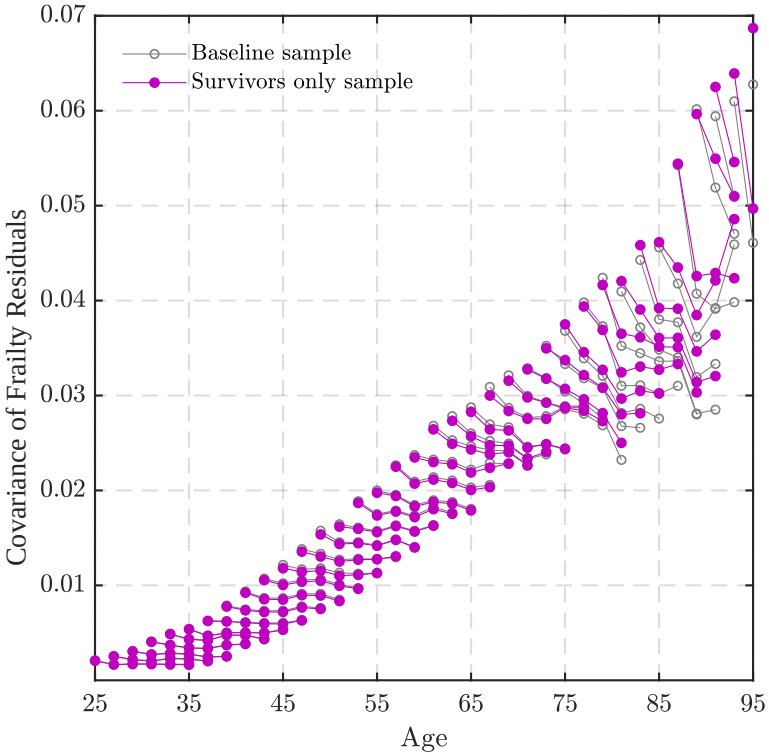


Figure 4: Cohort-adjusted variance-covariance matrix of the residuals, R_{it}, by age in PSID for two versions of the sample. The survivors-only sample is equivalent to the baseline sample except that it excludes individuals who exit the baseline sample due to death. The baseline sample variance-covariance matrix is provided for purposes of comparison.

where

$$var(z_{it}) = \rho^2 var(z_{it-1}) + \sigma^2_{\varepsilon,t} = \delta_{\varepsilon,0} \sum_{j=0}^{t-1} \rho^{2j} + \delta_{\varepsilon,1} \sum_{j=0}^{t-1} \rho^{2j}(t-j).$$

Equations (5) and (6) show how the variance and covariance age profiles depend on model parameters. From these equations one can see that the baseline statistical model has the ability to replicate the main qualitative properties of the empirical variance-covariance matrix presented in Figure 3. First, notice that $var(R_{it})$ will be increasing in age if $\delta_{\varepsilon,1} \geqslant 0$. Second, notice that the second term in $var(z_{it})$ is always convex in age. If ρ is greater than 1, then the first term will also be convex in age and so will $var(R_{it})$. If ρ is less than 1, then the first term is concave in age and the convexity of $var(R_{it})$ will depend on the parameterization. Third, notice that the model can generate covariances that decline with lag length, i.e., satisfy $cov(R_{it}, R_{it+k+1}) < cov(R_{it}, R_{it+k})$ for all k, if ρ is less than 1. Finally, notice that the rate at which the covariances decline with lag length will increase with age as long as $var(z_{it})$ is increasing in age, which is the case when $\delta_{\varepsilon,1} \geq 0$.

Allowing the variance of the persistent shock to vary with age is essential if the statistical model is to match the main qualitative features of the empirical variance-covariance matrix. To illustrate this point, we estimate two versions of the model. Our preferred version puts no

ρ	$\sigma_{\alpha}^{2\,\mathrm{a}}$	$\sigma_{u}^{2\,\mathrm{a}}$	${\delta_{\varepsilon,0}}^{\mathrm{a}}$	${\delta_{\varepsilon,1}}^{\mathrm{a}}$
A. Restricted				
1.008	4.911	5.533	3.968	–
(0.001)	(1.334)	(0.284)	(0.129)	–
B. Unrestricted				
0.989	16.290	4.231	1.900	0.279
(0.001)	(1.457)	(0.296)	(0.203)	(0.016)

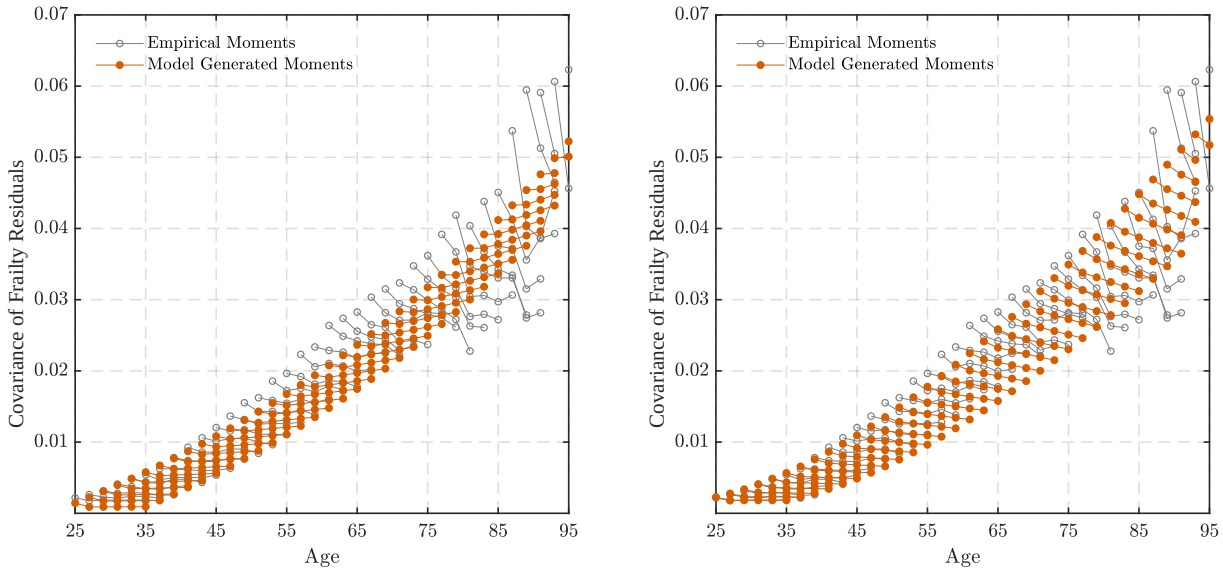
Table 7: Results from estimating the restricted ($\delta_{\varepsilon,1} = 0$) and unrestricted versions of the baseline model using the PSID sample. Standard errors are in parenthesis. The estimation targets the variance and covariance moments in Figure 3. [a]Estimates and standard errors are reported in tens of thousands.

a priori restriction on δ. We call this the *unrestricted* version. In our, alternative, *restricted* version we do not allow for a linear age-trend in the variances of the innovations to the persistent process by setting $\delta_{\varepsilon,1} = 0$. Under this version of the model, $\sigma_{\varepsilon,t}^2 = \delta_{\varepsilon,0}$ at each age t.

Table 7 presents the results from the GMM estimation of the restricted and unrestricted versions of the model. Figure 5 shows the model-predicted variance-covariance matrices together with the empirical variance-covariance matrix. Under the restricted specification, ρ is estimated to be larger than 1 and the hypothesis that $\rho = 1$ is rejected at standard significance levels. This value of ρ is driven by the slightly convex shape of the empirical variance profile. However, notice in the left panel of Figure 5, that under the restricted specification, the model-generated autocovariances are increasing with the lag order which is opposite the pattern in the data. There is a tension in the restricted model between the variance profile and the autocovariance structure. The gradually decaying autocovariances suggest that ρ lies between 0 and 1. However, the slightly convex variance profile can only be achieved if ρ is larger than 1.[23]

The unrestricted specification has the ability to simultaneously match both the slightly convex variance pattern and the decaying auto-covariance pattern. This is because, under the unrestricted specification, the positive linear trend in the variance of the persistent shock can also induce a convex variance pattern in the frailty residuals, as equation (5) shows. Therefore, with ρ less than one, the unrestricted specification can match both the slightly convex variance pattern and the decaying auto-covariance pattern. Consistent with this intuition, allowing the variance of the persistent shock to vary with age reduces the estimated value of ρ. The value of ρ estimated in the unrestricted model falls well below one and we can reject that these shocks are a permanent random walk. The estimated value of $\delta_{\varepsilon,1}$ is positive and statistical significant. As the right panel of Figure 5 shows, the model is able to generate both an increasing and slightly convex variance profile, and declining auto-covariance profiles that steepen with age. The rate of steepening of the auto-covariance profiles is equal to the rate of increase of the variance profile which is gradually converging

[23]As is the case for the restricted and unrestricted models in Guvenen (2009), if the true data-generating process is the unrestricted model, then estimating the restricted model introduces significant upward bias into the estimation of the persistence parameter.



(a) (b)

Figure 5: Fit of the baseline estimation. The orange closed circles are the autocovariance matrices generated by the restricted (left) and unrestricted (right) versions of the baseline model. The gray open circles are their empirical counterparts which are targeted in the GMM estimations.

