Review of Economic Dynamics 45 (2022) 237–263

Contents lists available at ScienceDirect

Review of Economic Dynamics

www.elsevier.com/locate/red

The evolution of health over the life cycle ☆

Roozbeh Hosseini [a], Karen A. Kopecky [b,*], Kai Zhao [c]

[a] Department of Economics, Terry College of Business, University of Georgia, 620 South Lumpkin Street, Athens, GA 30602, United States of America
[b] Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St. NE, Atlanta, GA 30309, United States of America
[c] Department of Economics, University of Connecticut, 365 Fairfield Way, Storrs, CT 06269-1063, United States of America

ARTICLE INFO

Article history:
Received 7 June 2019
Received in revised form 1 July 2021
Available online 9 July 2021

JEL classification:
I10
I14
C33

Keywords:
Health
Frailty index
Self-reported health status
Mortality
Simulated method of moments
Selection bias

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. Using this index, we document four stylized facts on health dynamics over the life cycle and show that they are robust to other ways of constructing the index. We also compare the frailty index with self-reported health status and find significant differences in their dynamics and ability to predict health-related outcomes. Finally, we propose and estimate a stochastic process for frailty dynamics over the life cycle accounting for mortality bias. Our frailty measure and dynamic process can be used to study the evolution of health over the life cycle and its economic implications.

1. Introduction

One of the most challenging aspects in the study of life cycle health dynamics is the lack of consensus on what health is and how exactly it should be measured. As a result, what researchers observe are noisy and incomplete signals about health. These signals can be divided into two broad categories. One category consists of people's self-assessment of their health or the extent to which health affects their activities. The other category consists of a long array of objective health-related indicators, ranging from past medical diagnostics to current health conditions and health-induced limitations. Both categories of signals have great potential to inform us about the evolution of the latent health status of an individual over her life cycle. The challenge, however, is to tease this information out in a concise and tractable way.

In this paper we study the evolution of health over the life cycle using a measure of health called the frailty index. The index aggregates information from the wide range of health indicators in the data into a unitary measure that is easy

☆ We would like to thank Mariacristina DeNardi and four referees for helpful feedback. We also thank Jordan Herring for outstanding research assistance and Toni Braun, Jesus Bueren, Eric French, John Jones, 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 seminars 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.

* Corresponding author.
E-mail addresses: roozbeh@uga.edu (R. Hosseini), karen.kopecky@atl.frb.org (K.A. Kopecky), kai.zhao@uconn.edu (K. Zhao).

https://doi.org/10.1016/j.red.2021.07.001
1094-2025/

to construct. Using the frailty index, we document four stylized facts about life cycle health dynamics. We then show that these findings are robust to variations in how the index is constructed that are inspired by other measures of health used in the literature. In particular, we show that the variations have relatively little impact on our findings on life cycle health dynamics or on the ability of the index to predict health-related outcomes. We also compare the frailty index to a commonly used measure of health status in the economics literature: self-reported health status (SRHS). We document significant differences in health dynamics and prediction ability between the two measures. Finally, we present a statistical model of frailty dynamics over the life cycle and show how to estimate the model using a simulated method of moments procedure that allows us to account for selection bias due to mortality.

The frailty index, or frailty for short, is simply the accumulated sum of all adverse health events that an individual has incurred. It can be treated as a continuous variable and is used extensively in the gerontology literature.[1] The idea behind its construction is as follows. As individuals age, they accumulate health problems, or deficits, ranging from symptoms to clinical signs, and laboratory abnormalities to diseases and disabilities. 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 it is constructed using different lists of deficits. They also find that the frailty index is a better predictor of mortality and institutionalization than age is.

Following the guidelines described in Searle et al. (2008), we construct a frailty index for individuals in the Panel Study of Income Dynamics (PSID). Since 2003, the PSID contains a rich set of survey questions on various aspects of individuals' health conditions, allowing us to include a broad range of deficit variables in our index. The variables span three general categories: restrictions and difficulties with activities of daily living (ADLs) and instrumental activities of daily living (IADLs); mental and cognitive impairments; and medical diagnoses and measurements. Examples of deficits in the first category are difficulties eating, dressing, walking across a room, or shopping without assistance. Examples from the second category are diagnoses of memory or psychological problems. Examples from the third category are diagnoses of high blood pressure, cancer, or obesity. The index is normalized by the sum of all deficits considered such that it lies 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 focus on the PSID dataset for two reasons. First, it allows us to document facts about health dynamics over the entire adult life cycle. Second, given the long panel structure of the PSID, it is an especially useful dataset for estimating a stochastic frailty process. However, the stylized facts that we present are not unique to the PSID data. In an extensive online appendix we show that the same set of facts emerge when obtained using a similarly constructed frailty index based on data from either the Health and Retirement Study (HRS) or the Medical Expenditure Panel Survey (MEPS). Given the similarities in frailty across the datasets, we also use the HRS and MEPS data to provide additional results that cannot be obtained using PSID data. For example, we use the HRS to study the relationship between frailty and mortality, and the MEPS to study the relationship between frailty and total medical expenditures.

Using frailty as our measure of health, we document the following facts about health dynamics over the life cycle. First, mean frailty increases with age, decreases with education, and is higher for women than for men. Also, differences in mean frailty are larger by education than by gender. Second, cross-sectional dispersion in frailty rises with age both overall and when controlling for gender or education, is declining in education, and is slightly higher for women than for men. Third, the cross-sectional distribution of frailty is right skewed both overall and when conditioning on age. This skewness is driven by two features of frailty in the data. First, many individuals have a frailty value of zero, meaning they have none of the deficits considered; the fraction of such individuals gradually declines with age. Second, at all ages, there is a long thin right tail of highly frail individuals; thus, cross-sectional variation in frailty is relatively larger among unhealthy individuals. Finally, frailty is highly persistent. We show that over a two-year period, individuals are much more likely to remain in the same frailty quintile than to move to a different one. Moreover, the probability of moving to a new quintile declines rapidly as its distance from the current one increases.

An alternative measure of health that is commonly used by researchers is SRHS. In the PSID (as in many other surveys) respondents are asked if their health is 'excellent', 'very good', 'good', 'fair', or 'poor'. Frailty and SRHS differ in several ways. Frailty is constructed using objective measures of health, while SRHS is a subjective one. In addition, frailty is a cardinal measure of health that can be treated as a continuous variable, whereas SRHS is a coarse ordinal measure of health. In addition to these qualitative differences between the two measures, we document several quantitative differences. First, SRHS indicates a less rapid decline in health with age as compared to frailty. Second, there is more variation in frailty than in SRHS. Specifically, we document substantial variation in frailty within all SRHS categories even at younger ages. The variation in frailty within SRHS categories is informative about several health-related outcomes, indicating that it reflects true variation in underlying health. Third, we find that frailty is more persistent than SRHS; the conditional probability of remaining in the same SRHS category two years later is lower than the probability of remaining in comparable frailty categories. Finally, we show that frailty tends to perform better than SRHS at predicting health-related outcomes. In particular, the frailty index outperforms SRHS in predicting total medical expenses, mortality, nursing home entry, and

[1] 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.

social security disability insurance recipiency. However, we also find that both measures have independent predictive power, indicating that each measure provides information about true underlying health that is not contained in the other measure.

Frailty aggregates information from several objective health indicators into a single index. However, it is not the only way to exploit multiple health variables to construct a unitary measure of health. One alternative approach used in the literature, see Poterba et al. (2017), is to take the first principal component of the indicators as the measure of health. Another approach, see Blundell et al. (2020), is to use a subjective indicator of health instrumented by objective indicators. Both approaches can be thought of as alternative ways to construct the weights in the frailty index. For instance, instead of equally weighting each deficit variable, the first principal component of the deficits chooses the weights to maximize the proportion of the variance in the variables that is captured by the index. Similarly, using a subjective indicator of health that has been instrumented by objective indicators is equivalent to choosing the deficit weights to minimize the distance between the index and the subjective health indicator. We show that our facts on health dynamics over the life cycle are robust to these two alternative ways of constructing the frailty index. Given our finding that SRHS has independent power in predicting health-related outcomes, we also show that the facts are robust to including it directly as an additional deficit variable in the frailty index. Finally, we find that all four versions of the frailty index that we consider perform similarly well in predicting health-related outcomes.

In the final section of the paper we propose a statistical model of frailty dynamics over the life cycle that is consistent with our four stylized facts. The model allows for a mass of individuals with zero frailty that gradually declines with age. The dynamics of nonzero log frailty are governed by a deterministic component common across individuals, a fixed individual-specific component, and a stochastic component consisting of an AR(1) shock and a purely transitory shock. The process can easily be embedded into quantitative life cycle models.[2] To estimate the model we employ a simulated method of moments procedure that accounts for selection bias due to mortality. This procedure targets several empirical moments. In particular, the parameters governing the individual-specific and stochastic components are identified by the variance-covariance profile of log frailty in the PSID data. We show that after having accounted for mortality effects, the model is able to replicate well both moments targeted in the data and the overall cross-sectional distributions of frailty by age. Finally, we estimate the model separately first by gender and then by education subgroups. The results indicate little variation in frailty dynamics by gender but significant variation across education groups. For instance, there is less overall variation in frailty and the shocks to frailty are significantly less persistent among college graduates than among those without a college degree.

Our paper contributes to a large literature that studies health dynamics over the life cycle and their economic implications. Most of the analyses reported in this literature use SRHS to measure health. For instance, Cole et al. (2019) use SRHS to study the effect of labor and health insurance market policies on the evolution of the cross-sectional health distribution. Capatina (2015) uses SRHS to calibrate a structural model and quantify the impact of health on labor supply, asset accumulation, and welfare. Braun et al. (2017) assess the welfare gains from means-tested old-age social insurance programs in a framework where health dynamics (measured as transitions across SRHS categories) impact medical expenses, mortality, and the risk of losing one's spouse. See also Kitao (2014), French and Jones (2011), and De Nardi et al. (2010), among others. One innovative paper in this literature is De Nardi et al. (2018), who use the persistence of the self-reported 'bad' health status to measure the severity of health shocks. They use their estimated health process to quantitatively evaluate the lifetime consequences of bad health. Another notable paper is Ameriks et al. (2020), who use SRHS augmented to include an additional category of health characterized by the need for help with ADLs to explore the drivers of old-age saving behavior.

In the empirical literature that studies the relationship between health and labor supply it is common to augment subjective measures of health, such as SRHS, with objective health indicators.[3] In particular, several papers in this literature essentially use the approach we describe above of instrumenting subjective measures with objective ones.[4] For instance, the approach is used by Blundell et al. (2016), Bound et al. (1999), and Disney et al. (2006) to document the differential impacts of various health dynamics on retirement. Blundell et al. (2020) compare objective health measures, subjective health measures, and subjective ones instrumented with objective ones and find that conditional on using a large enough set of objective measures, they all imply relatively similar effects of health on the labor supply of older workers.

Other papers in the literature use a more restrictive set of objective health indicators to measure health. For instance, Gilleskie et al. (2017) use body mass to measure health status and study its impact on wages in a life-cycle model. Robinson (1996) and Friedberg et al. (2014) use information on ADLs and IADLs to estimate transition rates across long-term care need states at the monthly frequency. Their model has been used both by researchers (see, for instance, Lockwood (2018)) and the insurance industry to measure long-term care risk. Amengual et al. (2017) use information on ADLs and IADLs to construct an objective discrete measure of health and estimate a panel Markov switching model of old-age health dynamics.[5] By using objective indicators of health conditions, these studies avoid the disadvantages of subjective health measures. However, as argued by Blundell et al. (2020), the objective health indicators used in these studies, while perhaps well suited to study

[2] For an example, see Hosseini et al. (2021), who use the frailty process in a model to study the relationship between health inequality and lifetime earnings inequality. Another example of using frailty to measure health in a structural model is Braun et al. (2019).

[3] See Currie and Madrian (1999), French and Jones (2017), and O'Donnell et al. (2015) for general surveys of this literature.

[4] Two pioneering examples are Stern (1989) and Bound (1991).

[5] 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 see it as a disadvantage, as it is less flexible than a continuous measure like ours. One can always discretize a continuous process, but it is not so obvious how to go the other way.

particular types of health-related events, provide an incomplete view of overall health since they cover only a subset of health conditions. The frailty index, in contrast, serves as a more comprehensive summary of an individual's health status.

Other notable approaches to measuring health include Dalgaard and Strulik (2014), who, also inspired by the gerontology literature, model health evolution over the life cycle 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; by Schünemann et al. (2017b) to study the impact of deteriorating health on the value of life; and by Kopecky and Koreshkova (2014) and Ozkan (2017) to estimate health shock processes by targeting survival probabilities and medical expenditures.

Our paper is also related to a small literature that has estimated stochastic processes of health dynamics using measures of health similar to the frailty index or the variations of it we consider. Using SRHS instrumented by objective measures, Bound et al. (1999) document that health shocks are highly persistent and estimate an autoregressive ordered probit model of health. In the context of studying the relationship between health and labor supply, Blundell et al. (2016) estimate a stochastic health process similar to ours using a health measure that is also based on instrumenting subjective measures with objective ones. They estimate the process using HRS data on 50- to 66-year-olds and do not control for mortality bias. Overall, their estimated parameters are similar to ours, although their AR(1) component is slightly less persistent. Finally, Mitnitski et al. (2006) estimate a stationary discrete Markov process of frailty using two waves of the Canadian Study of Health and Aging (CSHA).

Finally, this paper is related to the literature on estimating earnings and medical expenditure processes. 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), who estimate medical expenses processes. Our approach to estimating a frailty process draws heavily from this literature, which has focused on simpler statistical models that can be easily incorporated into quantitative life-cycle models.

The rest of the paper is organized as follows. In Section 2 we present the frailty index as a measure of health and document four stylized facts on life cycle frailty dynamics. In Section 3 we show that these facts are robust to alternative ways of constructing the frailty index, and we compare frailty and SRHS. In Section 4, we present and estimate a dynamic stochastic process for frailty over the life cycle first by using the whole sample and then separately for gender and education subgroups. Section 5 concludes.

2. The frailty index

As people age, they develop an increasing number of health problems, functional impairments, and abnormalities ranging from mild (e.g., reduced vision) to serious (e.g., heart disease). Regardless of the severity of these conditions, as their total number rises, a 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) demonstrated that the health status of a person can be represented by combining accumulated deficits into an index variable called the *frailty index*. The index is constructed as the ratio of deficits the 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 the simplicity of construction of this index 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.[6] Moreover, these findings have been shown to be robust to the choice of dataset used to construct the index and the number of potential deficits considered.[7] 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 over the life cycle at roughly 3 percent per year, predicts mortality better than age does (Mitnitski et al. (2005) and Goggins et al. (2005)), and hardly anyone in the sample accumulates more than two-thirds of total deficits considered. These findings suggest that the frailty index is a good and robust proxy for health status.

Motivated by these studies, we construct frailty indices for a sample of individuals in the PSID. 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 frailty indices in Yang and Lee (2010). We include the following broad categories of variables in our calculations:

- Restrictions or difficulty in activities of daily living (ADL) and instrumental ADL (IADL) such as difficulty eating, dressing, or walking across a room. We refer to these as ADL/IADL variables.
- Medical diagnosis or measurement such as has, or had, high blood pressure, diabetes, heart disease, cancer, or high BMI and is a current or former smoker. We refer to these as medical variables.
- Mental or cognitive impairment such as loss of memory or mental ability, or diagnosis of psychological problems. We refer to these as mental health variables.