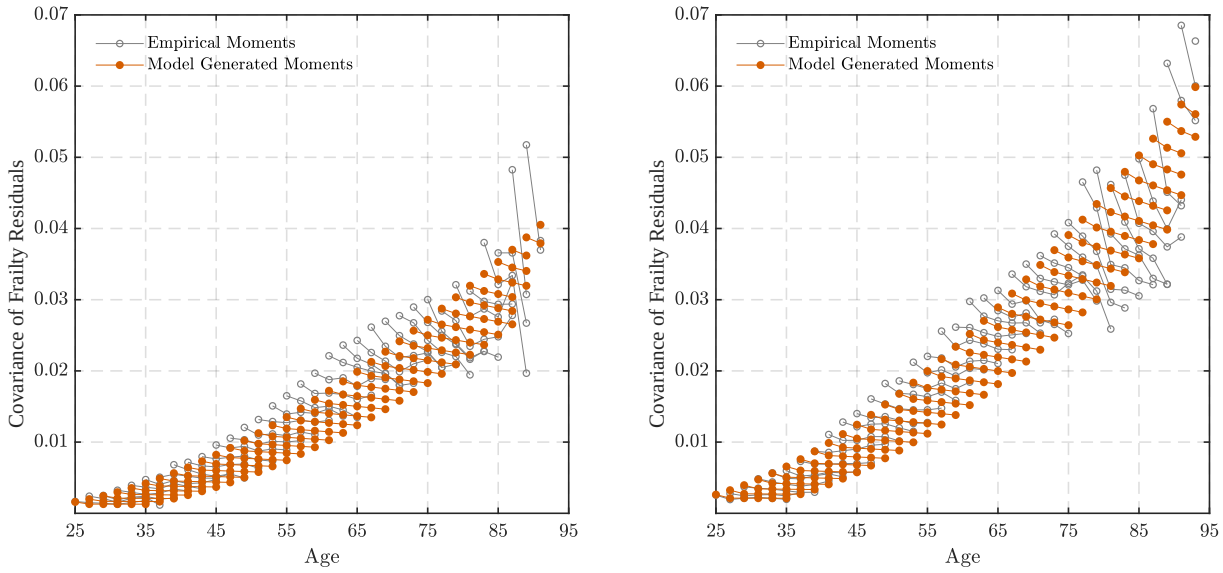
to $\delta_{\varepsilon,1}/(1-\rho^2)$ as age increases.[24]

3.2 Estimation Results from Subsamples

In this subsection, we present results from estimating the unrestricted models separately on subsamples that vary by gender and subsamples that vary by education.[25] Table 8 presents the estimation results for men and women and Figure 6 shows the model-predicted variance-covariance matrices together with their empirical counterparts. There is significantly more variation in frailty among women than men. This additional variation results in larger estimated values of σ_α^2, σ_u^2 and $\delta_{\varepsilon,0}$. The variance of frailty also increases faster with age for women. This is, in part, captured by the estimation through a higher value of $\delta_{\varepsilon,1}$.

[24]This suggests that, allowing the variance of the persistent shock to also be a function of higher-order age terms can improve the model's ability to match both the steeper rate of increase of the variance of frailty and the steeper rate of decline of the auto-covariance profiles at later ages.

[25]We provided results from estimating the restricted version of the model on the subsamples in the appendix. The appendix also contains results from estimating the model on subsamples that vary by both education and gender.



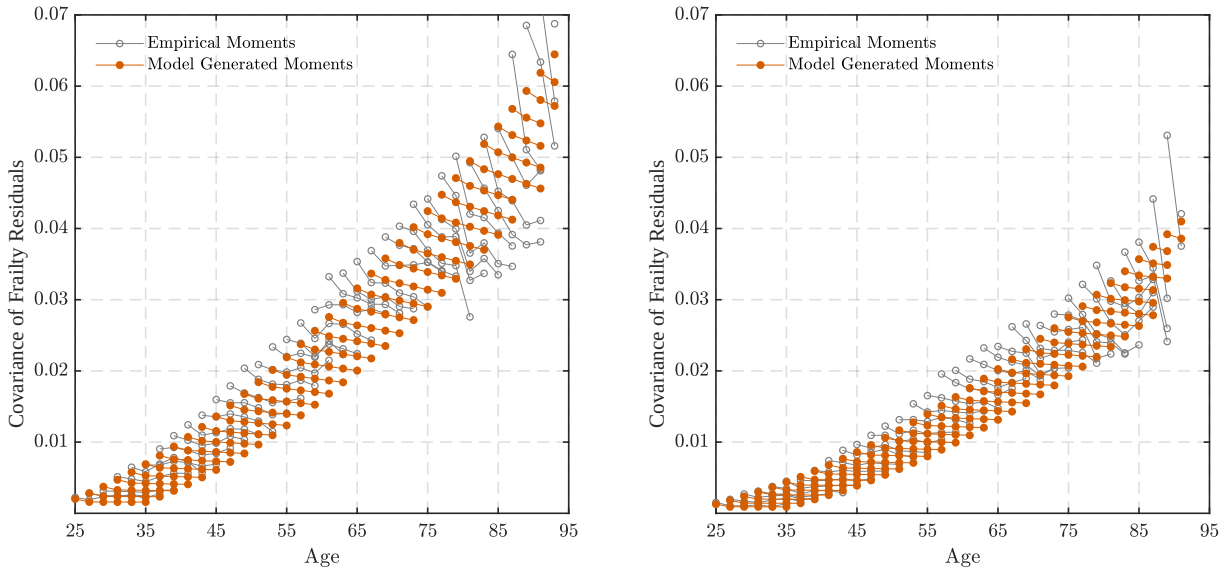
(a) (b)

Figure 6: Fit of the baseline estimation using subsample of males and females. The closed orange circles are the autocovariance matrices generated by the unrestricted version of the baseline model estimated on males only (left) and females only (right). The gray open circles are their empirical counterparts which are targeted in the GMM estimations.

	ρ	$\sigma_\alpha^{2\,\mathrm{a}}$	$\sigma_u^{2\,\mathrm{a}}$	${\delta_{\varepsilon,0}}^{\mathrm{a}}$	${\delta_{\varepsilon,1}}^{\mathrm{a}}$
All	0.989	16.290	4.231	1.900	0.279
	(0.001)	(1.457)	(0.296)	(0.203)	(0.016)
Men	0.993	11.243	2.988	1.519	0.191
	(0.002)	(1.256)	(0.316)	(0.192)	(0.017)
Women	0.992	18.577	4.931	2.483	0.269
	(0.001)	(2.202)	(0.416)	(0.272)	(0.021)

Table 8: Results from estimating the unrestricted version of the baseline model separately for men and women using PSID data. Standard errors are in parenthesis. The estimations target gender-specific . [a]Estimates and standard errors are reported in tens of thousands.

Table 9 and Figure 6 show the results of estimating the unrestricted version of the model separately for different education groups: people with a high school degree (or less) and people with at least some college. There are slightly larger differences between the college and non-college groups than between the gender groups. Variation in frailty is significantly lower within the college group even at age 25. Consistently, the estimation finds that they are ex-ante more homogeneous. Interestingly, the estimated degree of ex-ante heterogeneity is lower within each education group than in the baseline sample. This suggests that some of variation in the individual-specific effects are due to differences in education. Finally, the variation in frailty within the high school group increases more rapidly with age than within the college group. As a result, the growth rate of the conditional variance of their frailty



(a) (b)

Figure 7: Fit of the baseline estimation for different education groups. The closed orange circles are the autocovariance matrices generated by the unrestricted version of the baseline model estimated on the sample of individuals with at most a high school degree (left) and on the sample of individuals some college or more (right). The gray open circles are their empirical counterparts which are targeted in the GMM estimations.

shocks is nearly double.

	ρ	$\sigma_\alpha^{2\,\mathrm{a}}$	$\sigma_u^{2\,\mathrm{a}}$	${\delta_{\varepsilon,0}}^{\mathrm{a}}$	${\delta_{\varepsilon,1}}^{\mathrm{a}}$
All	0.989	16.290	4.231	1.900	0.279
	(0.001)	(1.457)	(0.296)	(0.203)	(0.016)
High school	0.993	12.242	4.233	3.482	0.268
	(0.001)	(2.490)	(0.505)	(0.315)	(0.024)
College	0.997	6.714	3.666	2.280	0.139
	(0.001)	(1.077)	(0.316)	(0.165)	(0.015)

Table 9: Results from estimating the unrestricted version of the baseline model separately for those with a high school degree or less and those with some college or more using PSID data. The estimation targets all the variance moments in Figure 3 and the age 25-65 covariance moments. [a]Estimates and standard errors are reported in tens of thousands.

3.3 Alternative Statistical Models

In this section we discuss two alternative specifications of the statistical model presented in Section 3.1. We show that while allowing for an age-varying conditional variance of the persistent shock is crucial for matching the qualitative properties of the empirical moments,

this is not the case for the other two alternatives we consider. Both allow for a small improvement in the fit of the baseline model when an age-varying variance is also present. However, neither on its own, is able to simultaneously generate both an increasing and convex variance profile and covariance profiles that decrease with age at an increasing rate.

3.3.1 Age-varying persistence

First, we consider a variation of the baseline statistical model that allows the persistence of the AR(1) component to vary with age. The model is identical to the baseline model except that we assume that

$$z_{it} = \rho_{z,t} z_{it-1} + \varepsilon_{it}, \tag{7}$$

and

$$\rho_{z,t} = \gamma_{z,1} t + \gamma_{z,0}, \tag{8}$$

where $\gamma_{z,0}$ is the initial level of persistence and $\gamma_{z,1}$ is the rate the persistence increases with age. Under this specification of the dynamic process, both the variance of ε_{it} and the persistence of z_{it} are allowed to vary linearly with age. The cross-sectional variances and covariances of the residual at age t are given by

$$var(R_{it}) = \sigma_{\alpha}^2 + var(z_{it}) + \sigma_u^2, \tag{9}$$

$$cov(R_{it}, R_{it+k}) = \sigma_{\alpha}^2 + var(z_{it}) \prod_{j=1}^{k} \rho_{z,t+j}, \tag{10}$$

where

$$var(z_{it}) = \rho_{z,t}^2 var(z_{it-1}) + \sigma_{\varepsilon,t}^2 = \delta_{\varepsilon,0} \sum_{j=0}^{t-1} \prod_{i=1}^{j} \rho_{z,t+1-i}^2 + \delta_{\varepsilon,1} \sum_{j=0}^{t-1} (t-j) \prod_{i=1}^{j} \rho_{z,t+1-i}^2.$$

Notice that with a linear time trend in the persistence alone ($\delta_{\varepsilon,1} = 0$), the model cannot simultaneously match both the convex variance profile and the pattern of declining covariances with age. The former requires $\rho_{z,t} > 1$ at all ages while the latter requires $\rho_{z,t} < 1$ at all ages. Moreover, allowing the persistence of the AR(1) shock to vary with age does little in terms of improving the model's ability to generate variance-covariance moments that match those constructed from the data. To demonstrate these two points, as in Section 3.1, we estimate both a *restricted* and *unrestricted* version of the model. Under the *restricted* version, we only allow for age-variation in the persistence of the AR(1) shock. In other words, we shut-down age-variation in its conditional variance by setting $\delta_{\varepsilon,1} = 0$. Under the *unrestricted* version, we put no *a priori* restrictions on the parameters. We estimate these two versions of the age-varying persistence model using the same procedure and targeted empirical moments as for the baseline model.

Table 10 presents the estimation results. Figure 8 plots the model-generated variance-covariance matrices together with their empirical counterparts. The estimated value of $\gamma_{z,1}$ under the restricted version of the model is zero and the estimated values of the other

$\gamma_{z,0}$	$\gamma_{z,1}$	$\sigma_{\alpha}^{2\,a}$	$\sigma_{u}^{2\,a}$	$\delta_{\varepsilon,0}{}^{a}$	$\delta_{\varepsilon,1}{}^{a}$
A. Restricted					
1.008	0.000	4.911	5.533	3.969	–
0.002	0.000	1.341	0.302	0.200	–
B. Unrestricted					
1.005	-0.0004	16.900	4.585	1.081	0.299
(0.003)	(0.0001)	(1.467)	(0.305)	(0.223)	(0.016)

Table 10: Results from estimating the restricted ($\delta_{\varepsilon,1} = 0$) and unrestricted versions of the age-varying persistence model using the PSID sample. Standard errors are in parenthesis. The estimation targets the variance and covariance moments in Figure 3. [a]Estimates and standard errors are reported in tens of thousands.

parameters are essentially the same as those under the restricted version of the baseline model. Allowing for a linear age trend in the persistence of the AR(1) shock, in the absence of an age trend in the variance, does nothing in terms of improving the model's ability to match the data.

Under the unrestricted specification, the estimated value of the initial persistence is above 1, while the estimated value of $\gamma_{z,1}$ is negative. Relative to the unrestricted version of the baseline model, there is a small improvement in overall fit. Compare Figure 8b with Figure 5b. The variance profile of the baseline unrestricted model is convex at all ages. In contrast, with age-varying persistence, the variance profile is relatively more convex at younger ages but concave at older ages. The time-varying persistence allows for a slightly better fit of the (more heavily weighted) variance-covariance moments at younger ages, at the cost of a worse fit at older ages.

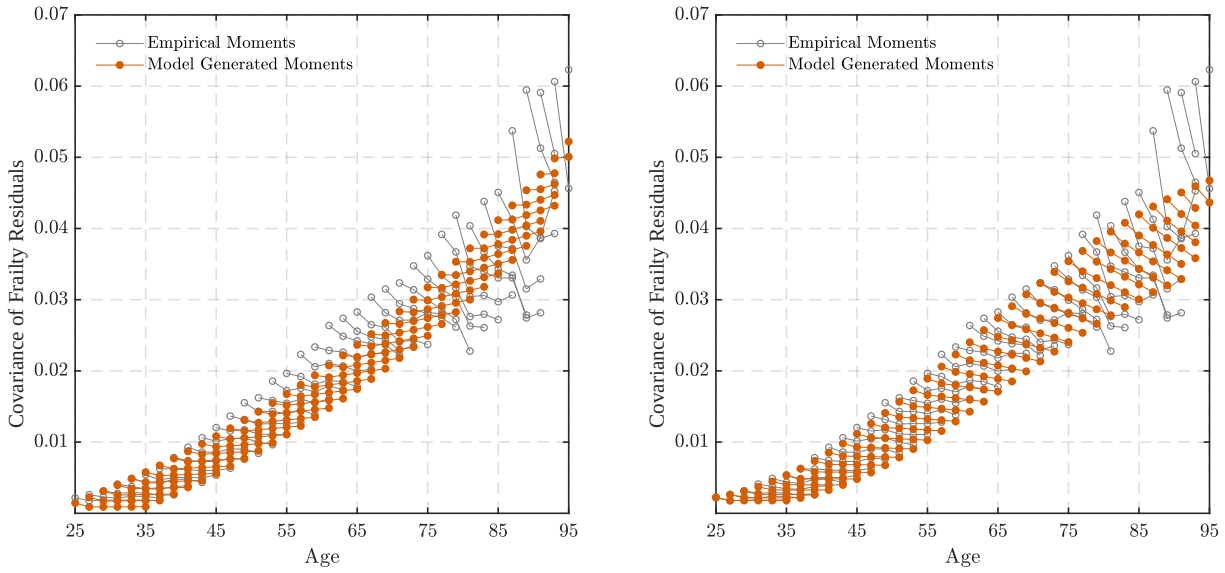
3.3.2 Heterogeneous profiles

In this section we consider a variation of the baseline model presented in Section 3.1 that allows for ex-ante heterogeneous frailty profiles. A highly persistence AR(1) shock is needed to match the empirical variance-covariance profile under the baseline model specification even when the conditional variance of the persistent shock is allowed to vary with age. However, the increasing and convex variance profile observed empirically could, in part, be due to ex-ante heterogeneity in individuals' frailty growth rates. Guvenen (2009) argues that ex-ante heterogeneity in earnings growth rates may be an important source of earnings inequality over the lifecycle. Similarly, individuals could have heterogeneous growth rates of frailty during their adult lives driven by differences in their genes and/or the investments made in their health as children.

We now consider a version of the baseline model that allows for ex-ante heterogeneous frailty profiles. The heterogenous profile model is identical to the baseline model except that we assume that the residual is given by

$$R_{it} = \alpha_i + \gamma_i t + z_{it} + u_{it}. \tag{11}$$

The term $\gamma_i t$ allows for individual-specific effects on the growth rate of frailty. We assume that (α_i, γ_i) is randomly distributed across individuals with mean zero, variances σ_{α}^2 and σ_{γ}^2,



(a) (b)

Figure 8: Fit of the estimation for model with age-varying persistence. The orange closed circles are the autocovariance matrices generated by the restricted (left) and unrestricted (right) versions. The gray open circles are their empirical counterparts which are targeted in the GMM estimations.

and covariance $\sigma_{\alpha\gamma}$. The individual-specific growth rate, γ_i, is assumed to be independent of ε_{it} and u_{it}. Under this specification of the dynamic process, the cross-sectional variances and covariances of the residual at age t are given by

$$var(R_{it}) = \sigma_\alpha^2 + 2\sigma_{\alpha\gamma}t + \sigma_\gamma^2 t^2 + var(z_{it}) + \sigma_u^2, \tag{12}$$

$$cov(R_{it}, R_{it+k}) = \sigma_\alpha^2 + \sigma_{\alpha\gamma}(2t+k) + \sigma_\gamma^2 t(t+k) + \rho^k var(z_{it}), \tag{13}$$

where

$$var(z_{it}) = \rho^2 var(z_{it-1}) + \sigma_{\varepsilon,t}^2 = \delta_{\varepsilon,0}\sum_{j=0}^{t-1}\rho^{2j} + \delta_{\varepsilon,1}\sum_{j=0}^{t-1}\rho^{2j}(t-j).$$

Recall that, under the baseline model with $\delta_{\varepsilon,1} = 0$, the variance of the residuals is only convex in age when ρ is larger than 1. This is not the case under the heterogeneous profile model. As equation (12) shows, when $\sigma_\gamma^2 > 0$, the variance of the residuals can be convex in age even if $\rho > 1$ and $\delta_{\varepsilon,1} = 0$. In addition, unlike the restricted version of the baseline model, the model with heterogeneous profiles does not require $\rho < 1$ to generate covariances that decline with age length (at least a early ages). To see this observe that

$$cov(R_{it}, R_{it+k+1}) - cov(R_{it}, R_{it+k}) = \sigma_{\alpha\gamma} + \sigma_\gamma^2 t + (\rho^{k+1} - \rho^k)var(z_{it}), \tag{14}$$

and note that even if $\rho > 1$, making the third term positive, the differential can be negative if $\sigma_{\alpha\gamma}$ is negative and sufficiently large in magnitude.

ρ	$\sigma_\alpha^{2\,a}$	$\sigma_\gamma^{2\,a}$	$\sigma_{\alpha\gamma}{}^{a}$	σ_u^{2a}	$\delta_{\varepsilon,0}{}^{a}$	$\delta_{\varepsilon,1}{}^{a}$
A. Restricted						
1.004	8.776	0.033	-0.541	5.148	4.413	–
(0.007)	(1.758)	(0.044)	(0.151)	(0.305)	(0.218)	–
B. Unrestricted						
0.931	12.434	0.046	0.753	4.206	0.00	0.407
(0.012)	(1.698)	(0.007)	(0.095)	(0.368)	(0.349)	(0.023)

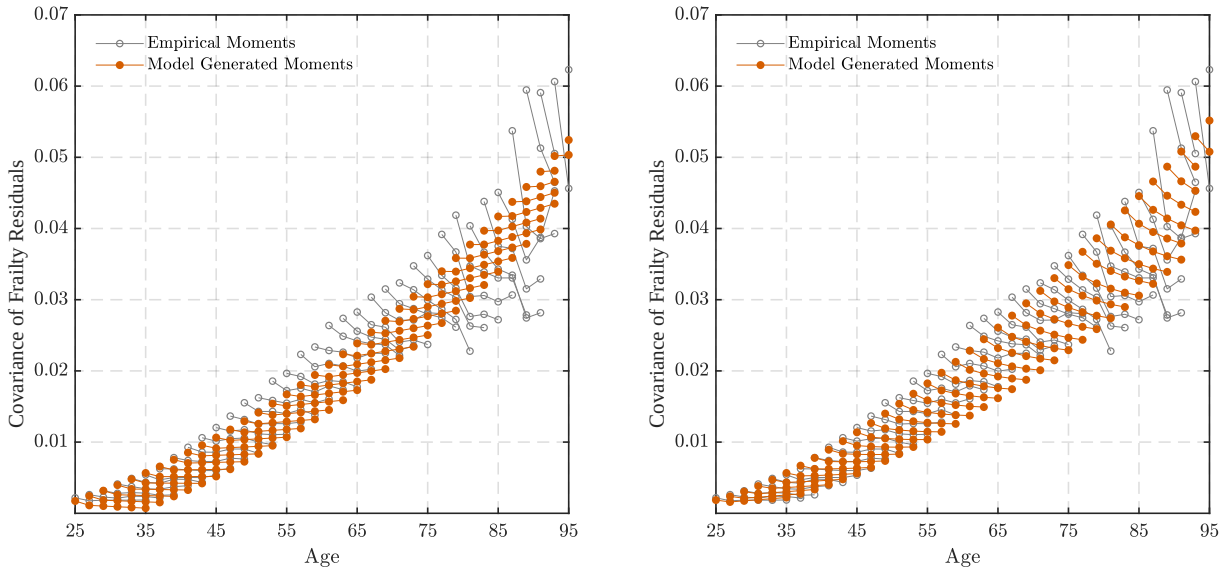
Table 11: Results from estimating the restricted ($\delta_{\varepsilon,1} = 0$) and unrestricted versions of the heterogeneous profiles model using the PSID sample. Standard errors are in parenthesis. The estimation targets the variance moments in Figure 3. [a]Estimates and standard errors are reported in tens of thousands.