[6] See also Kulminski et al. (2007b); Rockwood and Mitnitski (2007); Rockwood et al. (2007); Woo et al. (2005).
[7] Especially, when at least 30 conditions are included. See Kulminski et al. (2007a).

Table 1
Summary statistics for frailty constructed using 2003–2017 PSID data on household heads and their spouses ages 25 and older.

	All	Men	Women	Less than HS	HS grad	College grad
Mean	0.11	0.11	0.12	0.16	0.12	0.08
ages 25-49	0.08	0.08	0.09	0.12	0.09	0.06
ages 50-74	0.15	0.14	0.15	0.20	0.15	0.11
ages 75+	0.25	0.23	0.26	0.28	0.24	0.22
Standard deviation	0.11	0.10	0.12	0.14	0.11	0.09
ages 25-49	0.09	0.08	0.09	0.11	0.09	0.06
ages 50-74	0.12	0.12	0.13	0.15	0.12	0.10
ages 75+	0.16	0.15	0.17	0.18	0.15	0.14
Min	0.00	0.00	0.00	0.00	0.00	0.00
5th pctile	0.00	0.00	0.00	0.00	0.00	0.00
50th pctile	0.07	0.07	0.07	0.11	0.07	0.07
95th pctile	0.36	0.32	0.36	0.46	0.34	0.25
Max	0.89	0.86	0.89	0.89	0.86	0.82
Fraction zero	0.15	0.14	0.16	0.07	0.13	0.23
Observations	96,759	44,396	52,363	14,253	52,832	26,893
Individuals	19,996	9,257	10,739	2,957	10,828	5,605
% of total	100	46	54	15	54	28

In the PSID, we have a total of 28 deficits that we use to construct each individual's frailty index.[8] Each observed deficit takes a value of either 0 or 1 for each individual. The frailty index is then constructed as the fraction of 1s reported for each individual at each age.

We chose not to use any subjective indicators of health status such as SRHS in the construction of the frailty indices. We refer to this version of the frailty index constructed using only objective measures as "frailty". In Section 3 we compare frailty to a version of the index that includes SRHS among the list of deficit variables. As mentioned before, the frailty index is the sum of all deficits a person has accumulated by the time of each interview. Therefore, by construction the index assigns the same weight to each of the deficit variables. This is, of course, arbitrary. To gauge the importance of the weights, we also compare frailty to alternative weighted-versions of the index including one based on principal component analysis.

2.1. Stylized facts on frailty dynamics over the life cycle

We now document four stylized facts about health dynamics over the life cycle where health is measured by the frailty index as described above. The facts are documented using a sample constructed with 2003–2017 PSID data.[9] The sample consists of household heads and their spouses who are at least 25 years of age at the time of the interview. It contains 96,759 observations of 19,996 individuals (9,257 men and 10,739 women).

Fact 1. *Mean frailty increases with age, decreases with education, and is slightly higher for women than for men.*

Fact 1 can be seen in Table 1, which provides statistics summarizing the distribution of frailty in our PSID sample. The first column shows that frailty is increasing in age. Mean frailty of individuals ages 50 to 74 is nearly twice as large as mean frailty of individuals ages 25 to 49, and mean frailty of those 75 years of age or older is more than three times as large. A similar pattern can be seen within each gender and education subgroup presented in the second through sixth columns of the table.

Comparing the second and third columns of Table 1 shows that mean frailty of women (0.12) is only slightly higher than that of men (0.11). In all three age groups presented in the table, on average, women have only slightly higher frailty than men. Differences in frailty are larger by education than by gender. Mean frailty of high school graduates (0.12) is three-quarters the mean frailty of those without a high school degree (0.16), and mean frailty of college graduates (0.08) is half that amount. Large differences in frailty by education can also be seen within the three age groups shown in the table. Notice that the rate of growth of frailty with age is increasing in education. In other words, the frailty of college graduates increases more rapidly with age than that of high school graduates, whose frailty in turn increases more rapidly than the frailty of those without a high school degree. Consequently, the differences in mean frailty by education become smaller at older ages. As we discuss in more detail in Section 4.3, this observed convergence in the level of frailty across

[8] The complete list of deficits used in the construction of the frailty indices in the PSID is reported in Table 1 of the online appendix. Occasionally, respondents who in a previous wave answered "Yes" to a question about whether they had ever been diagnosed with a particular condition subsequently answer "No". We forward correct these reporting inconsistencies. For instance, once a respondent answers "Yes" to "Have you ever been diagnosed with diabetes?", we change any future "No" answers to "Yes".

[9] In 2003, the PSID expanded its disability- and health-related questions to include questions on specific medical conditions, ADLs, and IADLs. Since we rely on these questions to construct individuals' frailty indices, we do not use data prior to 2003.

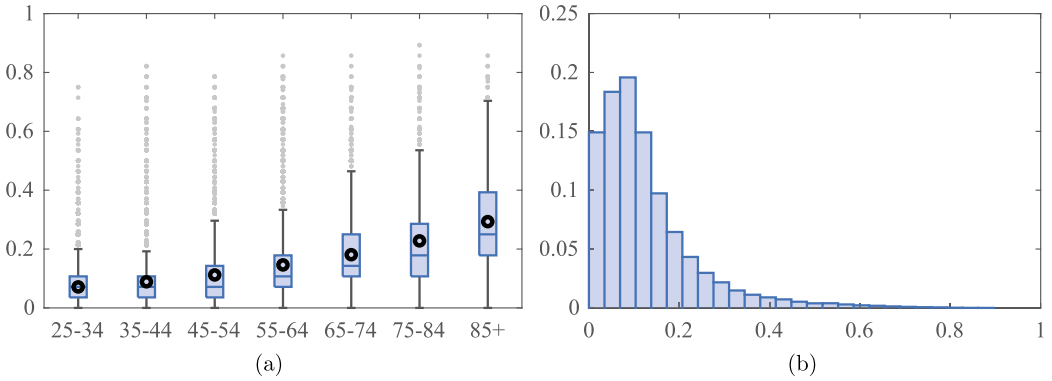


Fig. 1. Panel (a) is boxplots of frailty by 10-year age groups in the PSID sample. The bottom and top edges of the boxes indicate the 25th and 75th percentiles, respectively. The middle lines are medians, and the open circles are means. The upper whiskers extend to 1.5 times the inner quartile range. The lower whiskers extend to the minimum value of frailty in the age group. The dots are data points that lie beyond the whiskers. Panel (b) is a histogram showing the cross-sectional distribution of frailty in the sample.

education groups is due to mortality. Controlling for age, less educated individuals have, on average, higher frailty levels and, consequently, higher mortality rates.

Fact 2. *Cross-sectional dispersion in frailty increases with age.*

Fig. 1a shows box and whisker plots of frailty by 10-year age groups constructed using the PSID sample. The bottom and top edges of the box mark the 25th and 75th percentiles, respectively. The middle line in each box marks the median, and the dots indicate the means. Due to the shapes of the distributions, the upper whiskers extend to 1.5 times the inner quartile range, and the bottom whiskers extend to the minimum value of frailty for that age group. As the figure shows, the gap between the top and bottom quantiles of the distributions widens with age. Also, even though dispersion in frailty is increasing with age, there is significant variation in frailty in all age groups. The existence of variation in frailty even in the youngest age group suggests that the frailty index is capturing variation in health even among the younger individuals in the sample. In Section 3 we provide evidence that this variation is informative about individuals' likelihood of becoming beneficiaries of social security disability insurance, suggesting that it does in fact reflect true variation in their underlying health.

Fact 2 can also be seen in Table 1, which reports the standard deviations of frailty for three age groups: 25–49-year-olds, 50–74-year-olds, and individuals ages 75 and over. Notice that, consistent with the box and whisker plots in Fig. 1a, the standard deviation of frailty is increasing with age. It increases both overall and within each of the five gender and education subgroups shown in the table. Notice also that frailty dispersion is decreasing in education and is higher for women as compared to men.

Fact 3. *The cross-sectional distribution of frailty is right skewed both overall and by age.*

This fact can be seen in Fig. 1b, which presents a histogram of frailty for the entire PSID sample. The distribution is skewed for two reasons. First, there is a large mass of individuals with zero frailty. These are individuals who report having none of the 28 possible deficits considered, making them the healthiest individuals in the sample. They account for 15 percent of the frailty distribution, as indicated by the first bar in the histogram. Second, there is substantial variation in frailty among highly frail individuals (i.e., the frailty distribution has a long thin right tail). As the boxplots in Fig. 1a show, the distribution is also right skewed within each of the 10-year age groups. At younger ages much of the skewness is due to a large fraction of healthy individuals with zero frailty. This fraction is 24 percent for the 25–34-year-old age group, declining to less than 1 percent for the 85+ group.[10] However, in all the age groups, even the 25–34 year-old one, the distributions also exhibit thin right tails.

Fact 4. *Frailty is highly persistent.*

Table 2 shows two-year transition probabilities across frailty quintiles for three age groups: 25–49-year-olds, 50–74-year-olds, and individuals 75-years-old and older. Three patterns can be observed in the table. First, frailty is highly persistent. Within all three age groups, individuals are most likely to remain in the same frailty quintile, and the probability of moving

[10] See Fig. 5 in Section 4.

Table 2
Two-year transition probabilities (%) across frailty quintiles for three age groups: 25–49-year-olds, 50–74-year-olds, and 75-year-olds and older constructed using the PSID sample. Rows sum to 1.

	Bottom	2nd	3rd	4th	Top
Ages 25 to 49					
Bottom	80.3	13.8	3.9	1.6	0.4
2nd	4.1	71.4	17.6	5.6	1.3
3rd	0.7	9.2	66.2	19.1	4.7
4th	0.0	2.0	9.2	68.7	20.1
Top	0.1	0.1	0.8	8.1	90.9
Ages 50 to 74					
Bottom	79.2	16.1	3.4	0.8	0.4
2nd	5.4	67.8	20.7	5.0	1.1
3rd	0.2	7.4	67.1	22.1	3.2
4th	0.0	0.8	7.9	68.0	23.3
Top	0.1	0.2	0.8	10.0	89.0
Ages 75 and older					
Bottom	70.8	18.9	5.8	3.0	1.5
2nd	5.4	57.1	24.3	8.9	4.3
3rd	1.1	8.7	46.5	30.1	13.5
4th	0.3	2.1	15.9	50.9	30.9
Top	0.0	0.2	1.9	15.0	82.9

to another quintile is rapidly declining in its distance from the previous one. Second, the persistence of frailty declines with age. This is because as individuals age and accumulate deficits, those whose frailty shocks are most persistent become more likely to exit the sample via death. Since these individuals become less and less likely to contribute to the calculation of the transition probabilities, frailty's persistence appears to decline.[11] Third, in each age group, the persistence of frailty is highest in the bottom and top quintiles. Notice also that the probability of moving either up or down a quintile is increasing in the current quintile level, indicating that two-year changes in frailty are larger, on average, for individuals with higher current levels. This feature of the probability transition matrices indicates that the variance of frailty shocks is higher for individuals with a higher frailty level, and is consistent with the fact that both the mean and variance of frailty are increasing in age and decreasing in education, as Table 1 documents. That is, within each age group, older and less educated individuals are both more likely to be in the higher frailty quintiles and to experience larger changes in their frailty levels.

3. A comparison of the frailty index to other measures of health

In this section we explore how our findings on health dynamics over the life cycle are affected by using alternative measures of health. To this end, we first discuss differences and similarities between frailty (an objective health indicator) and SRHS (a widely used subjective health indicator). We then compare frailty to alternative versions of the index.

3.1. Self-reported health status

Even though SRHS is commonly used, there are known difficulties with it. For instance, due to its subjectivity, the extent to which it captures meaningful variation in health across individuals is questionable. What 'good' health means to one respondent might be different from what it means to another (or even to the same respondent at different moments in time).[12] A second and related difficulty is what is known as "justification bias", an issue that arises when estimating the effect of health on economic choices. For example, non-working individuals have been found to report lower levels of SRHS to justify why they are out of the labor force.[13] Subjectivity and justification bias can also be a concern for self-reports of objective health measures such as those used to construct the frailty index.[14] However, these biases should be less severe in the frailty index as compared to SRHS for two reasons. First, because the questions behind the deficit variables are much more specific in nature, there is less room for interpretation. For example, questions about medical conditions specifically ask if individuals have ever been diagnosed with the condition by a medical provider. Second, unlike SRHS, the frailty index

[11] We discuss the effect of mortality on the persistence of frailty in more detail in Section 4.3.

[12] In the medical literature, studies such as Dowd and Todd (2011) and Salomon et al. (2004) have found evidence of reporting biases by age, gender, education, and race in self-reported health assessments. These and related papers use responses to hypothetical questions on the impact of health conditions on health status to identify the reference points individuals use when assessing their health. However, the extent to which these reference points are informative is unclear. Crossley and Kennedy (2002) found that 28 percent of respondents in the 1995 Australian National Health Survey gave two different values when asked to report their SRHS twice during the same interview.

[13] See Butler et al. (1987), Benítez-Silva et al. (2004), and Blundell et al. (2020) for examples and further discussion.

[14] Baker et al. (2004) document evidence of both over- and under-reporting of self-reported objective health conditions in the Canadian National Population Health Survey.

combines information from several deficit variables, mitigating the overall impact of biases in responses related to specific variables.

A third difficulty with SRHS is that due to its lack of cardinality, it is cumbersome to determine the severity of health changes. For example, when a respondent reports 'excellent' health in one wave and 'fair' health in another, one can conclude that her assessment of her own health has deteriorated. However, it is difficult to deduct the severity of the health decline.[15] To remedy this shortcoming, many researchers aggregate individual responses to binary health measures.[16] For example, the worst two health ratings ('fair' and 'poor') are classified as *bad health,* and the rest are classified as *good health*. Converting SRHS into a binary health measure reduces the difficulty in interpreting changes in health states. However, this comes at the cost of discarding information from an already coarse health measure. The frailty index (and its variants that we discuss below) is, by construction, a cardinal measure of health. A person with frailty of 0.2 has accumulated twice as many deficits relative to a person with frailty of 0.1. The comparison across individuals and magnitude of differences has intrinsic meaning and, as we show below, is informative about health-related outcomes.

3.1.1. Comparing frailty and SRHS on health dynamics

In addition to the qualitative differences mentioned above, frailty and SRHS also paint different pictures of the evolution of health over the life cycle. Self-reported health status partitions individuals into different categories according to their own subjective health assessments. One way to track the evolution of SRHS is to document how the size of each partition changes with age. We do this with the shaded areas in Fig. 2. 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' and 'very good' SRHS falls with age (from 65 percent for age group 25–29 to 23 percent for age group 85+). At the same time, the fraction of individuals with 'fair' or 'poor' SRHS increases with age (from 8.5 percent for age group 25–29 to 43 percent for age group 85+).