We estimate two versions of the heterogeneous profiles model using the same procedure and targeted empirical moments as for the baseline model. Under the *restricted* version, we allow for heterogeneous frailty growth rates but remove the age-variation in the conditional variance of the AR(1) shock by setting $\delta_{\varepsilon,1} = 0$. Under the *unrestricted* version, we put no *a priori* restrictions on the parameters.

Table 7 presents the results from the GMM estimation. The model predicted variance-covariance matrices and their empirical counterpart are presented in Figure 9. Adding heterogeneous frailty growth rates to either the restricted or unrestricted versions of the baseline model does little in terms of improving the model's ability to match the data. This can be seen by comparing the parameter estimates in Table 11 to those in Table 7 or by comparing Figure 9 to Figure 5.

First, compare across the two restricted versions of the baseline and heterogeneous profiles models. The addition of heterogeneous profiles has little impact on the estimated values of the other model parameters or the model predicted variance-covariance matrix. Finally, notice that the estimated value of σ_γ^2 is not statistically different from zero. The primary reason that allowing for heterogeneous profiles does not significantly improve the model is that the profiles generate a covariance pattern that is opposite to the one in the data. Recall that the empirical covariance profiles are declining and that the rate of decline is increasing with age. In contrast, in the heterogeneous profiles model with $\delta_{\varepsilon,1} = 0$ the covariance profiles either increase with age or, if they decline, become less steep with age. To see this note that the first term in equation 14 is independent of age. If $\rho \geq 1$ then both the second and third terms increase with age. If $\rho < 1$, then the third term decreases with age. However, the third term decreases with age at a decreasing rate, while the second term increases at a constant rate. Hence, the covariance profiles become less steep with age.

Now, compare across the two unrestricted versions. Relative to the restricted baseline model, the addition of heterogeneous profiles to the unrestricted version has a bigger impact on the estimated parameter values. However, the effect on the model's fit of the empirical variance-covariance moments is still very small. Consistent with the intuition above, the estimated value of ρ declines because the convexity of the variance profile is now in part due to the heterogeneous frailty growth rates. The heterogeneous profiles model also generates a much higher estimated value of $\delta_{\varepsilon,1}$. Comparing across Figure 9 and 5 one can see a slight



(a) (b)

Figure 9: Fit of the estimation for heterogeneous profile model. The orange closed circles are the autocovariance matrices generated by the restricted (left) and unrestricted (right) versions. The gray open circles are their empirical counterparts which are targeted in the GMM estimations.

improvement in model fit. Essentially, the additional degrees of freedom provided under the heterogeneous profiles model allows for a slightly better fit of the shape of the covariance profiles at later ages without compromising the fit of the variance profile.

4 Conclusion

The most commonly used measure of health status in economics is self-reported health status. In this paper, we propose an alternative measure: the frailty index. We show that the frailty index is easy to construct and has several advantages over self-reported health status, especially when studying health dynamics over the life cycle. First, the frailty index is a more objective measure of health than self-reported health status and is more consistent across datasets. Second, self-reported health status underestimates the average rate of deterioration of objective health (as measured by the frailty index) with age. Third, the frailty index measures health on a finer scale than self-reported health status. Exploiting the richer variation in frailty compared to self-reported health status, we show that the dispersion in frailty increases with age and the frailty distribution is significantly right-skewed. We also document substantial variation in frailty in the 'poor' self-reported health status category. This finding suggests that 'poor' self-reported health status is a weak indicator that an individual is in the top quintile of the frailty distribution. Fourth, we demonstrate that the frailty index has an edge over self-reported health status in terms of predicting major health related outcomes.

Using the frailty index, we document facts about the dynamics of the cross-sectional health distribution over the lifecycle. We show that the variance of frailty increases with age (at an increasing rate). In addition, we show that the autocovariances of frailty at each age decline with lags and the rate of decline increases with age. We then turn to proposing and estimating a stochastic process for frailty that is consistent with these facts. In doing so we restrict attention to a model in which the stochastic component of frailty consists of a transitory shock and an AR(1) shock. This modeling choice is motivated by the fact that these models can be easily incorporated into structural analysis. We show that, within this class of stochastic processes, only models that allow the variance of shocks to vary with age can match the qualitative features of the empirical variance-covariance profile.

We have shown that the frailty index is a useful way to measure health when studying health dynamics. In addition, we have documented several facts about the evolution of frailty over the lifecycle. In related work, we use these findings to study the impact of health inequality and its evolution on earnings inequality and labor supply over the life cycle.

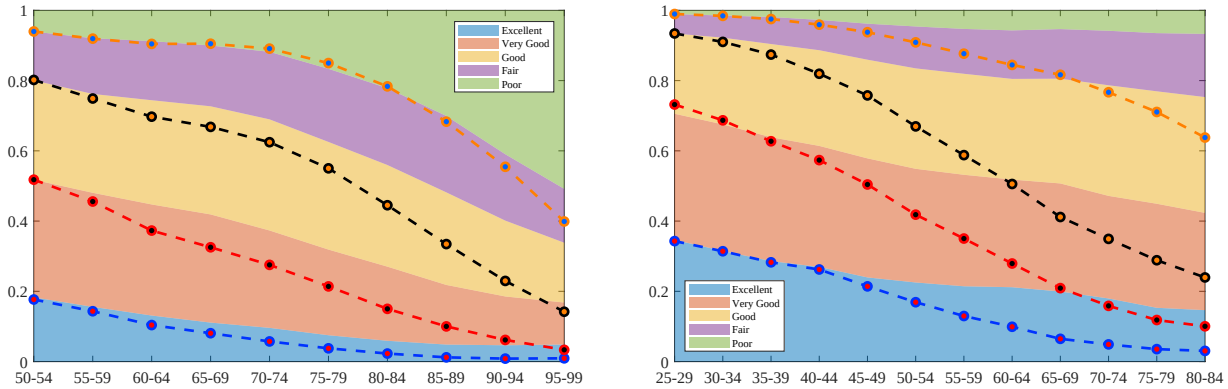
Appendix

A Frailty v. Self-reported Health Status in HRS and MEPS

In Section 2, we compare SRHS with the frailty index using PSID data. In this section we repeat this comparison using HRS and MEPS data and show that all the patterns we report for the PSID are also found in the HRS and the MEPS.

Figure 10a shows how the distributions of SRHS and frailty evolve as individuals age according to the HRS data. Since the HRS is a survey of Americans over age 50 and their spouses, the earliest age group we include is ages 50-54. However, because the HRS oversamples older Americans, we are able to include an additional age group relative to PSID: the 95-99 age group. For each age group, the height of each shaded area is the fraction of individuals in the corresponding SRHS category. We construct partitions of the HRS frailty distribution following the same procedure as is used for PSID. This procedure is described in Section 2. When constructing the partitions in the HRS, the frailty cutoffs are calculated so that the partitions of the frailty distribution match the distribution of SRHS for age group 50-54 (instead of age group 25-29). As in the PSID, the distribution of frailty in the HRS evolves more rapidly with age than the distribution of SRHS. For example, 17% of 95-99 year-olds assess their health to be 'excellent' or 'very good'. However, only 3% of 95-99 year-olds have a frailty index value below the cutoff for 'very good' frailty determined using the 50-54 year-old age group. Figure 10b shows similar patterns using MEPS respondents ages 25 to 84. Note that we can only observe these distributions for 5 year age groups up to age 84 because MEPS topcodes age at 85.

It is worth noting that the evolution of the frailty distribution in the MEPS and the HRS is remarkably similar to that in the PSID. To see this compare Figure 1 with Figures 10a and 10b. Note that the age ranges plotted differ across the three figures. In contrast, in the MEPS and the HRS, the fraction of people in different SRHS categories evolves very



(a) HRS. (b) MEPS

Figure 10: Distribution of health status by age. The colored areas show the fraction of individuals by SRHS at each age. The dashed line shows the fraction of health status according to frailty index. Panel (a) is based on calculation in the HRS. Panel (b) is based on calculations in MEPS.

differently with age as compared to PSID. For instance, in the MEPS these fractions are less variable with age as compared to the PSID and the HRS has a larger fraction of individuals at each age who report that they are in 'poor' health as compared to PSID. This is another indication that the frailty index is a more consistent indicator of health status than SRHS.

Finally Table 12 shows transition probabilities between different SRHS categories and frailty categories in the HRS and Table 13 shows the same transition probabilities in the MEPS. As in the PSID, the frailty index is more persistent than SRHS and the difference in persistence is the largest at the poorer health end of the spectrum in both the HRS and the MEPS.

B Computation of Cohort-adjusted Empirical Moments

This section describes how we obtain the cohort-adjusted variance and covariance moments that we target in the GMM estimations of the frailty process described in Section 3. The cohort-adjusted moments for the main PSID sample are shown in Figure 3. The corresponding raw moments are in Figure 11.

First, cohorts are defined. Cohort 1 consists of all individuals born before 1911 (this is 35 individuals with birth years from 1905–1910). Cohorts then increment with every two birth years, so that the second cohort includes birth years 1911 and 1912, and the third 1913 and 1914. Denote the number of cohorts by n_c. Next, age groups are defined. Each age group spans 2 years and age groups are non-overlapping. Denote an individual's age group by age_f. The initial OLS regression is then run on the sample of individuals meeting the desired age restrictions for the particular covariance matrix being computed. The means of the squared residuals R^2_{it} for each age group are the raw variance profile. Denote the raw variance of the 35–36 year old age group by R^2_{35}.

To obtain the cohort-adjusted variances and covariances, we first define the individual-

Transition Probabilities (%)					
Self Reported Health Status					
	'excellent'	'very good'	'good'	'fair'	'poor'
'excellent'	50.3	35.1	9.9	2.5	2.2
'very good'	11.3	54.6	25.6	5.3	3.2
'good'	2.8	21.7	50.2	17.8	7.6
'fair'	1.1	6.3	24.7	45.3	22.6
'poor'	0.4	1.9	7.5	25.5	64.7
Health Status by Frailty Index					
	'excellent'	'very good'	'good'	'fair'	'poor'
'excellent'	60.0	34.0	4.9	0.8	0.2
'very good'	11.1	57.3	27.9	3.3	0.5
'good'	0.9	16.3	58.5	21.9	2.3
'fair'	0.1	1.5	19.2	60.6	18.6
'poor'	0.0	0.2	1.4	20.0	78.5

Table 12: Transition probabilities for health status in the HRS. Left panel: Self Reported Health Status. Right panel: health status by frailty index. Source: authors' calculation using HRS data.

specific moments

$$m_i^{t,t+k} = R_{it}R_{it+k},$$

for $k \in \{0, 1, 2, \ldots, 5\}$. We regress these moments on age and cohort dummies as follows:

$$m_i^{t,t+k} = \eta + \sum_{a=0}^{35}\sum_{j=0}^{5} \delta_j^t I[k=j]I[a=25+2t] + \sum_{c=1}^{n_c} \theta_c I[C_i=c] + \varepsilon_{ik}^t.$$

Note that the omitted age effect dummy is δ_0^{35}. Individual-specific cohort-adjusted moments are then given by

$$\tilde{m}_i^{t,t+k} = \hat{\delta}_k^t + \varepsilon_{ik}^t + R_{35}^2 - \tilde{m}_0^{35},$$

where $\tilde{m}_{35}$ is the mean variance moment for individuals in the 35–36 year-old age group. The last two terms scale the moments so that the raw and cohort-adjusted variances are the same for this age group. The cohort-adjusted variance/covariance profiles are the means of these individual-specific cohort-adjusted moments for each age group:

$$\tilde{m}_k^t = \sum_i \tilde{m}_i^{t,t+k}.$$

C Estimation using Alternative Empirical Moments

In this section we show that our baseline estimation results are robust to concerns about selectivity bias due to mortality. To this end we run our estimations using two alternative

Transition Probabilities (%)					
Self Reported Health Status					
	'excellent'	'very good'	'good'	'fair'	'poor'
'excellent'	52.6	32.2	12.9	1.8	0.5
'very good'	19.2	50.5	25.6	3.9	0.8
'good'	8.8	30.0	47.7	11.5	2.1
'fair'	4.1	13.4	37.3	35.9	9.3
'poor'	1.5	5.3	18.8	34.5	39.8
Health Status by Frailty Index					
	'excellent'	'very good'	'good'	'fair'	'poor'
'excellent'	66.6	27.9	3.9	1.3	0.4
'very good'	15.0	61.9	18.4	3.8	0.9
'good'	9.3	17.8	52.5	17.8	2.7
'fair'	0.6	2.3	8.4	75.0	13.7
'poor'	0.4	1.3	3.2	23.8	71.3

Table 13: Transition probabilities for health status in the MEPS. Left panel: Self Reported Health Status. Right panel: health status by frailty index. Source: authors' calculation using MEPS data.

set of moments as target. In one version we estimate the model to match the moments in a subsample that excludes individuals who exit the baseline sample due to death. In another version we estimate the model using subsample that is restricted to individuals aged 25 to 65.

C.1 Estimation using the subsample of survivors only

Table 14 presents the results from estimating the restricted and unrestricted versions of the model by targeting the empirical variance-covariance matrix plotted in Figure 4. This matrix is constructed using a sample that is equivalent to the baseline sample except that it excludes individuals who exit the baseline sample due to death. The parameter estimates are very similar to those of the baseline estimation.

C.2 Estimation using subsample of working age individuals.

Table 15 presents results from estimating the baseline model using the variances and auto-covariances only of working-age individuals (those ages 25 to 65) as target moments. The results are also very similar to our baseline estimation results. The estimated values of ρ are slightly above their counterparts in the estimation where we target the entire empirical variance-covariance matrix (Table 15). This is likely due to the fact that in the data the age profile of variance is slightly more convex between ages 25 and 65 than for the entire age distribution.

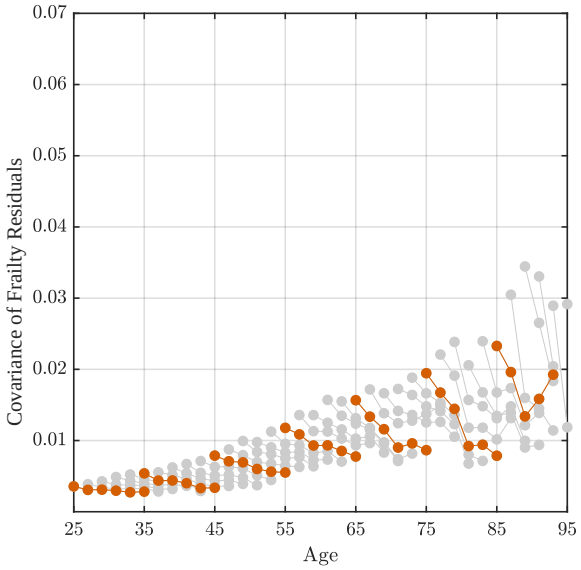


Figure 11: Raw covariances of the residuals, R_{it}, by age in PSID.

ρ	$\sigma_\alpha^{2\,\mathrm{a}}$	$\sigma_u^{2\,\mathrm{a}}$	$\delta_{\varepsilon,0}{}^{\mathrm{a}}$	$\delta_{\varepsilon,1}{}^{\mathrm{a}}$
A. Restricted				
1.009	4.373	5.562	3.921	–
(0.001)	(1.298)	(0.280)	(0.127)	–
B. Unrestricted				
0.993	14.227	4.284	2.114	0.250
(0.001)	(1.412)	(0.293)	(0.193)	(0.015)

Table 14: Results from estimating the restricted and unrestricted versions of the baseline model using the survivors-only PSID sample. Standard errors are reported in parenthesis. The estimation targets the variance and covariance moments in Figure 4. [a]Estimates and standard errors are reported in tens of thousands.

D Additional Subsample Estimations

Tables 16 and 17 report the estimation results by education for each gender separately. The pattern in these tables are mostly similar to the ones observed in Section 3.2. The estimates reflect larger and more rapidly increasing variation in frailty for women relative to men and for the high school group relative to the college group. Interestingly, under the unrestricted model, the estimated degree of persistence of the AR(1) shock is significantly larger for college-educated women relative to the other three groups.

E Estimation using Principal Component

One interpretation of the frailty index is that it is essentially a weighted average of all available health indicators with all indicators assigned an equal weight. Throughout the

ρ	$\sigma_\alpha^{2\,a}$	$\sigma_u^{2\,a}$	$\delta_{\varepsilon,0}{}^{a}$	$\delta_{\varepsilon,1}{}^{a}$
A. Restricted				
1.014	6.352	5.410	3.337	–
(0.001)	(1.351)	(0.295)	(0.153)	–
B. Unrestricted				
0.999	15.823	4.418	1.363	0.271
(0.002)	(1.490)	(0.306)	(0.222)	(0.019)

Table 15: Results from estimating the restricted and unrestricted versions of the baseline model using the PSID sample but a limited set of target moments. Standard errors are reported in parenthesis. The estimation targets the variance and covariance moments for individuals ages 25-65 in the right panel of Figure 3. [a]Estimates and standard errors are reported in tens of thousands.

paper we do not provide any justification for this simple weighting scheme other than its simplicity and usage in the gerontology literature. To what extent does our analysis depend on this particular weighting scheme? In this section we show that the results of our main estimation are robust to using an index constructed via a principal component analysis.

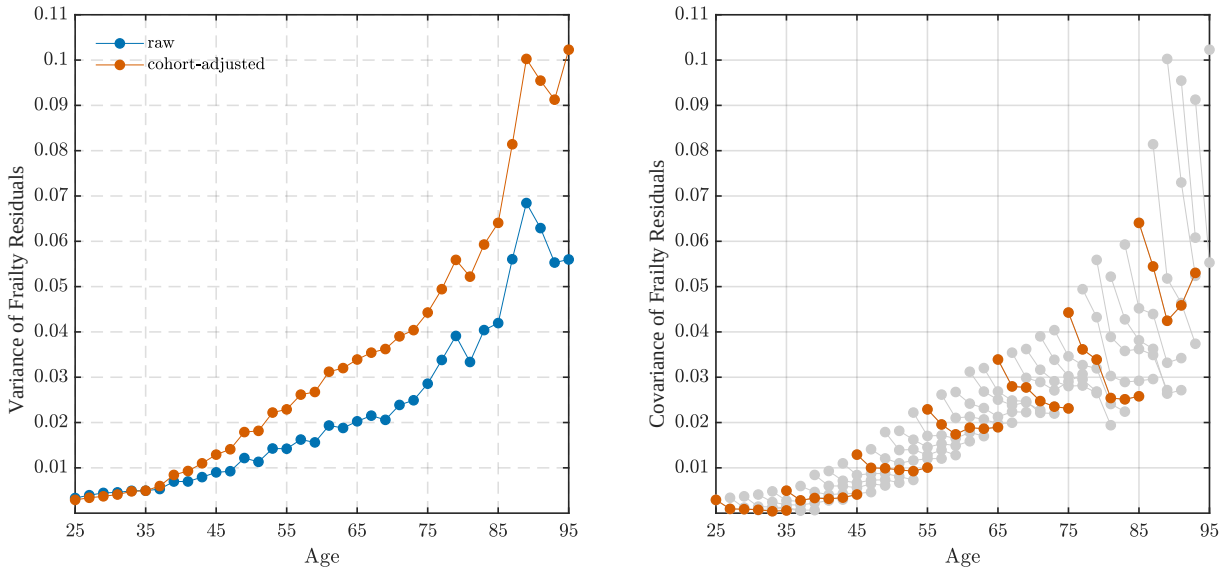
We start by obtaining the first principal component of the health indicators reported in Table 23. The first principal component accounts for as much variability in the indicators as possible. We construct a health index by taking a weighted average of the health indicators using the first principal component weights. We normalize the weights so that they sum to 1. We refer to this index as the principal component index.

The principal component loadings (weights) are reported in Table 18. Notice that the index places most of the weight on ADL and IADL variables as compared to the variables relating to specific health conditions. This is consistent with the principal component weights documented by Poterba et al. (2017) for health indicator variables in the HRS.

We now repeat the estimation of the baseline model presented in Section 3.1 with the frailty index replaced by the principal component index. Table 19 presents the results of the initial OLS regression. The estimated effects of age, years of schooling, gender, and marital status on health as measure by the principal component index are very similar to those found when health is measure by frailty.