To compare the evolution of frailty (a continuous variable) with categorical SRHS we likewise partition individuals into categories using frailty. Specifically, we partition individuals within each five-year age group into five frailty categories labeled 'excellent', 'very good', 'good', 'fair', and 'poor' (same as the SRHS categories). The cutoff values of frailty that determine which category is assigned are age-independent and determined such that the distribution of individuals across frailty and SRHS categories is the same for the age group 25 to 29. For example, the fraction of 25- to 29-year-olds with SRHS of 'excellent' is 26.5 percent. We set the cutoff value for 'excellent' frailty such that 26.5 percent of 25- to 29-year-olds are also in the 'excellent' frailty category. The resulting cutoff value for frailty is 0.036. At each age, individuals with a frailty value less than 0.036 are assigned to the frailty category 'excellent'.[17] Next, we find the cutoff value for the 65th percentile (65 percent of the sample have a SRHS of 'excellent' or 'very good'): 0.071. In each age group, anyone whose frailty measure is larger than 0.036 but smaller than 0.071 is assigned to the frailty category of 'very good', and anyone whose frailty value is exactly 0.071 is randomly assigned to either the 'very good' or 'good' category. The other two cutoffs, whose values are reported in the table below Fig. 2, are chosen accordingly at the 91st and 99th percentiles and determine the assignment of the remaining individuals to the 'good', 'fair', and 'poor' frailty categories. Using this procedure, the frailty and SRHS categories of the age group 25 to 29 are perfectly aligned (by construction).[18]

The dashed lines in Fig. 2 show how the distribution of each health measure evolves with age when individuals are assigned to health categories using the method described above. As we see, the overall pattern is similar to that of SRHS. However, the decline in 'excellent'/'very good' shares and rise in 'fair'/'poor' shares happen more rapidly with age when health is measured by frailty instead of SRHS. This is especially true after age 49. One possible reason for this finding could be that at older ages, subjectivity bias leads individuals to overstate the quality of their health. For instance, older individuals may be more optimistic about their health relative to what is implied by objective measures, or as individuals age they may adjust the reference point they use when assessing their health. Another possibility is that individuals have private information about 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 Section 3.1.2.) 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 discussed in Section 3.1.2, if anything, suggest the opposite, namely, that when individuals have private information about their health it is private information that their health is worse than what is inferred from their frailty index. Regardless of the explanation for this discrepancy, health status depreciates much more rapidly with age when measured by frailty as opposed to SRHS.[19]

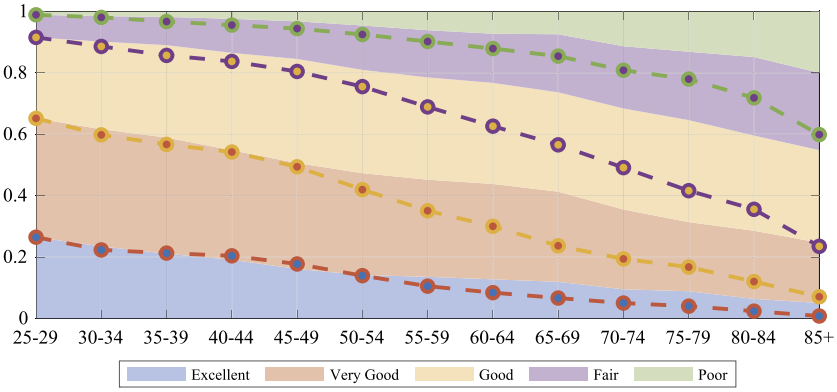
[15] De Nardi et al. (2018) overcome this loss by using the persistence of the bad health state as a measure of the severity of health shocks.

[16] See, for instance, Braun et al. (2017), De Nardi et al. (2018), French and Jones (2011), and Kitao (2014).

[17] Individuals with a frailty value equal to 0.036 are randomly assigned to either the 'excellent' or 'very good' frailty category such that, on average, 26.5 percent are assigned to the 'excellent' category.

[18] The choice to perfectly align the 25–29-year-old age group is not essential and is done to facilitate visualization. We have repeated this exercise with thresholds chosen instead to perfectly align the distribution of frailty and SRHS for the 50–54-year-old age group and found similar results to those reported here. See Section A.3.3 of the online appendix for details.

[19] These findings are not driven by noise, due to small sample sizes. In Section A.3.3 of the online appendix we provide confidence bands for the shares reported in Fig. 2. We also see similar results when we repeat this exercise in the HRS and MEPS, where sample sizes are larger. See Sections B.3.1 and C.6 of the online appendix.



	Excellent	Very good	Good	Fair	Poor
% of SRHS in age 25–29	26.5	38.7	26.4	7.4	1.1
Frailty range	[0, 0.036)	[0.036, 0.071)	[0.071, 0.143)	[0.143, 0.286)	[0.286, 1]

Fig. 2. 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 the corresponding frailty category at each age. Source: authors' calculations using the PSID. (For interpretation of the colors in the figures, see the web version of the article.)

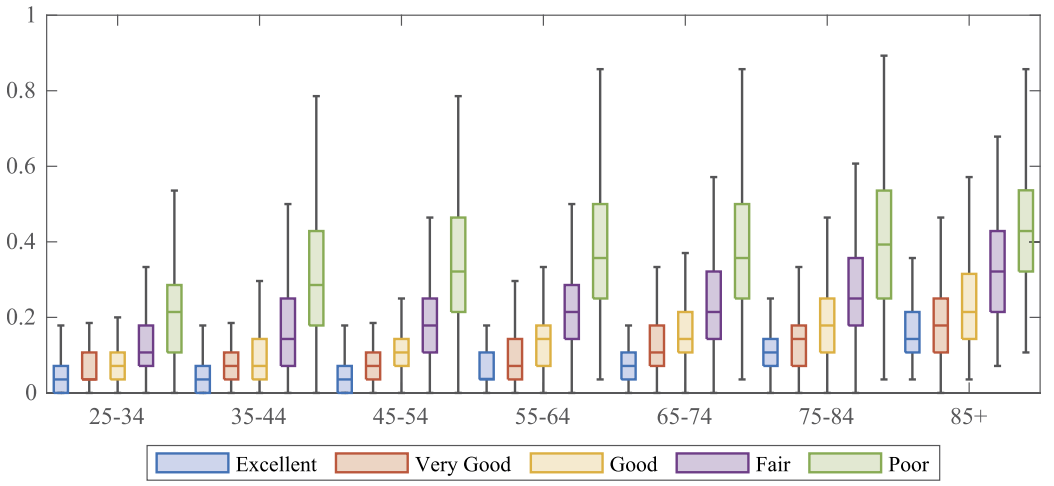


Fig. 3. Boxplots of frailty by SRHS and 10-year age groups in the PSID sample. The bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The middle lines are medians. The upper whiskers extend to 1.5 times the inner quartile range. The lower whiskers extend to the minimum value of frailty in the subgroup.

Next we show that frailty varies significantly more than SRHS at all ages. Fig. 3 shows boxplots for the distribution of frailty by SRHS categories and 10-year age groups. The bottom and top edges of the boxes indicate the 25th and 75th percentiles, respectively. The middle line indicates the median. Within each age group, median frailty increases across SRHS categories, indicating that the two health measures are positively correlated.[20] Notice that in all the age groups the inter-quartile range also increases as SRHS declines, indicating that there is more variation in frailty among individuals who self report worse health. This feature of the data also indicates that if individuals have private information about their health it is more likely that the information is that their health is worse than what is captured by the frailty index. In Section 3.1.2 we show that the dispersion in frailty within SRHS categories is informative about health-related outcomes.

Finally, we compare the persistence of SRHS with that of frailty. To do this we compute counterparts to Table 2 showing two-year transition probabilities between SRHS categories. We compare these conditional probabilities to counterparts constructed using the frailty categories defined above. We find that, even though there is more cross-sectional variation in frailty within SRHS categories, frailty is more persistent than SRHS and that the difference in persistence is largest at the poor-health end of the spectrum. See Section A.3.4 in the online appendix.

[20] Consistently, Idler and Benyamini (1997) and Van Doorsaler and Gerdtham (2002) document that SRHS is highly correlated with objective measures of health.

Table 3
Results of OLS regressions of log total medical expenditures on frailty and SRHS in wave $t-1$ using data from the MEPS. Panel A shows results from the full sample, and Panel B shows results using subsamples based on SRHS in wave $t-1$. Controls are gender, education, marital status, and a quadratic in age.

	Panel A. Everyone			**Panel B.** By SRHS at $t-1$				
	(1)	(2)	(3)	excellent (1)	very good (2)	good (3)	fair (4)	poor (5)
very good$_{t-1}$	0.377***		0.186***					
	(0.018)		(0.017)					
good$_{t-1}$	0.646***		0.195***					
	(0.018)		(0.018)					
fair$_{t-1}$	1.354***		0.379***					
	(0.023)		(0.024)					
poor$_{t-1}$	2.337***		0.755***					
	(0.033)		(0.036)					
frailty$_{t-1}$		13.335***	12.515***	15.017***	15.673***	14.929***	10.887***	5.879***
		(0.132)	(0.140)	(0.432)	(0.319)	(0.271)	(0.316)	(0.442)
frailty$^2_{t-1}$		−13.505***	−13.170***	−20.161***	−21.759***	−18.111***	−10.563***	−3.171***
		(0.223)	(0.229)	(1.070)	(0.708)	(0.507)	(0.489)	(0.558)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	124,730	124,730	124,730	30,018	37,082	35,145	16,635	5,850
R^2	0.247	0.308	0.310	0.254	0.274	0.308	0.292	0.199

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

3.1.2. Comparing frailty and SRHS on predicting health-related outcomes

We have shown that health is more persistent and deteriorates more rapidly with age when measured by frailty as compared to SRHS. Also, there is substantial variation in frailty within each of the five SRHS categories. Now, we explore how frailty and SRHS compare in predicting future health-related outcomes. We are particularly interested in the extent to which there may be independent information about health in each of these health measures. We focus on five health-related outcomes: total medical expenditures, becoming a recipient of social security disability insurance (SSDI), retiring, experiencing a nursing home stay, and mortality.

The results for total medical expenditures are reported in Table 3. Since total medical expenditures (at the individual level) are not available in the PSID, the regressions in the table were estimated using data from the MEPS.[21] The table shows the results of a series of OLS regressions with log total medial expenditures as the dependent variable.[22] All the regressions include gender, education, marital status, and age polynomials as controls. Panel A in the table shows regression results for the entire sample. In column (1) the regressors include dummies for lagged self-reported health categories. In column (2), the SRHS dummies are replaced with a quadratic in lagged frailty. Column (3) includes both sets of health indicators. Lagged self-reported health categories and frailty are the values in the previous wave (approximately two years earlier). Columns (1) and (2) indicate that variations in both frailty and SRHS are informative about variations in future medical expenditures. However, the R-squared in column (2) is higher than in column (1), indicating that frailty is more informative about future medical expenditures than SRHS is. Column (3) shows that both health measures remain statistically significant when included together, indicating that they both have independent predictive power. However, notice the large decline in the coefficients on the SRHS dummies moving from column (1) to (3), as well as the magnitude of the changes in the R-squared across columns. Both indicate that SRHS's additional impact is relatively small.

In panel B of Table 3 we repeat the same regression as in column (2) of panel A. However, instead of estimating the regression on the entire sample, we estimate five different regressions, one for each SRHS category. In each of these regressions, SRHS has no explanatory or predictive power. However, in each category, the coefficients of the frailty index are highly significant. Notice that frailty is informative about future medical expenses even among individuals who report 'excellent' health. This result is suggestive evidence that even among healthier individuals, frailty is informative about true underlying health.

We estimate similar (probit) regressions for each of the other health-related outcomes of interest; see the online appendix.[23] Remarkably, the results of these regressions show a very similar pattern to that seen for total medical expenditures in Table 3. Both frailty and SRHS have independent power in predicting outcomes. However, frailty has an edge when used alone, and it is a strong predictor of outcomes even within each SRHS category. Interestingly, the only deviation from these

[21] Recall, we construct comparable frailty indices in the MEPS and HRS. In Sections B.3 and C.3 of the online appendix we show that all of the findings regarding the relationship between frailty and SRHS reported in Section 3.1.1 also hold in these other datasets.

[22] When taking logs, medical expenditures less than $5 are set to $5.

[23] Specifically, we generated versions of Table 3 where the dependent variable is 'became an SSDI beneficiary in period t' (using both PSID and HRS data), 'entered a nursing home in period t' (using HRS data), 'died in period t' (using HRS data), and 'retired in period t' (using HRS data). In addition, we have generated versions of the tables were the dependent variable is regressed on frailty and SRHS in period $t-2$ or $t-3$ instead of period $t-1$. We also repeated the regressions in Table 3 using a sample consisting only of individuals under the age of 45. See Sections A.3, B.3, and C.3 of the online appendix.

findings occurs in the case of nursing home entry, where we find that SRHS has no independent prediction power. When frailty is included together with SRHS in the nursing home regressions, all the coefficients on SRHS become statistically insignificant. To show that all these findings are not driven by the older population alone, we repeated the regressions of total medical expenditures and SSDI recipiency using a sample consisting only of individuals under the age of 45 years and again found the same pattern of results. Finally, we also repeated all the health-related outcome regressions discussed above but with health-age interactions included. We found that adding these interaction terms does not change any of the main findings: frailty still outperforms SRHS in all cases. See the online appendix for details.

3.2. Variants of the frailty index

In this section we explore the extent to which the stylized facts documented in Section 2.1 are robust to variations in the construction of the frailty index. We focus on three variants because of two reasons. One, our regression analysis above indicates that SRHS has power in predicting health-related outcomes that is independent of our baseline frailty index; therefore, we explore the impact of adding SRHS to the index as an additional deficit variable. Two, our baseline frailty index equally weights the deficit variables. To explore how the choice of weights affects the properties of the index, we consider two different weighting schemes. One uses principal component analysis to construct the weights following Poterba et al. (2017). The other uses subjective measures of health to determine the weights based on a procedure outlined in Blundell et al. (2020).

Frailty with SRHS included as a deficit variable. This index is very similar to the baseline frailty index outlined in Section 2. The only difference is that SRHS is included as an additional deficit variable. To do this we follow De Nardi et al. (2018) and Poterba et al. (2017) (among many others) and use a value of 1 indicating *bad health* if SRHS is 'fair' or 'poor' and a value of 0 indicating *good health* if SRHS is 'good', 'very good', or 'excellent'.[24] In what follows, we refer to this index as "frailty with SRHS."

Frailty using first principal component weighting of deficits. This index includes all the deficit variables in frailty with SRHS. However, following Poterba et al. (2017), the index is given by the first principal component of the deficits. Thus, the index is effectively a weighted average of the deficits with weights chosen to maximize the proportion of the captured variance in the deficit variables. In what follows, we refer to this index as the "first principal component index", or "FPC index" for short.

Consistent with the index constructed by Poterba et al. (2017), our FPC index gives the highest weights to ADL/IADL variables. It gives the lowest weights to the high BMI and smoking variables. In particular, aggregating the weights across deficits by category we find that it puts 63 percent of the total weight on the ADL/IADL variables, 27 percent on the medical variables, and 6 percent on the mental health variables with the remaining 4 percent on SRHS. In contrast, frailty (and frailty with SRHS) equally weights each deficit. Consequently, frailty puts 46 percent of the weight on the ADL/IADL variables, 46 percent on the medical variables, 7 percent on the mental health variables, and zero weight on SRHS (because it is not included). Similarly, the distribution of weights for frailty with SRHS is 45, 45, 7, and 3 percent for the four aforementioned variables, respectively.[25] Thus, the main difference between the FPC index and both frailty and frailty with SRHS is that the FPC index puts more weight on the ADL/IADL variables, while the other indices put more weight on the medical variables.