	ρ	$\sigma_\alpha^{2\,a}$	$\sigma_u^{2\,a}$	$\delta_{\varepsilon,0}{}^{a}$	$\delta_{\varepsilon,1}{}^{a}$
A. Restricted					
All	1.009	4.815	3.973	2.818	–
	(0.001)	(1.137)	(0.305)	(0.124)	–
High school	1.010	5.241	3.781	3.510	–
	(0.001)	(1.838)	(0.480)	(0.184)	–
College	1.010	0.502	2.257	2.093	–
	(0.001)	(1.009)	(0.260)	(0.099)	–
B. Unrestricted					
All	0.993	11.243	2.988	1.519	0.191
	(0.002)	(1.256)	(0.316)	(0.192)	(0.017)
High school	0.998	10.947	2.933	2.397	0.177
	(0.002)	(2.010)	(0.496)	(0.273)	(0.026)
College	0.999	4.872	1.595	1.421	0.097
	(0.002)	(1.153)	(0.273)	(0.147)	0.013

Table 16: PSID samples: separate estimation by education, only men. The estimation targets all the variance moments in Figure 3 and the age 25-65 covariance moments. [a]Estimates and standard errors are reported in tens of thousands.



(a) (b)

Figure 12: Raw and cohort-adjusted variances (left) and cohort-adjusted covariances (right) of the residuals, R_{it}, by age using the principal component index as the measure of health and the PSID sample.

The qualitative properties of the variance-covariance matrix of the principal component index are very similar to those of the frailty index. Figure 12 presents the moments. As we did with frailty, we adjust the moments for cohort effects. The left panel of the figure

	ρ	$\sigma_{\alpha}^{2\,a}$	σ_{u}^{2a}	${\delta_{\varepsilon,0}}^{a}$	${\delta_{\varepsilon,1}}^{a}$
A. Restricted					
All	1.008	5.701	6.369	4.480	–
	(0.001)	(2.012)	(0.397)	(0.180)	–
High school	1.004	0.000	6.054	6.952	–
	(0.001)	(3.318)	(0.653)	(0.310)	–
College	1.014	3.289	4.996	2.994	–
	(0.001)	(1.353)	(0.418)	(0.145)	–
B. Unrestricted					
All	0.992	18.577	4.931	2.483	0.269
	(0.001)	(2.202)	(0.416)	(0.416)	(0.021)
High school	0.993	11.401	4.309	5.358	0.224
	(0.002)	(3.646)	(0.694)	(0.413)	(0.028)
College	1.002	6.191	4.301	2.642	0.108
	(0.002)	(1.459)	(0.433)	(0.202)	(0.019)

Table 17: PSID samples: separate estimation by education, only women. The estimation targets all the variance moments in Figure 3 and the age 25-65 covariance moments. [a]Estimates and standard errors are reported in tens of thousands.

shows both the raw and cohort-adjusted variances. The right panel of the figure shows the entire variance-covariance matrix adjusted for cohort effects. Comparing Figure 12 to Figure 3 reveals that properties of the variance profile and autocovariance profiles for the principal component index are very similar to those for the frailty index. The variance increases monotonically with age and the auto-covariances decay as the lag order increases. In addition, the auto-covariances decay faster at older ages. The biggest difference between the two set of moments is that the variance profile of the principal component index is more convex than the one for frailty. Notice also that the difference between the variance and the first autocovariance at each age is larger for the principal component index than for the frailty index. These differences will lead to slightly different parameter estimates.

The GMM estimation results for the restricted and unrestricted versions of the baseline model are presented in Table 20. The fits of the two versions of the model along with the age profiles of variances and covariances are plotted in Figure 13. Recall that in the restricted model, convexity of the variance profile is generated by a persistence that is larger than one. Since the principal component index has a more convex variance profile, the estimated value for ρ is higher (relative to the one estimated using the frailty index) under the restricted model. Moreover, the principal component index process has a much larger variance of the transitory shock and much smaller variance of the fixed effect relative to the frailty index process. This is true under both the restricted and unrestricted specifications. This is consistent with the larger difference between the variances and first autocovariances observed for the principal component index as compared to the frailty index.

Overall, the dynamics of the frailty index and the principal component index appear to be extremely similar. Both have variance-covariance profiles that strongly support a statistical process featuring a highly persistent shock with an age-varying conditional variance. The

Table 18: Principal component loadings on health variables in the PSID

Variable	Loading
Some difficulty with bathing/showering	0.285
Some difficulty with walking	0.278
Some difficulty with dressing	0.276
Some difficulty with getting outside	0.275
Some difficulty with getting in/out of bed or chair	0.268
Some difficulty with shopping for personal items	0.265
Some difficulty with preparing meals	0.258
Some difficulty with light housework	0.250
Some difficulty with using the toilet	0.245
Some difficulty with heavy housework	0.232
Some difficulty with managing money	0.190
Ever had loss of memory or mental ability	0.181
Some difficulty with using the telephone	0.180
Ever diagnosed with arthritis	0.170
Some difficulty with eating	0.166
Ever diagnosed with stroke	0.153
Ever diagnosed with heart disease	0.146
Ever diagnosed with high blood pressure	0.130
Ever had a heart attack	0.126
Ever diagnosed with other serious, chronic condition	0.122
Ever diagnosed with psychological problems	0.122
Ever diagnosed with diabetes	0.114
Ever diagnosed with lung disease	0.107
Ever diagnosed with cancer	0.086
Ever diagnosed with asthma	0.059
Has ever smoked	0.053
BMI ≥ 30	0.036

estimated parameters of the process are also very similar. Given, the simple and intuitive construction of the frailty index, we see no obvious advantage to using a more sophisticated weighting, like principal component.

F Estimation using the HRS Sample

Table 21 provides results from running the first stage OLS regression on HRS respondents aged 51–95. The regression yields similar estimates to those found for the PSID. Frailty is increasing in age and declining in years of schooling. On average, males and married individuals are less frail then females and singles. The coefficients on male and married are similar in magnitude to those from the OLS regression on the PSID data.

Figure 14 shows the empirical variances and covariances of the frailty residuals. As we did with the PSID moments, we adjust the HRS moments for cohort effects. The left panel

Variable	Coefficient (×100)	Std. Err. (×100)
Age	-0.239	(0.015)
Age^2	0.005	(0.0001)
Years of School	-0.525	(0.015)
Male	-1.113	(0.076)
Married	-2.948	(0.084)
Const.	13.236	(0.418)
Year dummies included		
	$N = 81,815$, $R^2 = 0.179$	

Table 19: OLS regression results for the health index constructed using the principal component of health deficits as weights and the PSID sample for ages 25–95.

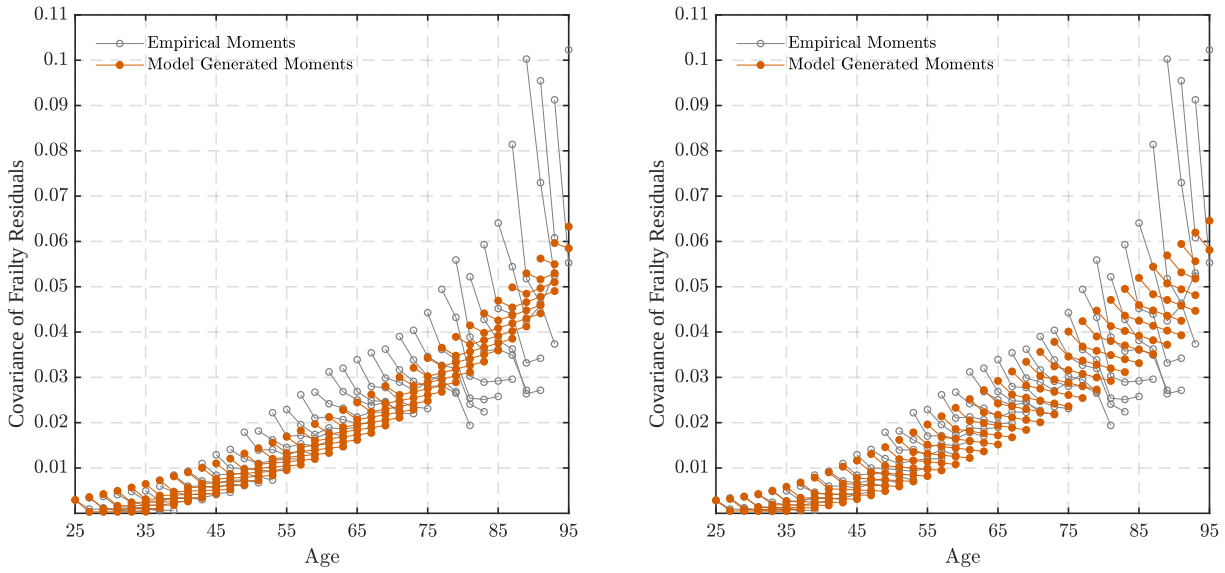
ρ	$\sigma_\alpha^{2\,a}$	$\sigma_u^{2\,a}$	${\delta_{\varepsilon,0}}^a$	${\delta_{\varepsilon,1}}^a$
A. Restricted				
1.013	0.000	26.44	3.050	–
(0.001)	(1.130)	(1.107)	(0.138)	–
B. Unrestricted				
0.987	3.985	23.25	0.758	0.393
(0.002)	(1.229)	(1.13)	(0.288)	(0.029)

Table 20: Estimation results for the health index using the principal component of health deficits as weights. The variance-covariance moments are shown in Figure 13 . [a]Estimates and standard errors are reported in tens of thousands.

of the figure shows both the raw and cohort-adjusted variances. The right panel of the figure shows the entire variance-covariance matrix adjusted for cohort effects. We rescale the variances and covariances after removing the cohort effects so that the adjusted variance at age 58 is the same as the raw variance. Comparing Figure 14 to Figure 3 reveals that both the variances and the auto-covariances show patterns similar to those in the PSID data. The variance increases monotonically with age from ages 65 to 95,[26] and the auto-covariances also decay as the lag order increases. In addition, the auto-covariances decay faster at older ages.

Table 22 shows the results from estimating the restricted and unrestricted versions of the baseline model on the HRS data. The model predicted variance-covariance matrices are shown in Figure 15. The unrestricted process has much lower persistence in the HRS estimation relative to the PSID one. Also the estimated variances are much higher relative to those for the PSID. This is expected. Individuals enter the HRS at a much older age on average than the PSID. The initial variance in frailty in the HRS sample is higher due to the fanning out of frailty that occurs at younger ages. The rate of growth of the conditional variance of the persistent shock is also larger in HRS relative to PSID. This is likely due to two facts. First, the covariances decline much more rapidly with lag length in the HRS. Second, in the PSID estimation most of the weight is put on the variances and covariances

[26]Note that the non-monotonicity of the variance over the 50–65 age range is a robust feature of the HRS sample that is most likely due to survey design.