Frailty using subjective measures to inform the weights. The idea behind the third index is to construct a weighted sum of deficits where SRHS and other subjective measures of health are used to inform the weights instead of being included directly into the index as additional deficit variables. The idea for the index is inspired by the use of objective measures of health to instrument subjective ones as a way to control for justification bias in subjective health measures when studying the impact of health on labor supply.[26] We call it the "subjective instrumenting objective index", or "SIO index" for short.

Our procedure for constructing the SIO index is similar to that in Blundell et al. (2016) and Blundell et al. (2020).[27] The construction follows three steps. In the first step we use principal component analysis to combine three subjective health measures: SRHS, whether health limits daily activities, and whether health limits the ability to work. Let H_{it} denote the first principal component of these three variables for individual i at age t. In the second step we estimate the regression

$$H_{it} = \alpha + \sum_{k=1}^{K} \gamma_k D_{it}^k + \epsilon_{it}, \tag{1}$$

using OLS, where D_k $(k = 1, \ldots, K)$ is a list of K deficits. The set of deficit variables used in this second step is the same set we use to construct the baseline frailty index. In the PSID, these are the 28 deficit variables discussed at the beginning

[24] As we mentioned above, SRHS is not a cardinal measure and assigning numeric values to each health category would be arbitrary. For this reason, we instead follow the literature and aggregate the information in SRHS into a 0/1 deficit variable.

[25] Table 2 in the online appendix compares the full set of weights across the three indices.

[26] See Bound (1991), Disney et al. (2006), and Stern (1989).

[27] See Section B.6 of the online appendix for a detailed comparison of our SIO index to their instrumental variables approach.

Table 4
Summary statistics for four versions of the frailty index: frailty, frailty with SRHS, FPC index, and SIO index using PSID data on household heads and their spouses ages 25 and older.

	Frailty (1)	Frailty w/SRHS (2)	FPC (3)	SIO (4)
Mean	0.11	0.12	0.08	0.10
men	0.11	0.11	0.07	0.08
women	0.12	0.12	0.09	0.11
ages 25-49	0.08	0.08	0.05	0.06
ages 50-74	0.15	0.15	0.11	0.13
ages 75+	0.25	0.25	0.24	0.28
less than HS	0.16	0.17	0.13	0.15
HS grad	0.12	0.12	0.08	0.10
college grad	0.08	0.08	0.06	0.06
Standard dev.	0.11	0.12	0.12	0.16
men	0.10	0.11	0.11	0.14
women	0.12	0.12	0.13	0.17
ages 25-49	0.09	0.09	0.08	0.11
ages 50-74	0.12	0.13	0.13	0.18
ages 75+	0.16	0.16	0.20	0.23
less than HS	0.14	0.15	0.16	0.20
HS grad	0.11	0.11	0.12	0.16
college grad	0.09	0.09	0.09	0.12
Min	0.00	0.00	0.00	0.00
5th percentile	0.00	0.00	0.00	0.00
50th percentile	0.07	0.07	0.04	0.04
95th percentile	0.36	0.34	0.33	0.50
Max	0.89	0.90	0.95	0.97
Fraction at zero	0.15	0.15	0.15	0.15
Observations	96,759	96,759	96,759	96,416
Individuals	19,996	19,996	19,996	19,963

of Section 2. In the third step we normalize the estimated coefficients $\hat{\gamma}_k$ such that they add up to 1. The SIO index is then defined as the sum of deficit variables weighted by the normalized coefficient estimates, or

$$H_{it}^{SIO} \equiv \frac{\sum_{k=1}^{K} \hat{\gamma}_k D_{it}{}^k}{\sum_{k=1}^{K} \hat{\gamma}_k}.$$

The SIO index is essentially a weighted sum of the deficit variables where the weights are chosen to minimize the distance between the weighted sum and the subjective health measure.[28] Similar to the FPC index, the SIO index puts more weight on the ADL/IADL variables (66 percent) than on the medical variables (26 percent). However, while the FPC index gives a relatively equal share of weight to each of the 13 ADL/IADL variables, the SIO index assigns more than two-thirds of the ADL/IADL weight (44 percent of the total weight) to only two deficit variables: difficulty with heavy housework and difficulty walking, indicating that compared to the other deficits, these two deficits are more highly correlated with individuals' subjective assessment of their health. Interestingly, compared to the FPC index, the SIO index also puts a higher weight on the mental health variables (9 percent).

Table 4 shows summary statistics for the three variants of the frailty index we introduced above. In column (1) we repeat the summary statistics for frailty. Columns (2) through (4) show statistics for frailty with SRHS, the FPC index, and the SIO index, respectively. Fig. 4 shows boxplots of these indices by age and histograms showing their cross-sectional distributions. Columns (1) and (2) are almost indistinguishable; therefore, adding SRHS as a deficit variable has very little impact on the frailty index. This fact can also be seen by comparing the boxplots and histogram in panels (a) and (b) of Fig. 4 with those in Fig. 1.

Comparing columns (1) and (3) of Table 4 reveals that the FPC index has a lower mean relative to frailty. As the table shows, this is primarily due to lower values of the FPC index relative to frailty at younger ages. The FPC index assigns lower values at younger ages because, as we discussed above, compared to frailty, it puts more weight on the deficit variables capturing difficulties with ADL/IADLs and these deficits are more prevalent among older individuals. The FPC index also has slightly higher dispersion than frailty. This is driven by a higher standard deviation at older ages (ages 85 and above), which is again because relative to frailty, FPC puts more weight on deficits that disproportionately affect older individuals. Thus,

[28] A few of the weights on individual deficit variables in the SIO index are negative, indicating that controlling for other deficits, having these particular deficits is negatively correlated with individuals' subjective health measure. See Table 2 in the online appendix for details.

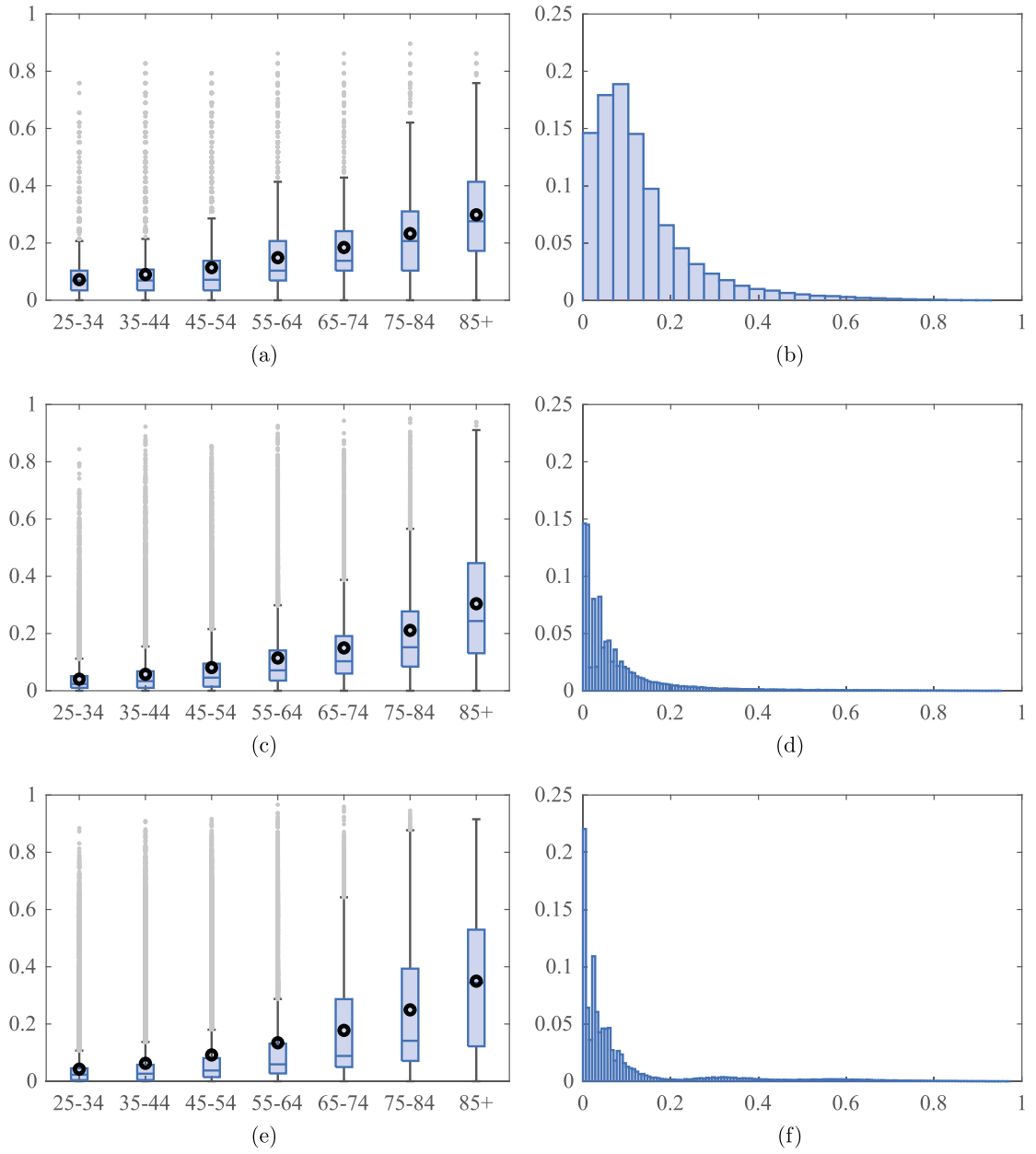


Fig. 4. Boxplots of frailty with SRHS (a), FPC index (c), and SIO index (e) by 10-year age groups in the PSID sample. The bottom and top edges of the boxes indicate the 25th and 75th percentiles, respectively. The middle lines are medians, and the open circles are means. The upper whiskers extend to 1.5 times the inner quartile range. The lower whiskers extend to the minimum value of frailty in the age group. The dots are data points that lie beyond the whiskers. Histograms showing the cross-sectional distribution of frailty with SRHS (b), FPC index (d), and SIO index (f) in the sample.

unlike with frailty, the difference in the FPC index between individuals with and without deficits is larger at older ages than at younger ones. As a result, dispersion in the FPC index rises relatively more rapidly with age.

Finally, as Fig. 4 and Table 4 show, the SIO index exhibits both more variation (particularly at older ages) and more skewness (particularly at younger ages) as compared to the other three indices. These features of the SIO index are due to its relatively more unequal weighting of deficits. Recall that the SIO index puts 44 percent of the weight on just two deficit variables (difficulty doing heavy housework and difficulty walking), both of which are more common at older ages. The differences in frailty between individuals with and without these particular deficits are larger than in the frailty index, where deficits are equally weighted, and in the FPC index, where the distribution of weights is also relatively more equal. This increases skewness at younger ages when the deficits most individuals are assigned have relatively low weights and, as in the case of the FPC index, leads to a faster rate of increase in dispersion in the index with age. Notice that, interestingly, the SIO index also implies slightly larger differences in health by gender than the other three indices do.

Table 5
OLS regressions of log total medical expenditures on SRHS, frailty, frailty with SRHS, the FPC index, and the SIO index at $t-1$ using data from the MEPS. Panel A shows results using the full sample, and Panel B shows results using a subsample consisting of individuals younger than 45. Controls are gender, education, marital status, and a quadratic in age.

	Panel A. All ages					**Panel B.** Younger than 45				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
very good$_{t-1}$	0.377***					0.344***				
	(0.018)					(0.027)				
good$_{t-1}$	0.646***					0.550***				
	(0.018)					(0.028)				
fair$_{t-1}$	1.354***					1.189***				
	(0.023)					(0.040)				
poor$_{t-1}$	2.337***					2.623***				
	(0.033)					(0.072)				
frailty$_{t-1}$		13.335***					15.355***			
		(0.132)					(0.298)			
frailty$^2_{t-1}$		−13.505***					−16.608***			
		(0.223)					(0.660)			
frailty w/SRHS$_{t-1}$			13.017***					14.752***		
			(0.130)					(0.288)		
frailty w/SRHS$^2_{t-1}$			−12.847***					−15.256***		
			(0.217)					(0.628)		
FPC$_{t-1}$				11.196***					14.714***	
				(0.131)					(0.309)	
FPC$^2_{t-1}$				−10.763***					−15.727***	
				(0.188)					(0.521)	
SIO$_{t-1}$					8.870***					11.044***
					(0.126)					(0.280)
SIO$^2_{t-1}$					−7.555***					−9.806***
					(0.169)					(0.424)
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	124,730	124,730	124,730	124,730	116,745	52,571	52,571	52,571	52,571	49,132
R^2	0.247	0.308	0.308	0.285	0.279	0.153	0.197	0.198	0.185	0.182

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Despite the differences highlighted above, Table 4 and Fig. 4 show that all three variants of the frailty index convey the same facts about the evolution of health over the life cycle as our baseline frailty measure: Frailty increases with age, decreases with education, and is slightly higher for women than for men. Dispersion in frailty increases with age, and the cross-sectional distribution of frailty is right skewed both overall and by age. Like frailty, frailty with SRHS, the FPC index, and the SIO index are also highly persistent (see Tables 25, 26 and Table 27 in the online appendix).

3.3. Predicting health-related outcomes

We have shown that the stylized facts on health dynamics over the life cycle documented using the frailty index are robust to variations in its construction. We now examine the impacts of these variations on the index's performance in predicting future health-related outcomes. Similar to Section 3.1.2, we focus on the following health-related outcomes: total medical expenditures, becoming an SSDI recipient, experiencing a nursing home stay, and mortality.[29]

Consider first the relationship between each health measure and total medical expenditures. Using MEPS data, we estimate a series of OLS regressions of log total medical expenditures on lagged values of each of the health measures and a set of controls: for gender, education, marital status, age, and age-squared. The results for the full sample (individuals ages 25 and older) are presented in Panel A of Table 5. In column (1), the regressors include dummies for the SRHS categories. In columns (2) to (4) the regressors include a quadratic polynomial in frailty, frailty with SRHS, the FPC index, and the SIO index, respectively. Columns (1) and (2) are identical to columns (1) and (2) in Table 3, but we include them here for ease of comparison. Comparing the R-squared across columns shows that all four variants of the frailty index perform similarly well.[30] Each of them also performs better than SRHS. In particular, all four measures predict about 30 percent of the total variation in medical expenditures two years ahead, while SRHS predicts slightly less (25 percent).[31]

[29] We also compare the indices on their ability to predict retirement and find similar results. See Section C.4 in the online appendix.

[30] We use identical samples to compare across measures except in the case of the SIO index. In each dataset, we lose some observations for this index because of missing values of the additional subjective health variables used in its construction. However, we have repeated all the outcome regressions in this section holding the sample fixed and found that the relative performance of the measures does not change.

[31] Our analysis focuses on comparisons of predictive power as measured by the R-squared. However, there may be other factors worth considering when comparing across indices. For instance, one potential advantage of the SIO index is that its construction removes uncorrelated measurement error in

Table 6
Probit regressions of 'becomes an SSDI recipient at date t' on SRHS, frailty, frailty with SRHS, the FPC index, and the SIO index at $t-1$ using data from the PSID. Panel A shows results for all individuals under age 66, and Panel B shows results using a subsample consisting of individuals younger than 45. Controls are gender, education, marital status, and a quadratic in age.

	Panel A. Younger than 66					**Panel B.** Younger than 45				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
very good$_{t-1}$	0.080					0.113				
	(0.073)					(0.100)				
good$_{t-1}$	0.487***					0.330***				
	(0.067)					(0.097)				
fair$_{t-1}$	1.013***					0.999***				
	(0.069)					(0.099)				
poor$_{t-1}$	1.622***					1.550***				
	(0.078)					(0.125)				
frailty$_{t-1}$		7.380***					6.964***			
		(0.385)					(0.651)			
frailty$^2_{t-1}$		−5.558***					−4.370***			
		(0.654)					(1.175)			
frailty w/SRHS$_{t-1}$			7.503***					7.143***		
			(0.382)					(0.647)		
frailty w/SRHS$^2_{t-1}$			−5.668***					−4.584***		
			(0.635)					(1.146)		
FPC$_{t-1}$				8.199***					8.540***	
				(0.323)					(0.572)	
FPC$^2_{t-1}$				−7.365***					−7.658***	
				(0.526)					(0.995)	
SIO$_{t-1}$					5.659***					6.014***
					(0.273)					(0.482)
SIO$^2_{t-1}$					−4.506***					−4.758***
					(0.413)					(0.760)
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	45,906	45,906	45,906	45,906	45,780	23,475	23,475	23,475	23,475	23,419
Pseudo R^2	0.187	0.232	0.241	0.260	0.239	0.153	0.218	0.229	0.255	0.230

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

In Panel B we repeat the exercise on a subsample that is younger than 45. We do this to demonstrate that even among the younger individuals in our sample all four frailty variants are informative about medical expenditures. This finding indicates that variation in frailty among younger individuals is likely reflecting true variation in their underlying health. Even for individuals under 45, all four frailty variants remain more informative than SRHS about future total medical expenditures. Both overall and within the younger subsample, frailty and frailty with SRHS marginally outperform the other measures.

Next, consider the health measures' ability to predict the probability of becoming an SSDI recipient. To explore the relationship between each health measure and this probability, we use a restricted version of the PSID sample consisting of observations for individuals younger than age 66 who were not SSDI beneficiaries in the previous wave. Using this sample, we estimate a series of probit regressions of SSDI beneficiary status on lagged values of each health measure and the same set of controls as described above. The dependent variable is 1 at age t if the individual has become an SSDI beneficiary and 0 otherwise. Panel A of Table 6 shows the results for individuals younger than 66. Panel B repeats the exercise for those younger than 45. To evaluate and compare the performance of each health measure we report the *pseudo R-squared* for each regression. It is calculated as 1 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.[32] We use the pseudo R-squared as a measure of predicted variation in the dependent variable.

In terms of predicting future SSDI recipiency, FPC performs better than the other health measures, both for the full sample of individuals under age 66 and the subsample of those under age 45. Recall that the FPC index puts higher weight than frailty and frailty with SRHS on the ADL/IADL variables. So, it is not surprising that having several difficulties with ADLs and IADLs is highly correlated with becoming an SSDI beneficiary. Interestingly, the SIO index, which disproportionately

subjective and objective health measures. For this reason, if measurement error biases are significant, the SIO index may provide a relatively more accurate measure of health effects. See Blundell et al. (2020) for details.

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

Table 7
Probit regressions of 'dead at date t' and 'nursing home entry at date t' on SRHS, frailty, frailty with SRHS, the FPC index, and the SIO index at $t-1$ using data from the HRS. Panel A shows the results for mortality, and Panel B shows the results for nursing home entry. Controls are gender, education, marital status, and a quadratic in age.

	Panel A. Mortality					**Panel B.** Nursing home entry				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
very good$_{t-1}$	0.053**					0.008				
	(0.024)					(0.044)				
good$_{t-1}$	0.293***					0.139***				
	(0.023)					(0.042)				
fair$_{t-1}$	0.649***					0.360***				
	(0.023)					(0.043)				
poor$_{t-1}$	1.186***					0.700***				
	(0.024)					(0.045)				
frailty$_{t-1}$		2.970***					1.975***			
		(0.098)					(0.211)			
frailty$^2_{t-1}$		−0.490***					0.160			
		(0.120)					(0.269)			
frailty w/SRHS$_{t-1}$			3.010***					1.981***		
			(0.098)					(0.209)		
frailty w/SRHS$^2_{t-1}$			−0.524***					0.133		
			(0.119)					(0.265)		
FPC$_{t-1}$				2.349***					1.774***	
				(0.074)					(0.159)	
FPC$^2_{t-1}$				−0.318***					−0.092	
				(0.088)					(0.196)	
SIO$_{t-1}$					0.376***					0.656***
					(0.110)					(0.230)
SIO$^2_{t-1}$					1.157***					0.374
					(0.137)					(0.281)
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	212,978	212,978	212,978	212,978	133,707	168,412	168,412	168,412	168,412	113,769
Pseudo R^2	0.217	0.241	0.244	0.237	0.193	0.231	0.261	0.262	0.260	0.266

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

weights two ADL/IADLs (difficulty with heavy housework and difficulty walking) does not perform quite as well. As was the case with total medical expenditures, SRHS again performs the worst relative to the four frailty variants.

Next, using HRS data, we estimate mortality and nursing home entry probits for each health measure. In each mortality probit estimation we regress a variable that is equal to 1 if an individual died between wave $t-1$ and wave t on one of the health measures and the set of controls. The results are reported in Panel A of Table 7. In each nursing home entry probit we regress a variable that is equal to 1 if an individual has a nursing home stay at age t but did not have one at $t-1$ on each health measure plus the controls. For the nursing home entry regressions we restrict the sample to observations for individuals who did not have nursing home stays at $t-1$. The results are reported in Panel B. Again, we evaluate and compare the performance of each health measure using the *pseudo R-squared* for each regression.

Notice that frailty, frailty with SRHS, and the FPC index each predict approximately 24 percent of mortality variation. All three indices perform significantly better than the SIO index and SRHS. When it comes to mortality, the SIO index has the worst performance overall, predicting 19 percent of the variation in mortality, whereas SRHS predicts 22 percent. Interestingly, while the SIO index does not predict mortality well relative to the other health measures, it is the best predictor overall of nursing home entry, predicting nearly 27 percent of its variation. In comparison, frailty, frailty with SRHS, and the FPC index do marginally worse (predicting closer to 26 percent). However, as was the case with medical expenditures and becoming an SSDI beneficiary, when it comes to nursing home entry again all four variants of frailty perform better than SRHS (which predicts 23 percent).

3.4. Taking stock

We have shown that frailty, frailty with SRHS, the FPC index, and the SIO index all imply qualitatively similar life cycle health dynamics. Yet, we have also noted several quantitative differences between frailty and frailty with SRHS on the one hand, and the FPC and SIO index on the other. Despite these quantitative differences, we found that the differences between the four indices in predicting health-related outcomes are relatively small. In addition, while some variants perform better than others at predicting specific outcomes, no variant is consistently better overall.

However, frailty has one advantage over the others: it is the simplest to construct. Thus, in the next section, we use frailty to estimate a stochastic process of health dynamics over the life cycle. We have also estimated the process using the other variants and have found that, perhaps not surprisingly, the quantitative differences between frailty and the SIO and

FPC index lead to quantitatively different estimation results.[33] We discuss these findings and their implications further at the end of the section.

4. Frailty dynamics over the life cycle

In this section we propose and estimate a statistical model of frailty dynamics over the life cycle. The model is as parsimonious as possible while still being flexible enough to capture the main qualitative properties of frailty documented in Section 2.1. We first present the model and describe the estimation procedure which accounts for selection due to mortality. Then we present results from the baseline estimation of the model. Finally, we present results from estimating the model separately for men and women, and then separately for different education groups.

4.1. Statistical model

4.1.1. Zero frailty probability

One important feature of the data is that many young people have no health deficits. For instance, nearly 30 percent of 25-year-olds in our PSID sample have zero frailty.[34] The fraction of individuals with zero frailty gradually declines with age. To capture this feature of the data, we allow for a mass of people with zero frailty at age 25. Each period, these people have either a probability of staying at zero frailty or 1 minus that probability of going to a positive frailty value. Once a person's frailty becomes positive, we assume it never goes back to zero.[35]

Let f_{ij} denote the frailty of individual i at age j. Each period, the probability that the individual's frailty is zero follows a probit model, that is,

$$Prob(f_{ij} = 0|Z_{ij}) = \Phi\left(Z'_{ij}\gamma\right), \tag{2}$$

where Φ is the cdf of the standard normal distribution and Z_{ij} is a set of covariates: age, age^2, high school graduate and college graduate education dummies, and a male gender dummy. The probability of having zero frailty conditional on having zero frailty in the previous period is given by

$$Prob(f_{ij} = 0|f_{ij-1} = 0) = Prob(f_{ij} = 0|Z_{ij})/Prob(f_{ij-1} = 0|Z_{ij-1}). \tag{3}$$

Thus, at age 25 the probability of having zero frailty is given by equation (2). At age $j > 25$ this probability is given by equation (3) if frailty is zero at age $j-1$ and zero otherwise.

4.1.2. Nonzero frailty dynamics

Our statistical model of the dynamics of frailty at nonzero values is similar to ones used for earnings processes.[36] In particular, we assume that the log of the frailty index, $\ln f_{ij}$, for individual i at age j is the sum of a deterministic component whose effect is common to all individuals and a residual that is individual-specific:

$$\ln f_{ij} = X'_{ij}\beta + R_{ij}, \tag{4}$$

where β is an 8×1 vector of coefficients and X_{ij} is a set of covariates that includes a fourth order age polynomial, high school graduate and college graduate education dummies, and a male gender dummy.[37] The residual consists of two components and is given by

$$R_{ij} = \alpha_i + z_{ij} + u_{ij}. \tag{5}$$

The first component, α_i, is individual specific and allows us to capture ex ante heterogeneity in individuals' initial frailty levels. We assume that α_i is normally 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_{ij}.[38] Thus

$$z_{ij} = \rho z_{ij-1} + \varepsilon_{ij}, \tag{6}$$

[33] The results from estimating the stochastic process using the other three variants of frailty are reported in Section E of the online appendix.

[34] The most common deficits of 25–29-year-olds with positive frailty are the following: has ever smoked, currently smokes, has a BMI over 30, and has a serious, chronic health condition other than those that are asked about explicitly.

[35] This is a realistic assumption given that less than 1 percent of individuals in our PSID sample with positive frailty have zero frailty next period.

[36] See, for instance, Guvenen (2009), Karahan and Ozkan (2013), and Storesletten et al. (2004).

[37] Age values are re-centered by subtracting 25 and then normalized by dividing by 100 to improve the numerical performance of the simulated method of moments estimation procedure.

[38] The variance of frailty increases with both age and time in the PSID sample. However, the increase with age is much more dramatic. Therefore, we chose a specification with an age-dependent but time-invariant stochastic component.

where $z_{i,24} = 0$ and if $f_{ij-1} = 0$, then $z_{ij-1} = 0$. The shocks ε_{ij} and u_{ij} are assumed to be independent of each other and over time, and independent of α_i. We assume that u_{ij} is normally distributed with mean zero and variance σ_u^2 and that ε_{ij} is normally distributed with mean zero and variance σ_ε^2. 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 variance of the innovations to the persistent process, σ_ε^2, determine individuals' exposure to persistent health shocks.[39] Our procedure for estimating ρ places no restrictions on its value; that is, we do not rule out the possibility that shocks to frailty are permanent.

4.2. Frailty process estimation

Frailty is highly correlated with mortality. Thus, as we show below, it is important to control for selection bias stemming from sample attrition due to mortality when estimating the frailty process. For this reason, we estimate the frailty process via simulated method of moments (SMM).[40] That is, we choose the parameters of the frailty process by minimizing the distance between moments constructed in the data and counterparts constructed using artificial data. The artificial data is generated by simulating a dynamic model of frailty and mortality over the life cycle for a sample of individuals with the same distribution across education and gender at ages 25–26 as in our PSID sample.

In the model, frailty dynamics are governed by equations (5) and (6). To simulate mortality risk we assume that the probability individual i with frailty index value in $[0, 1]$ and who is alive at age j, $d_{ij} = 0$, is dead at age $j+1$ is given by

$$Prob(d_{ij+1} = 1 | d_{ij} = 0) = \Phi\left(Y_{ij}'\delta\right), \tag{7}$$

where Φ is the standard normal CDF. If an individual's frailty index value is above 1, we assume he dies immediately.[41] The covariates in Y_{ij} are age, age^2, frailty, frailty^2, education, and a gender dummy.

We estimate equation (7) outside the SMM procedure using HRS data. The HRS has much larger sample of older individuals relative to the PSID, making it better suited for estimating mortality rates.[42] The results of the estimation are reported in Table 89 in the online appendix.[43] The estimation results indicate that frailty, age, and gender all have a statistically significant effect on mortality and that mortality rates increase in age over the range 25 to 95 and frailty over the range $[0, 1]$. Interestingly, having controlled for frailty, we do not find a statistically significant effect of education on mortality.

The probability of having zero frailty, equation (2), is also estimated directly as opposed to through the SMM procedure. This choice is made because the overall impact of ignoring the effect of mortality bias on this probability is small. At young ages, when the probability of zero frailty is large, mortality risk is very low, and at later ages, when mortality effects become important, the probability of zero frailty is small. Fig. 5 plots the fraction of individuals with zero frailty by age in the PSID data, the predicted fraction from the probit regression, and the predicted fraction, accounting for mortality, from the simulation model. As the figure shows, mortality has little effect on the simulated fraction. Table 8 provides the probit regression estimation results. Notice that, on average, the probability of zero frailty is increasing in education and is higher for women than for men.

Nonzero frailty dynamics in our model are governed by 12 parameters: the eight coefficients of the deterministic component contained in β and the four parameters that characterize the stochastic component, specifically, σ_u^2, σ_ε^2, σ_α^2, and ρ. Let θ denote a vector of these parameters. To estimate θ via SMM we construct target moments, denoted by m^T, using our PSID sample. To construct the target moments we first adjust log frailty for year and sample duration effects.[44] We then construct two sets of target moments.

The first set of targets consists of 175 moments which are used to determine the eight parameters governing the deterministic component of the nonzero log frailty process. These moments are mean log frailty by education (less than high school, high school graduates, and college graduates) and two-year age groups, and mean log frailty by gender and two-year age groups for individuals ages 25–94. The moments are presented as the open circles in Figs. 6a and 6b. Notice that conditional on age, there is much more variation in log frailty by education than by gender in the data.

[39] Under this specification $var(f_{ij}|f_{ij-1})$ is proportional to f_{ij-1}. Thus, individuals with a higher frailty level will also have a higher conditional variance of shocks consistent with the observed patterns in the probability transition matrices presented in Table 2.

[40] See Duffie and Singleton (1993), McFadden (1989), and Pakes and Pollard (1989).

[41] This assumption has negligible effects on life expectancy. Given our estimated mortality rates, frailty values near but slightly below 1 are relatively rare (see Fig. 7) and values above 1 are even rarer. For instance, absent this assumption, in artificial samples generated using our baseline estimation results, values of frailty above 1 account for less than 0.4 percent of observations.

[42] This choice presumes that the relationship between mortality and frailty is the same in both PSID and HRS data. Section C.6 of the online appendix shows that the dynamics of the frailty distributions in the two datasets are similar, suggesting that this is a reasonable assumption.