(a) (b)

Figure 13: Fit of the model estimated using the health index constructed with the principal components of health deficits as weights. The orange closed circles are the autocovariance matrices generated by the restricted (left) and unrestricted (right) versions of the baseline model estimated using principal component. The gray open circles are their empirical counterparts which are targeted in the GMM estimations.

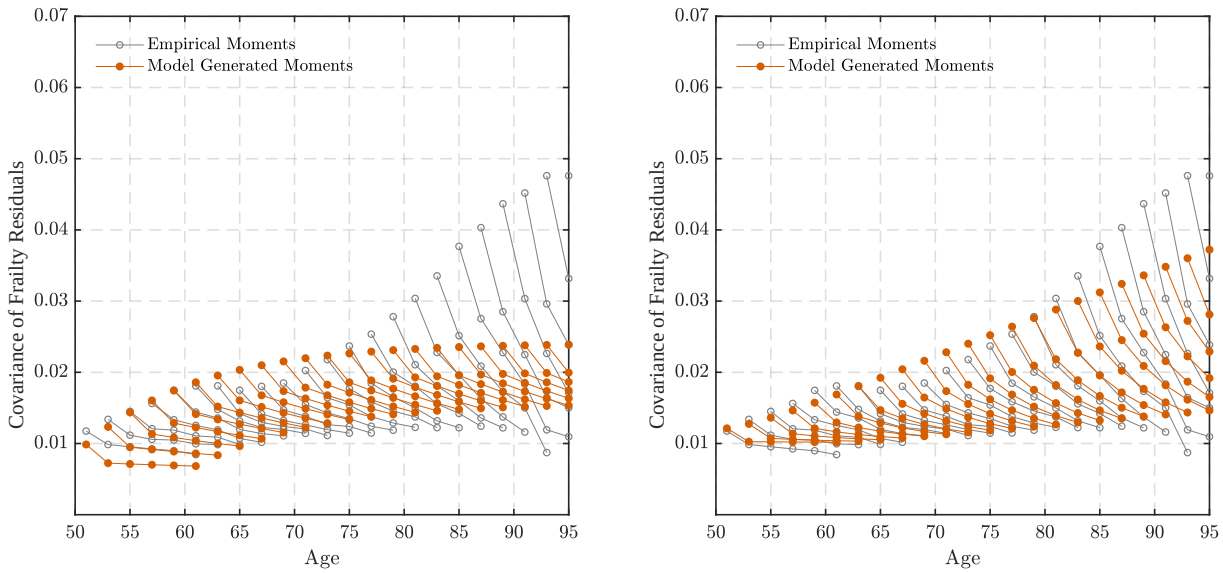
of the young because the sample sizes at older ages are much smaller. In contrast, the HRS estimation does not target any moments constructed off of individuals under age 51.

HRS ages 51–95 sample		
Variable	Coefficient (×100)	Std. Err. (×100)
Age	-1.256	(0.038)
Age^2	0.012	(0.0001)
Years of School	-1.135	(0.009)
Male	-1.585	(0.064)
Married	-3.254	(0.068)
Const.	66.570	(1.302)
Year dummies included		
	$N = 202,914,\ R^2 = 0.185$	

Table 21: OLS regression results for frailty using HRS sample: ages 51–95.

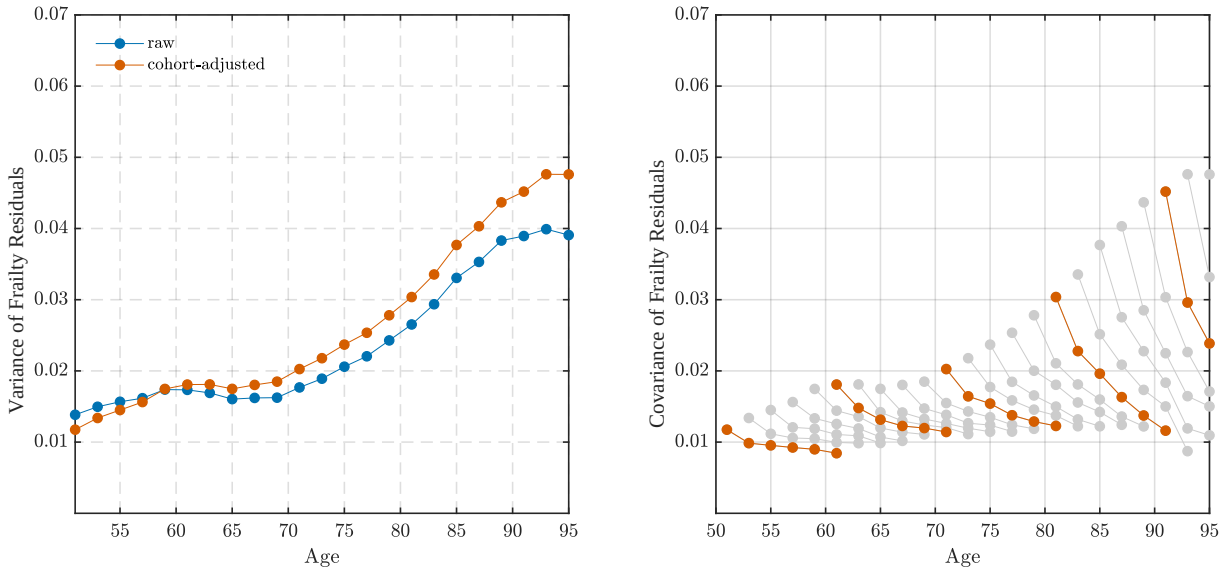
ρ	$\sigma_\alpha^{2\,\text{a}}$	$\sigma_u^{2\,\text{a}}$	${\delta_{\varepsilon,0}}^{\text{a}}$	${\delta_{\varepsilon,1}}^{\text{a}}$
A. Restricted				
0.954	59.50	24.98	14.18	–
(0.002)	(2.270)	(0.482)	(0.390)	–
B. Unrestricted				
0.865	101.48	18.33	0.00	1.516
(0.004)	(1.814)	(0.619)	(0.700)	(0.044)

Table 22: Estimation results using the HRS sample and all the variance-covariance moments in Figure 14 . [a]Estimates and standard errors are reported in tens of thousands.



(a) (b)

Figure 15: Fit of the baseline estimation using HRS sample. The orange closed circles are the autocovariance matrices generated by the restricted (left) and unrestricted (right) versions of the model. The gray open circles are their empirical counterparts which are targeted in the GMM estimations.



(a) (b)

Figure 14: Raw and cohort-adjusted variances (left) and cohort-adjusted covariances (right) of the residuals, R_{it}, by age in HRS.

G Health Variables used in Frailty Indices

Table 23 lists the health variables used to construct the frailty index for PSID respondents, Table 24 lists the variables used to construct the index for HRS respondents, and Table 25 lists the variables used for MEPS respondents.

Table 23: Health Variables used to construct frailty index for PSID respondents

Variable	Value
Some difficulty with ADL/IADLs:	
Eating	Yes=1, No=0
Dressing	Yes=1, No=0
Getting in/out of bed or chair	Yes=1, No=0
Using the toilet	Yes=1, No=0
Bathing/showering	Yes=1, No=0
Walking	Yes=1, No=0
Using the telephone	Yes=1, No=0
Managing money	Yes=1, No=0
Shopping for personal items	Yes=1, No=0
Preparing meals	Yes=1, No=0
Heavy housework	Yes=1, No=0
Light housework	Yes=1, No=0
Getting outside	Yes=1, No=0
Ever had one of following conditions:	
High Blood Pressure	Yes=1, No=0
Diabetes	Yes=1, No=0
Cancer	Yes=1, No=0
Lung disease	Yes=1, No=0
Heart disease	Yes=1, No=0
Heart attack	Yes=1, No=0
Stroke	Yes=1, No=0
Arthritis	Yes=1, No=0
Asthma	Yes=1, No=0
Loss of memory or mental ability	Yes=1, No=0
Psychological problems	Yes=1, No=0
Other serious, chronic condition	Yes=1, No=0
BMI ≥ 30	Yes=1, No=0
Has ever smoked	Yes=1, No=0

Notes: for "Ever had one of following conditions", we make the following adjustment to the raw data: if in any wave an individual has a positive answer to any the conditions below, we assign the value of 1 to that conditions when calculating frailty in all future waves.

Table 24: Health Variables used to construct frailty index for HRS respondents

Variable	Value
Some difficulty with ADL/IADLs:	
Eating	Yes=1, No=0
Dressing	Yes=1, No=0
Getting in/out of bed	Yes=1, No=0
Using the toilet	Yes=1, No=0
Bathing/shower	Yes=1, No=0
Walking across room	Yes=1, No=0
Walking several blocks	Yes=1, No=0
Using the telephone	Yes=1, No=0
Managing money	Yes=1, No=0
Shopping for groceries	Yes=1, No=0
Preparing meals	Yes=1, No=0
Getting up from chair	Yes=1, No=0
Stooping/kneeling/crouching	Yes=1, No=0
Lift/carry 10 lbs	Yes=1, No=0
Using a map	Yes=1, No=0
Taking medications	Yes=1, No=0
Climbing 1 flight of stairs	Yes=1, No=0
Picking up a dime	Yes=1, No=0
Reaching/ extending arms up	Yes=1, No=0
Pushing/pulling large objects	Yes=1, No=0
Cognitive Impairment:	
Immediate Word Recall	+.1 for each word not recalled (10 total)*
Delayed Word Recall	+.1 for each word not recalled (10 total)*
Serial 7 Test	+.2 for each incorrect subtraction (5 total)
Backwards Counting	Failed=1, 2nd attempt=.5, 1st attempt=0
Identifying obejcts & Pres/VP	.25 for each incorrect answer (4 total)
Identifying date	.25 for each incorrect answer (4 total)
Ever had one of following conditions:	
High Blood Pressure	Yes=1, No=0
Diabetes	Yes=1, No=0
Cancer	Yes=1, No=0
Lung disease	Yes=1, No=0
Heart disease	Yes=1, No=0
Stroke	Yes=1, No=0
Psychological problems	Yes=1, No=0
Arthritis	Yes=1, No=0
BMI ≥ 30	Yes=1, No=0
Has ever smoked	Yes=1, No=0

*For the 1994 HRS cohort, 40 questions were asked (instead of 20) for word recall. In this year, each missed question receives weight 0.05.