[43] The mortality rates implied by our estimation results and simulated data are consistent with ones computed using the year 2000 period life-table in Bell and Miller (2005).

[44] Recall from footnote 8 that we forward correct reporting inconsistencies in the 'have you ever' deficit questions in the PSID data. Removing sample duration effects eliminates any spurious trends in frailty by sample duration that this procedure may introduce. This is done by regressing log frailty on year and sample duration dummies and then removing their estimated effects. Log frailty is then rescaled so that it has the same mean as before.

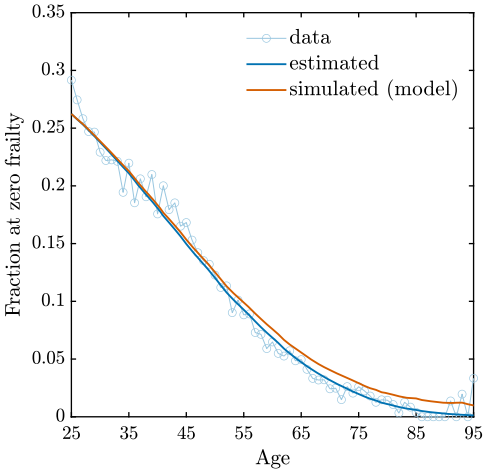


Fig. 5. Fraction with zero frailty by age in the data (blue open circles), predicted from the probit regression (solid blue line), and predicted, after accounting for mortality, by the model (solid orange line).

Table 8
Estimation results from the zero frailty probit regression.

Variable	Age	Age^2	High school	College	Male	Constant
Value	4.92e−4	−2.98e−4***	0.104***	0.299***	−0.063***	−0.494***
Std. error	(2.73e−3)	(2.95e−5)	(0.017)	(0.017)	(0.010)	(0.062)

Pseudo $R^2 = 0.0832$, No. of observations = 96,252

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

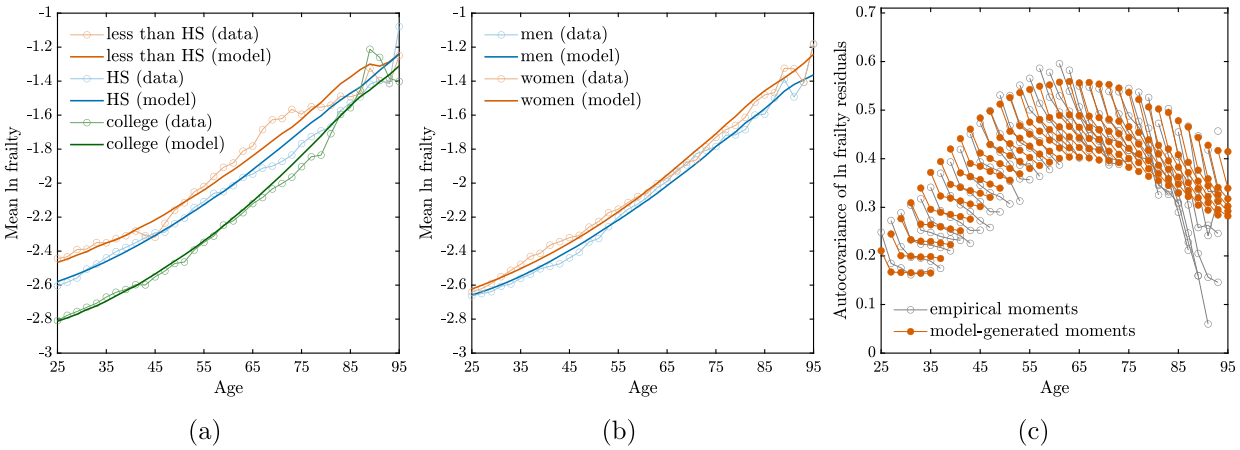


Fig. 6. Fit of the estimated model: empirical moments targeted in the baseline SMM estimation and their model counterparts. Panel (a) is mean log frailty by age in the data (open circles) and in the model (solid lines) for those without a high school degree (orange), high school graduates (blue), and college graduates (green). Panel (b) is mean log frailty by age in the data (open circles) and in the model (solid lines) for women (orange) and men (blue). Panel (c) is autocovariance profiles of log frailty residuals by age in the data (open gray circles) and in the model (closed orange circles).

The second set of targets consists of autocovariance age profiles conditional on survival which are used to determine the four parameters governing the stochastic component.[45] To construct these moments we first remove variation in observed log frailty due to age, education, and gender by regressing log frailty on a quadratic in age, education dummies, and a gender dummy using OLS. Taking the residuals from this regression, to construct the variances we group individuals into two-year non-overlapping age groups (25–26-year-olds, 27–28-year-olds, and so on).[46] The variance profile is then given

[45] Autocovariance age profiles have been used extensively to estimate stochastic processes similar to the one in equation (5) especially in the literature on life cycle earnings dynamics. See, for instance, Abowd and Card (1989), Baker and Solon (2003), and Guvenen (2009).

[46] We estimate the model at the annual frequency. However, since our data is biennial, we can compute empirical covariances only between current frailty residuals and lagged values of frailty residuals that are multiples of two.

by the means of the squared residuals for each age group. As is commonly done in the literature on earning dynamics, we adjust the variances for cohort effects.[47] We do this by regressing the raw variances on a full set of age and cohort dummies and removing the estimated cohort effects. 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 45 is the same as the raw variance at age 45.

For each age group, we construct up to five autocovariances. The covariances are constructed similarly to the variances. For instance, the first autocovariance at each age j is given by the means of the product of the age j and age $j+1$ residuals. To cohort-adjust the 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.[48] The resulting empirical variance-covariance matrix targeted in the SSM estimation is shown as the gray open circles in Fig. 6c.[49] The first open circle on each gray line in the figure is the variance of that age group's log frailty residual. The next open circle is the first autocovariance, followed by the second autocovariance, and so on.

The cross-sectional variance of log frailty in Fig. 6c is hump-shaped over the life cycle. Recall that Fig. 1a in Section 2.1 shows that the variance of frailty in levels is increasing over the life cycle. The reason the variance profile in logs has a different shape to the variance profile in levels is mortality. As individuals age, dispersion in frailty increases due to the persistence of frailty shocks. However, their frailty levels also increase, driving up mortality rates at the top of the frailty distribution and putting downward pressure on frailty dispersion conditional on survival. In levels, rising dispersion at the top is still large enough to generate rising overall dispersion in frailty even at older ages. However, the log function places more weight on the bottom of the distribution. Thus, in logs, at later ages when mortality rates are high, declining dispersion at the bottom of the distribution is the dominant effect.

Also observable in Fig. 6c is that the autocovariance profiles are declining in the lag length. Specifically, the covariance between frailty at age t and frailty at age $t+k$ is declining in k. Notice that the rate at which they decline with lag length is increasing in age, resulting in autocovariance profiles that become increasingly steeper at older ages. We show below that this feature of the autocovariance profiles is also due to mortality.

Given values for the parameters in θ, we can construct counterparts to the target moments by using the model to generate artificial data. We do this by simulating the life cycle frailty and mortality dynamics of 50,000 individuals. Each individual starts life at age 25 and we simulate their life cycle for 72 periods or until age 94. The initial distribution of the individuals at age 25 across gender, education, and frailty is given by the distribution of 25–26-year-olds across these variables in the PSID data. Simulated moments are constructed in the same way as target moments. First, using the artificial data, we compute mean log frailty conditional on survival by education and two-year age groups, and then by gender and two-year age groups. Then, just as we did for the target moments, conditional on individuals' survival, we estimate an OLS regression of log frailty on a quadratic in age, education dummies, and a gender dummy. We use the residuals from these regressions to construct counterpart variance-covariance moments by two-year age groups. Denote the simulated moments by $m^S(\theta)$.

The SMM procedure finds the vector of parameters θ that solves

$$\min_{\theta}[m^T - m^S(\theta)]' W [m^T - m^S(\theta)], \tag{8}$$

where the weighting matrix W is set to a diagonal matrix of the inverse of the estimated variance of each moment. This weighting matrix puts less weight on moments that are less precisely measured, due, for instance, to small sample size. In general, the matrix puts less weight on the higher-order autocovariances and the moments constructed using older individuals.[50] We obtain standard errors on the parameters in θ from an approximation of its Jacobian. The approximation is computed using a 4th-order centered difference.

4.3. Baseline estimation results

Table 9 provides the results from the SSM estimation on our main PSID sample. Consider first the left side of the table, which shows the estimates of the parameters governing the deterministic component of the log frailty process. The moments targeted to identify these parameters are the age profiles of mean log frailty by education and gender. These moments and their model counterparts can be seen in Figs. 6a and 6b. Overall the fit of the model to the moments in the data is excellent. Notice in Fig. 6a that in both the model and the data, the differences in log frailty by education decline with age. This convergence in observed frailty, which occurs even though the specification in equation (4) assumes that the effect of education on frailty is independent of age, is due to a differential effect of education (via frailty) on mortality. Controlling for age, less educated individuals have higher frailty on average. Since mortality is increasing in frailty, these individuals consequently have higher mortality rates.

[47] See Deaton and Paxson (1994) and Storesletten et al. (2004).

[48] See Section D.1 of the online appendix for additional details on the construction of the cohort-adjusted variance-covariance matrix.

[49] Moments constructed with 30 or less observations are dropped.

[50] We found that the estimation results are very similar if equal weighting is used instead.

Table 9
Baseline results from estimating the log frailty process via SMM using a sample of individuals aged 25–95 from the PSID.

Deterministic component			**Stochastic component**		
Variable	Value	Std. error	Variable	Value	Std. error
age*	1.008	(0.008)	σ_u^2	0.042	(0.001)
age*^{2}	2.073	(0.023)	σ_ε^2	0.019	(0.001)
age*^{3}	−1.154	(0.040)	σ_α^2	0.140	(0.004)
age*^{4}	5.393	(0.039)	ρ	0.991	(0.001)
HS degree	−0.116	(0.001)			
college degree	−0.339	(0.001)			
male	−0.030	(0.001)			
constant	−2.460	(0.001)			

Note: age* ≡ (age − 25)/100. Stochastic component estimates are annual.

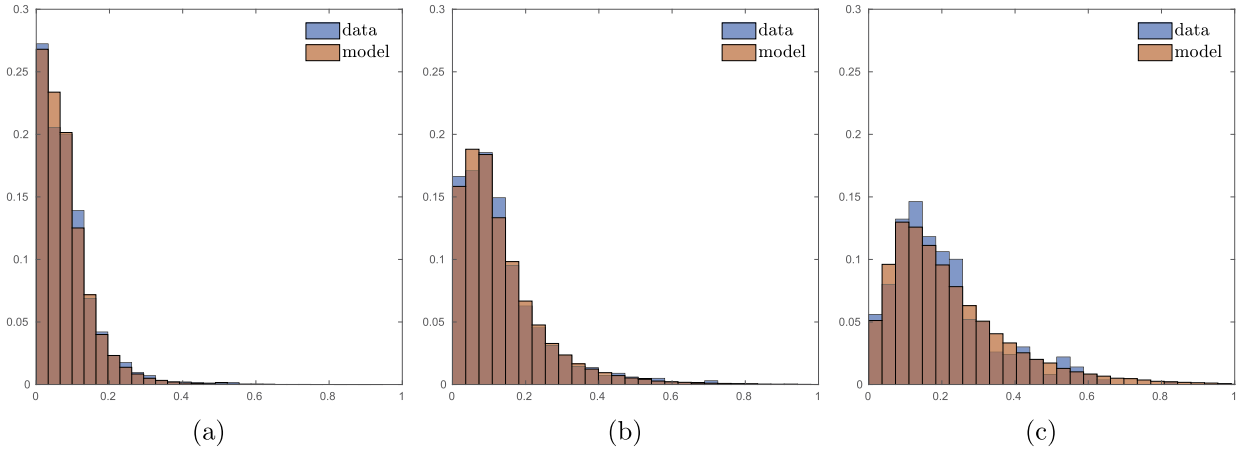


Fig. 7. Histograms showing the cross-sectional distributions of frailty for 35–36-year-olds (a), 55–56-year-olds (b), and 75–76-year-olds (c) in the model and in the data.

Consistent with the mean log frailty profiles presented in the figures, the estimates in Table 9 show that education has a much larger impact on frailty than gender does. Controlling for age and education, being male instead of female reduces frailty by only 3 percent. In contrast, controlling for age and gender, being a high school graduate as opposed to not having a high school degree reduces frailty by 11.6 percent. Interestingly, being a college graduate nearly triples the percentage decline, reducing frailty by 33.9 percent.

The right side of Table 9 presents the estimates of the parameters governing the stochastic component of the log frailty process. The parameters are identified by the structure of the empirical autocovariance age profiles. Fig. 6c shows these profiles together with their model counterparts. Notice that the moments generated by the model have the same properties as those in the data. The variance profile is hump-shaped. The autocovariance profiles decline in the lag length, and the rate at which they do so is increasing in age.

The parameter estimates indicate that log frailty innovations are highly persistent, but not permanent, with an autocorrelation of 0.99. This result implies that controlling for age, less frail individuals accumulate deficits at a higher rate than frailer ones. The autocorrelation coefficient is identified off both the curvature of the variance age-profile and the rate of decline of the autocovariance profile, both of which are also affected by mortality. Notice that even at younger ages when mortality effects are relatively small, the empirical autocovariance profiles in Fig. 6c decline push down the estimated value of ρ.

The parameter estimates also indicate that both ex ante heterogeneity in log frailty at age 25 and frailty shocks during the life cycle are an important driver of cross-sectional variation in health. Controlling for gender and education, ex ante heterogeneity accounts for 70 percent of the overall variation in log frailty at age 25, with shocks accounting for the remaining 30 percent. However, the contribution of the shocks increases with age. By age 45, shocks are responsible for 73 percent of the overall variation, and by age 65, for 81 percent.

To illustrate the overall fit of the estimated frailty process to the data, Fig. 7 provides histograms showing cross-sectional distributions of frailty in the model and in the data for three age groups: 35–36-year-olds, 55–56-year-olds, and 75–76-year-olds. We have already shown that our estimated frailty process replicates well the targeted moments: the fraction of individuals with zero frailty, as well as the mean, variance, and autocovariances of log frailty by two-year age groups. Fig. 7 illustrates that, in addition, the model generates age-conditional cross-sectional distributions of frailty that are very similar to those observed in the data. In particular, the figure shows that the cross-sectional distributions of frailty generated by

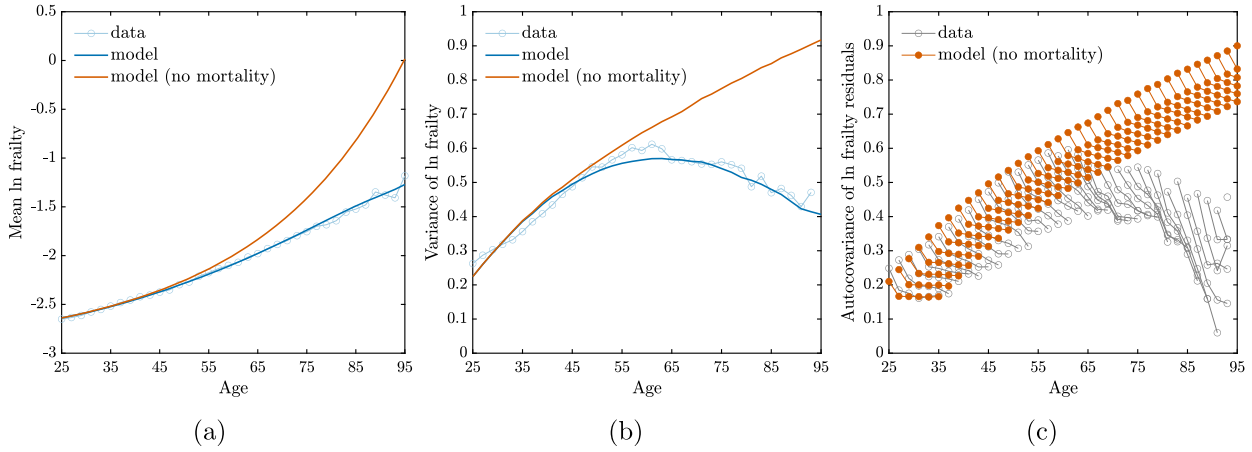


Fig. 8. Moments illustrating the effect of mortality. Panel (a) is mean log frailty by age in the data (open blue circles) and in the model with mortality (solid blue line) and without mortality (solid orange line). Panel (b) is the variance of log frailty by age in the data (open blue circles) and in the model with mortality (solid blue line) and without mortality (solid orange line). Panel (c) is autocovariance profiles of log frailty residuals by age in the data (open gray circles) and in the model without mortality (closed orange circles).

the model generally match well non-targeted higher-order moments of the distributions in the data, such as their skewness and kurtosis. While we show only three age groups in the figure, the fit of the model to the data is similar for the other groups (see Fig. 22 in the online appendix).

We have claimed above that mortality has a large impact on the targeted moments. Fig. 8 illustrates this point by showing the implications of assuming that no one ever dies. Panel (a) shows the impact of mortality on mean log frailty by age, panel (b) shows the impact of mortality on the variance of log frailty by age, and panel (c) shows the impact of mortality on the autocovariance profiles. Notice that the effects of mortality on all the moments are negligible until about age 50 but quickly become sizable at later ages. For instance, as panels (a) and (b) show, due to their high correlation, mortality reduces both the mean and variance of frailty after age 50 at a rate that is increasing with age.

Panel (c) of Fig. 8 shows that mortality is the reason that the model is able to generate autocovariance profiles that steepen over the life cycle at an increasing rate. Indeed, the autocovariance profiles of log frailty absent mortality have the opposite feature of those in panel (c) of Fig. 6. They still steepen with age but the rate at which they do declines with age instead of rising with it.[51] In the baseline model with mortality risk, individuals whose frailty shocks are more persistent are more likely to die and, as a result, are less likely to contribute to higher-order autocovariances conditional on survival. This selection bias becomes more severe as the lag length increases, generating downward pressure on the autocovariance profile. The selection bias also becomes more severe as individuals age, become frailer, and die at higher rates. These findings indicate that ignoring mortality would lead substantially underestimating not only both the mean and variance of frailty at later ages but also its persistence.[52]

4.4. Estimation results for gender and education subgroups

In this subsection, we present results from estimating the frailty process separately for men and women, and for each of the three education groups: less than high school degree, high school graduates, and college graduates. For each subgroup, the frailty process estimation is done exactly as the baseline estimation. In particular, the survival probabilities are based on the probit estimation of equation (7) (see Table 89 in the online appendix), and each subgroup's probability of nonzero frailty is determined by equation (6) and the probit regression results in Table 8. Then, for each subgroup, counterparts to the baseline empirical moments are constructed from the data. The SMM procedure finds the vector of parameters θ for each subgroup that minimizes the distance between these moments and their model counterparts. Just as before, this is done by solving the minimization problem in equation (8) using a diagonal weighting matrix consisting of the inverse of the estimated variances of the moments.

Consider first the findings from estimating the model separately for men and women. Columns (1) and (2) of Table 10 show the estimation results for each gender, and Fig. 9 shows the overall fits of the models to their empirical targets. Panels

[51] Given the specification of the log frailty shock process, the rate of decrease of the autocovariance profiles with age must be proportional to the rate of increase of the variance of the persistent shock. With $\rho < 1$, the variance of the persistent shock is increasing at a decreasing rate and, hence, the autocovariance profiles are decreasing at a decreasing rate.

[52] The persistence of the log frailty innovations is identified off the curvature of the variance age-profile and the rate of decline of the autocovariance profile (both of which increase when ρ declines). If the data was targeted without accounting for selection due to mortality, the hump-shaped variance age-profile and steeper autocovariance profiles would result in a much lower value of ρ.

Table 10
Results from estimating the log frailty process via SMM separately for each gender and separately for each education group. Gender and education subsamples are constructed from the baseline PSID sample. Standard errors in parentheses.

		Men (1)	Women (2)	No HS degree (3)	HS degree (4)	College degree (5)
Deterministic component	age*	0.996	0.964	1.215	1.330	0.726
		(0.014)	(0.015)	(0.032)	(0.021)	(0.027)
	age*^{2}	2.000	1.875	2.520	0.820	2.320
		(0.036)	(0.045)	(0.119)	(0.065)	(0.099)
	age*^{3}	−0.615	−0.852	−0.403	−2.909	−0.803
		(0.064)	(0.058)	(0.221)	(0.078)	(0.191)
	age*^{4}	4.923	4.837	2.591	11.427	3.495
		(0.056)	(0.053)	(0.175)	(0.056)	(0.156)
	HS	−0.083	−0.194			
		(0.002)	(0.001)			
	college	−0.317	−0.390			
		(0.001)	(0.001)			
	male			−0.111	−0.019	−0.035
				(0.003)	(0.001)	(0.002)
	const.	−2.541	−2.373	−2.447	−2.590	−2.775
		(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
Stochastic component	σ_u^2	0.040	0.039	0.038	0.044	0.028
		(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
	σ_ε^2	0.018	0.022	0.029	0.015	0.026
		(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
	σ_α^2	0.092	0.176	0.153	0.145	0.129
		(0.005)	(0.005)	(0.010)	(0.004)	(0.005)
	ρ	0.992	0.987	0.984	1.001	0.963
		(0.001)	(0.001)	(0.002)	(0.001)	(0.002)

Note: age* $\equiv$ (age − 25)/100. Stochastic component estimates are annual.

(a) and (d) of the figure show the fraction of men and women with zero frailty by age in the data, predicted by the probit regression, and predicted from the model simulations. Similarly to the findings for the whole sample, the impact of ignoring mortality bias on the probabilities is small overall. Panels (b) and (e) of the figure show the age profiles of mean log frailty by education for each gender in the model and in the data. Notice that, controlling for age, there is more variation in log frailty across education groups for women than for men. For both genders, the impact on frailty of having a college degree relative to having only a high school degree is similar. However, women without a high school degree have much higher frailty, on average, relative to both men without a high school degree and more educated women. This feature of the data is reflected in the estimation results provided in the table. For women, being a high school graduate reduces frailty by 19.4 percent relative to not having a high school degree. For men, the effect is less than half that size. Specifically, relative to not having a high school degree, being a high school graduate reduces frailty by only 8.3 percent.

Panels (c) and (f) of Fig. 9 show the autocovariance profiles of the log frailty residuals for men and women separately in both the model and data. While the profiles are overall fairly similar, there are three main differences between them that indicate slightly different life cycle frailty dynamics. First, the level of the variation in the residuals is higher for women than for men at all ages. This is why, as Table 10 shows, the estimated variance of the fixed effect, σ_α^2, is higher for women. Second, women's variance profile increases slightly more rapidly before age 55, and their autocovariance profiles decline more rapidly with lag length, as compared to men's, leading to a larger estimated variance of the persistent shock, σ_ε^2, for women. Finally, women's variance profile has more curvature. While much of the curvature differences are driven by differences in mortality, according to the estimation results, men's frailty shocks are also slightly more persistent. That is, the value of ρ for men is 0.992, compared to 0.987 for women.

The results from estimating the model separately for each of the three education groups are presented in columns (3) through (5) of Table 10. Plots illustrating the fit of each estimated model to the data are presented in Fig. 10. In the left column, panels (a), (d), and (e) show the fraction with zero frailty for each education group. In the middle column, panels (b), (e), and (h) show mean log frailty by age and gender. In the right column, panels (c), (f), and (i) present the autocovariance profiles. Overall, there are larger differences in the target moments and estimation results across education groups than there are across genders. Focusing on the middle column in the figure, notice that the effects of gender on the age profiles of mean log frailty are significantly larger in the less than high school degree group. According to the table, among those without a high school degree, men's frailty is 11.1 percent lower than women's on average. In contrast, men's frailty is only 1.9 percent and 3.5 percent lower than women's among high school and college graduates, respectively. This result is consistent with the finding above that the effect on frailty of not having a high school degree is much larger for women than for men. It is unclear to what extent this finding is due to truly larger variation in health by gender among those without a high school degree versus differential awareness of or reporting of health conditions by less educated men as compared to less educated women.

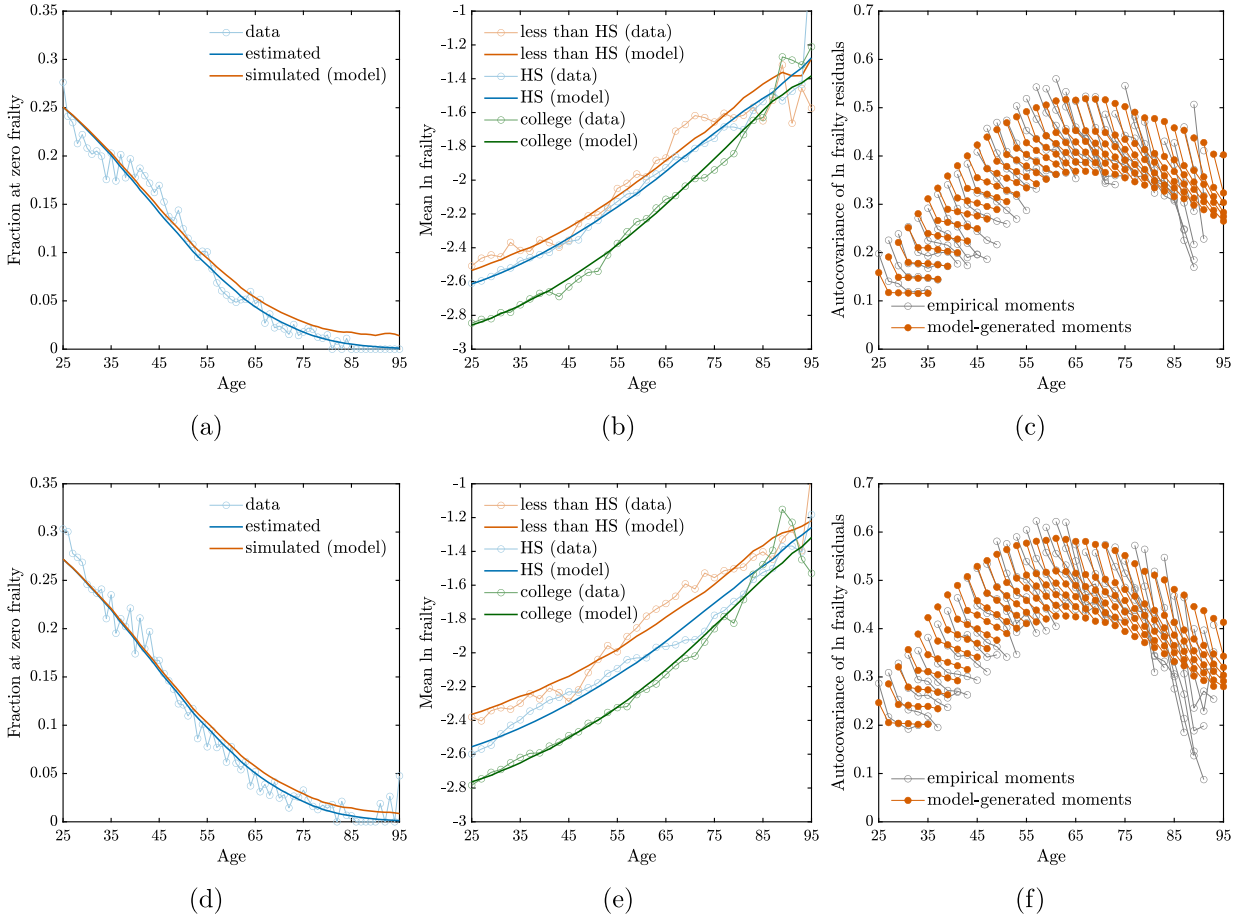


Fig. 9. Fit of the model estimation separately for men and women. Fraction of men (a) and women (b) with zero frailty by age in the data (blue open circles), predicted from the probit regression (solid blue line), and predicted, after accounting for mortality, by the model (solid orange line). Mean log frailty for men (b) and women (e) without a high school degree (orange), with only a high school degree (blue), and with a college degree (green) by age in the data (open circles) and in the model (solid lines). Autocovariance profiles of log frailty residuals for men (c) and women (f) by age in the data (open gray circles) and in the model (closed orange circles).

The autocovariance profiles of each education group's log frailty residuals shown in the last column of Fig. 10 suggest that the dynamics of frailty vary considerably by education group. First, the level of variation in the residuals at age 25 is decreasing in education. The estimates of the variance of the fixed effect, σ_α^2, are also declining in education, reflecting this feature of the data moments. Second, the shapes of the variance age profiles and autocovariance profiles differ by education in a non-monotonic way. The autocovariance profiles decline less rapidly with lag length in the high school graduate group as compared to the less than high school and college graduate groups. The variance profile for high school graduates also exhibits less curvature. These two features of the moments lead to a relatively lower estimate of the variance, σ_ε^2, and a relatively higher estimate of the autocorrelation, ρ, of the persistent shock for the high school graduate group. Interestingly, these features also indicate that shocks to the frailty of high school graduates are permanent with an effect that slowly grows larger over time.

The estimation results for the high school and college graduate groups suggest that the persistence of health shocks may be declining in education. While the estimation results from the less than high school degree group are not consistent with this result, the findings for this group should be treated cautiously as its moments are measured with the least precision. This is because it is the smallest group, with about half the number of people and observations of the college graduate group and about a third of the high school graduate group.[53]

[53] Specifically, the less than high school degree group consists of 2,767 unique individuals and 11,767 observations, while the college graduates group consists of 5,383 individuals and 26,755 observations, and the high school graduate group consists of 9,176 individuals and 43,403 observations.