Table 25: Health Variables used to construct frailty index for MEPS respondents

Variable	Value
Need help with ADLs	Yes=1, No=0
Need help with IADLs	Yes=1, No=0
Use assistive technology	Yes=1, No=0
Limitation impacts work/housework/school	Yes=1, No=0
Any difficultly with the following:	
Walking three blocks	Yes=1, No=0
Standing for 20 minutes	Yes=1, No=0
Bending/Stooping	Yes=1, No=0
Lifting 10 pounds	Yes=1, No=0
Walking up 10 steps	Yes=1, No=0
Using fingers to grasp	Yes=1, No=0
Reaching over head	Yes=1, No=0
Ever been diagnosed with:	
High Blood Pressure	Yes=1, No=0
Diabetes	Yes=1, No=0
Cancer	Yes=1, No=0
Emphysema	Yes=1, No=0
Angina	Yes=1, No=0
Coronary Heart Disease	Yes=1, No=0
Other Heart Disease	Yes=1, No=0
Heart Attack	Yes=1, No=0
Stroke	Yes=1, No=0
Asthma	Yes=1, No=0
Arthritis	Yes=1, No=0
High Cholesterol	Yes=1, No=0
Other serious, chronic condition	Yes=1, No=0
BMI ≥ 30	Yes=1, No=0
Cognitive Limitations	Yes=1, No=0
K6 Depression Score	0–25, rescaled to 0–1

References

Amengual, D., J. Bueren and J. A. Crego, "Endogenous Health Groups and Heterogeneous Dynamics of the Elderly," Working Paper, June 2017. 1, 6

Baker, M. and G. Solon, "Earnings dynamics and inequality among Canadian men, 1976–1992: Evidence from longitudinal income tax records," *Journal of Labor Economics* 21 (2003), 289–321. 1

Blundell, R., J. Britton, D. MC and E. French, "The impact of health on labor supply near retirement," mimeo, 2017. 1

Blundell, R., M. Graber and M. Mogstad, "Labor income dynamics and the insurance from taxes, transfers, and the family," *Journal of Public Economics* 127 (2015), 58–73. 1

Bound, J., "Self-reported versus objective measures of health in retirement models," *Journal of Human Resources* 26 (1991), 106–138. 5

Capatina, E., "Life-cycle effects of health risk," *Journal of Monetary Economics* 74 (2015), 67–88. 1

Chamberlain, G., "Panel Data," in Z. Griliches and M. Intriligator, eds., *Handbook of Econometrics* (Elsevier, 1984), 1247–1318. 3.1

Cole, H. L., S. Kim and D. Krueger, "Analyzing the effects of insuring health risks: On the trade-off between short-run insurance benefits versus long-run incentive costs," *The Review of Economic Studies* 86 (2019), 1123–1169. 1

Dalgaard, C.-J. and H. Strulik, "Optimal Aging and Death: Understanding the Preston Curve," *Journal of the European Economic Association* 12 (2014), 672–701. 3

De Nardi, M., G. Fella and G. P. Pardo, "Nonlinear household earnings dynamics, self-insurance, and welfare," *Journal of the European Economic Association, forthcoming*, 2018. 1

De Nardi, M., E. French and J. B. Jones, "Why do the Elderly Save? The Role of Medical Expenses," *Journal of Political Economy* 118 (2010), 37–75. 1

De Nardi, M., S. Pashchenko and P. Porapakkarm, "The Lifetime Costs of Bad Health," Working Paper, 2017. 1

Deaton, A. and C. Paxson, "Intertemporal Choice and Inequality," *Journal of Political Economy* 102 (1994), 437–467. 21

Fella, G., G. Gallipoli and J. Pan, "Markov-chain approximations for life-cycle models," Working Paper, School of Economics and Finance, Queen Mary University of London, 2017. 1

French, E. and J. B. Jones, "On the Distribution and Dynamics of Health Care Costs," *Journal of Applied Econometrics* 19 (2004), 705–721. 8

———, "The Effects of Health Insurance and Self-Insurance on Retirement Behavior," *Econometrica* 79 (2011), 693–732. 1

Gilleskie, D., E. Han and E. Norton, "Disentangling the contemporaneous and dynamic effects of human and health capital on wages over the life cycle," *Review of Economic Dynamics* 25 (2017), 350–383. 1

Goggins, W. B., J. Woo, A. Sham and S. C. Ho, "Frailty index as a measure of biological age in a Chinese population," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 60 (2005), 1046–1051. 1, 2, 2, 9

Guner, N., Y. Kulikova and J. Llull, "Marriage and Health: Selection, Protection, and Assortative Mating," working paper, 2017. 3.1

Guvenen, F., "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics* 12 (2009), 58–79. 1, 8, 3.1, 21, 23, 3.3.2

Guvenen, F., F. Karahan, S. Ozkan and J. Song, "What do data on millions of US workers reveal about life-cycle earnings risk?," National Bureau of Economic Research, 2015. 1

Hubbard, R. G., J. Skinner and S. P. Zeldes, "Precautionary Savings and Social Insurance," *Journal of Political Economy* 103 (1995), 360–399. 8

Idler, E. L. and Y. Benyamini, "Self-rated health and mortality: a review of twenty-seven community studies," *Journal of health and social behavior* (1997), 21–37. 4

Jung, J. and C. Tran, "Medical consumption over the life-cycle," *Empirical Economics* 47 (2014), 927–957. 8

Karahan, F. and S. Ozkan, "On the persistence of income shocks over the life cycle: Evidence, theory, and implications," *Review of Economic Dynamics* 16 (2013), 452–476. 1, 3.1

Kopecky, K. and T. Koreshkova, "The Impact of Medical and Nursing Home Expenses on Savings," *American Economic Journal: Macroeconomics* 6 (2014), 29–72. 1

Kulminski, A., A. Yashin, K. Arbeev, I. Akushevich, S. Ukraintseva, K. Land and K. Manton, "Cumulative index of health disorders as an indicator of aging-associated processes in the elderly: results from analyses of the National Long Term Care Survey," *Mechanisms of ageing and development* 128 (2007a), 250–258. 2, 9, 10

Kulminski, A. M., S. V. Ukraintseva, I. V. Akushevich, K. G. Arbeev and A. I. Yashin, "Cumulative index of health deficiencies as a characteristic of long life," *Journal of the American Geriatrics Society* 55 (2007b), 935–940. 2, 9

McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior" in P. Zarembka Eds. Frontiers in Econometrics, New York Acedemic press," (1974). 18

Meghir, C. and L. Pistaferri, "Income variance dynamics and heterogeneity," *Econometrica* 72 (2004), 1–32. 1

Mitnitski, A., L. Bao and K. Rockwood, "Going from bad to worse: a stochastic model of transitions in deficit accumulation, in relation to mortality," *Mechanisms of ageing and development* 127 (2006), 490–493. 1

Mitnitski, A., X. Song, I. Skoog, G. A. Broe, J. L. Cox, E. Grunfeld and K. Rockwood, "Relative fitness and frailty of elderly men and women in developed countries and their relationship with mortality," *Journal of the American Geriatrics Society* 53 (2005), 2184–2189. 1, 2, 2, 9

Mitnitski, A. B., J. E. Graham, A. J. Mogilner and K. Rockwood, "Frailty, fitness and late-life mortality in relation to chronological and biological age," *BMC geriatrics* 2 (2002), 1. 1, 2

Mitnitski, A. B., A. J. Mogilner and K. Rockwood, "Accumulation of deficits as a proxy measure of aging," *The Scientific World Journal* 1 (2001), 323–336. 1, 2, 2, 9

Mitnitski, A. B., X. Song and K. Rockwood, "The estimation of relative fitness and frailty in community-dwelling older adults using self-report data," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 59 (2004), M627–M632. 2

O'Donnell, O., E. Van Doorslaer and T. Van Ourti, "Health and inequality," in *Handbook of Income Distribution*volume 2 (Elsevier, 2015), 1419–1533. 1

Ozkan, S., "Preventive vs. Curative Medicine: A Macroeconomic Analysis of Health Care over the Life Cycle," Working Paper, University of Toronto, 2017. 3

Poterba, J. M., S. F. Venti and D. A. Wise, "The Asset Cost of Poor Health," *The Journal of the Economics of Ageing* 9 (2017), 172–184. 1, 7, E

Rockwood, K., M. Andrew and A. Mitnitski, "A comparison of two approaches to measuring frailty in elderly people," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 62 (2007), 738–743. 2, 9

Rockwood, K. and A. Mitnitski, "Frailty in relation to the accumulation of deficits," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 62 (2007), 722–727. 2, 9

Schünemann, J., H. Strulik and T. Trimborn, "The gender gap in mortality: How much is explained by behavior?," *Journal of Health Economics* 54 (2017a), 79–90. 3

———, "Going from bad to worse: Adaptation to poor health health spending, longevity, and the value of life," *Journal of Economic Behavior & Organization* 140 (2017b), 130–146. 3

Searle, S., A. Mitnitski, E. Gahbauer, T. Gill and K. Rockwood, "A standard procedure for creating a frailty index," *BMC Geriatrics* 8 (2008), 1. 1, 2, 2.1, 9

Smith, J., "Unraveling the SES: Health connection," *Population and Development Review* 30 (2004), 108–132. 5

Storesletten, K., C. I. Telmer and A. Yaron, "Consumption and Risk Sharing over the Life Cycle," *Journal of Monetary Economics* 51 (2004), 609–633. 1, 8, 3.1, 21

Van Doorsaler, E. and U. Gerdtham, "Does Inequality in Self-Assessed Health Predict Inequality in Survival by Income? - Evidence from Swedish Data," *Social Science and Medicine* 57 (2002), 1621–1629. 4

Woo, J., W. Goggins, A. Sham and S. Ho, "Social determinants of frailty," *Gerontology* 51 (2005), 402–408. 2, 9

Wood, R. G., S. Avellar and B. Goesling, *The Effects of Marriage on Health: A Synthesis of Recent Evidence* (New York: Nova Science Publishers Inc., 2009). 20

Yang, Y. and L. C. Lee, "Dynamics and heterogeneity in the process of human frailty and aging: evidence from the US older adult population," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* (2009), gbp102. 2.1