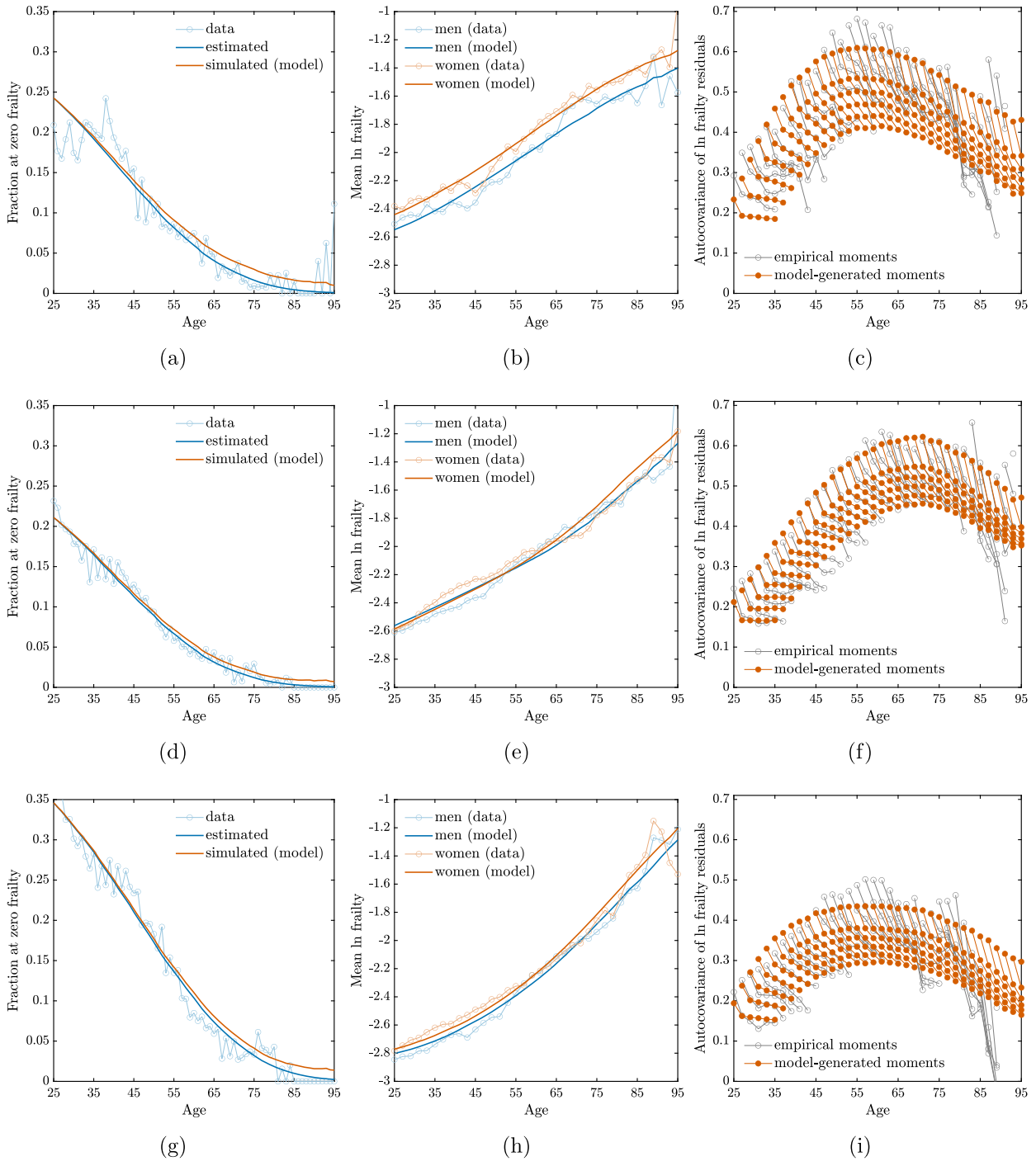


Fig. 10. Fit of estimated model by education: less than high school degree (top row), high school graduates (middle row), and college graduates (bottom row). Panels (a), (d), and (g) are fraction with zero frailty by age in the data (blue open circles), predicted from the probit regression (solid blue line), and predicted, accounting for mortality, by the model (solid orange line). Panels (b), (e), and (h) are mean log frailty by age in the data (open circles) and in the model (solid lines) for women (orange) and men (blue). Panels (c), (f), and (i) are autocovariances of log frailty residuals by age in the data (open gray circles) and in the model (closed orange circles).

4.5. Estimation results using other variants of frailty

In Section E of the online appendix we report results from estimating the statistical model in Section 4.1 using the other three variants of frailty presented in Section 3.2. The results from estimating the model using frailty with SRHS are

extremely similar to the baseline estimation results. The only notable difference is that when SRHS is added as a deficit variable, the average effect on frailty of being male is even smaller than in the baseline.

However, the results from estimating the model using the FPC and SIO indices have important quantitative differences from those using frailty. Recall from Section 3.2 that even though the FPC and SIO index are qualitatively similar to frailty and perform similarly well at accounting for variation in health-related outcomes, quantitatively they exhibit several differences. In particular, there is more variation at later ages in the FPC and SIO index than in frailty, and their means increase relatively faster with age. These differences are reflected in both the empirical moments targeted in the SMM estimation and the estimation results. The estimated age profiles of both indices are steeper and the overall variance of the stochastic component is larger. Both estimations also attribute a larger share of variation to ex ante heterogeneity especially at older ages. Controlling for gender and education, in the FPC index, ex ante heterogeneity accounts for 77 percent of the variation at age 25, 44 percent at age 45, and 42 percent at age 65. Similarly, in the SIO index, ex ante heterogeneity accounts for 72 percent, 41 percent, and 38 percent, respectively. Recall that in frailty, in contrast, the fractions are 70 percent, 27 percent, and 19 percent.

These findings are interesting and warrant further investigation in future work. On the one hand, they indicate that the choice of index may matter for measuring the overall level of residual variation in health and the rate at which health uncertainty is resolved with age. However, on the other hand, our findings in Section 3.2 indicate that these differences in the dynamics of the various health measures do not translate into significant differences in their ability to predict several health-related outcomes. Thus, the extent to which they capture meaningful differences in underlying health dynamics and their economic implications is unclear.

5. Conclusion

We present the frailty index and use it to document several facts about health dynamics over the life cycle. Frailty increases with age, declines with education, and is slightly higher, on average, for women than for men. Dispersion in frailty increases with age, and the cross-sectional distribution of frailty is highly skewed. Also, frailty innovations are highly persistent. We find that these facts are robust to alternative approaches to constructing the frailty index inspired by measures of health used in the literature. We also document several differences between frailty and self-reported health status. In particular, we find that self-reported health status indicates a slower rate of deterioration of health with age as compared to frailty. In addition, we find that frailty (and its variants) outperforms the use of self-reported health status alone in predicting many health-related outcomes.

We then show that frailty dynamics over the life cycle can be well captured by a statistical model featuring a shock process governed by AR(1) and purely transitory components, a structure that can easily be embedded into life cycle models. We estimate the model using data from the Panel Study of Income Dynamics first for all the individuals in the study and then separately by gender and education subgroups. The estimation is done via simulated method of moments and accounts for sample selection bias due to mortality. Our frailty measure and estimated stochastic process can be used to further study life cycle health dynamics and their economic implications.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.red.2021.07.001.

References

Abowd, J.M., Card, D., 1989. On the covariance structure of earnings and hours changes. Econometrica 57, 411–445.
Amengual, D., Bueren, J., Crego, J.A., 2017. Endogenous health groups and heterogeneous dynamics of the elderly. Working Paper.
Ameriks, J., Briggs, J., Caplin, A., Shapiro, M.D., Tonetti, C., 2020. Long-term-care utility and late-in-life saving. Journal of Political Economy 128, 2375–2451.
Baker, M., Solon, G., 2003. Earnings dynamics and inequality among Canadian men, 1976–1992: evidence from longitudinal income tax records. Journal of Labor Economics 21, 289–321.
Baker, M., Stabile, M., Deri, C., 2004. What do self-reported, objective, measures of health measure? The Journal of Human Resources 39, 1067–1093.
Bell, F.C., Miller, M.L., 2005. Life Tables for the United States Social Security Area. Social Security Administration Publications.
Benítez-Silva, H., Buchinsky, M., Man Chan, H., Cheidvasser, S., Rust, J., 2004. How large is the bias in self-reported disability? Journal of Applied Econometrics 19, 649–670.
Blundell, R., Dias, M.C., Britton, J., French, E., 2020. The impact of health on labor supply near retirement. The Journal of Human Resources, 1217–9240R4.
Blundell, R.W., Britton, J., Costa Dias, M., French, E., 2016. The Dynamic Effects of Health on the Employment of Older Workers. Michigan Retirement Research Center Research Paper.
Bound, J., 1991. Self-reported versus objective measures of health in retirement models. The Journal of Human Resources 26, 106–138.
Bound, J., Schoenbaum, M., Stinebrickner, T.R., Waidmann, T., 1999. The dynamic effects of health on the labor force transitions of older workers. Labour Economics 6, 179–202.
Braun, R.A., Kopecky, K.A., Koreshkova, T., 2017. Old, sick, alone, and poor: a welfare analysis of old-age social insurance programmes. The Review of Economic Studies 84, 580–612.
Braun, R.A., Kopecky, K.A., Koreshkova, T., 2019. Old, frail, and uninsured: accounting for features of the US long-term care insurance market. Econometrica 87, 981–1019.
Butler, J.S., Burkhauser, R.V., Mitchell, J.M., Pincus, T.P., 1987. Measurement error in self-reported health variables. Review of Economics and Statistics, 644–650.
Capatina, E., 2015. Life-cycle effects of health risk. Journal of Monetary Economics 74, 67–88.

Cole, H.L., Kim, S., Krueger, D., 2019. 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, 1123–1169.
Crossley, T.F., Kennedy, S., 2002. The reliability of self-assessed health status. Journal of Health Economics 21, 643–658.
Currie, J., Madrian, B.C., 1999. Health, health insurance and the labor market. In: Handbook of Labor Economics, vol. 3, pp. 3309–3416.
Dalgaard, C.-J., Strulik, H., 2014. Optimal aging and death: understanding the Preston curve. Journal of the European Economic Association 12, 672–701.
De Nardi, M., French, E., Jones, J.B., 2010. Why do the elderly save? The role of medical expenses. Journal of Political Economy 118, 37–75.
De Nardi, M., Pashchenko, S., Porapakkarm, P., 2018. The lifetime costs of bad health. Working Paper 23963. National Bureau of Economic Research.
Deaton, A., Paxson, C., 1994. Intertemporal choice and inequality. Journal of Political Economy 102, 437–467.
Disney, R., Emmerson, C., Wakefield, M., 2006. Ill health and retirement in Britain: a panel data-based analysis. Journal of Health Economics 25, 621–649.
Dowd, J.B., Todd, M., 2011. Does self-reported health bias the measurement of health inequalities in US adults? Evidence using anchoring vignettes from the Health and Retirement Study. The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences 66, 478–489.
Duffie, D., Singleton, K.J., 1993. Simulated moments estimation of Markov models of asset prices. Econometrica 61, 929–952.
French, E., Jones, J.B., 2004. On the distribution and dynamics of health care costs. Journal of Applied Econometrics 19, 705–721.
French, E., Jones, J.B., 2011. The effects of health insurance and self-insurance on retirement behavior. Econometrica 79, 693–732.
French, E., Jones, J.B., 2017. Health, health insurance, and retirement: a survey. Annual Review of Economics 9, 383–409.
Friedberg, L., Hou, W., Sun, W., Webb, A., Li, Z., 2014. New evidence on the risk of requiring long-term care. Boston College Center for Retirement Research WP 12.
Gilleskie, D., Han, E., Norton, E., 2017. Disentangling the contemporaneous and dynamic effects of human and health capital on wages over the life cycle. Review of Economic Dynamics 25, 350–383.
Goggins, W.B., Woo, J., Sham, A., Ho, S.C., 2005. Frailty index as a measure of biological age in a Chinese population. The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences 60, 1046–1051.
Guvenen, F., 2009. An empirical investigation of labor income processes. Review of Economic Dynamics 12, 58–79.
Hosseini, R., Kopecky, K.A., Zhao, K., 2021. How important is health inequality for lifetime earnings inequality? Working papers series 2021-1. Federal Reserve Bank of Atlanta.
Hubbard, R.G., Skinner, J., Zeldes, S.P., 1995. Precautionary savings and social insurance. Journal of Political Economy 103, 360–399.
Idler, E.L., Benyamini, Y., 1997. Self-rated health and mortality: a review of twenty-seven community studies. Journal of Health and Social Behavior, 21–37.
Karahan, F., Ozkan, S., 2013. On the persistence of income shocks over the life cycle: evidence, theory, and implications. Review of Economic Dynamics 16, 452–476.
Kitao, S., 2014. A life-cycle model of unemployment and disability insurance. Journal of Monetary Economics 68, 1–18.
Kopecky, K., Koreshkova, T., 2014. The impact of medical and nursing home expenses on savings. American Economic Journal: Macroeconomics 6, 29–72.
Kulminski, A., Yashin, A., Arbeev, K., Akushevich, I., Ukraintseva, S., Land, K., Manton, K., 2007a. 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, 250–258.
Kulminski, A.M., Ukraintseva, S.V., Akushevich, I.V., Arbeev, K.G., Yashin, A.I., 2007b. Cumulative index of health deficiencies as a characteristic of long life. Journal of the American Geriatrics Society 55, 935–940.
Lockwood, L.M., 2018. Incidental bequests and the choice to self-insure late-life risks. The American Economic Review 108, 2513–2550.
McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), Frontiers in Econometrics. New York. Academic Press.
McFadden, D., 1989. A method of simulated moments for estimation of discrete response models without numerical integration. Econometrica 57, 995–1026.
Mitnitski, A., Bao, L., Rockwood, K., 2006. Going from bad to worse: a stochastic model of transitions in deficit accumulation, in relation to mortality. Mechanisms of Ageing and Development 127, 490–493.
Mitnitski, A., Song, X., Skoog, I., Broe, G.A., Cox, J.L., Grunfeld, E., Rockwood, K., 2005. Relative fitness and frailty of elderly men and women in developed countries and their relationship with mortality. Journal of the American Geriatrics Society 53, 2184–2189.
Mitnitski, A.B., Graham, J.E., Mogilner, A.J., Rockwood, K., 2002. Frailty, fitness and late-life mortality in relation to chronological and biological age. BMC Geriatrics 2, 1.
Mitnitski, A.B., Mogilner, A.J., Rockwood, K., 2001. Accumulation of deficits as a proxy measure of aging. The Scientific World Journal 1, 323–336.
Mitnitski, A.B., Song, X., Rockwood, K., 2004. 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, M627–M632.
O'Donnell, O., Van Doorslaer, E., Van Ourti, T., 2015. Health and inequality. In: Handbook of Income Distribution, vol. 2. Elsevier, pp. 1419–1533.
Ozkan, S., 2017. Preventive vs. curative medicine: a macroeconomic analysis of health care over the life cycle. Working Paper. University of Toronto.
Pakes, A., Pollard, D., 1989. Simulation and the asymptotics of optimization estimators. Econometrica 57, 1027–1057.
Poterba, J.M., Venti, S.F., Wise, D.A., 2017. The asset cost of poor health. The Journal of the Economics of Ageing 9, 172–184.
Robinson, J., 1996. A long-term care status transition model. In: Bowles Symposium on the Old-Age Crisis: Actuarial Opportunities. Georgia State University.
Rockwood, K., Andrew, M., Mitnitski, A., 2007. A comparison of two approaches to measuring frailty in elderly people. The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences 62, 738–743.
Rockwood, K., Mitnitski, A., 2007. Frailty in relation to the accumulation of deficits. The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences 62, 722–727.
Salomon, J.A., Tandon, A., Murray, C.J., 2004. Comparability of self rated health: cross sectional multi-country survey using anchoring vignettes. BMJ. British Medical Journal 328, 258.
Schünemann, J., Strulik, H., Trimborn, T., 2017a. The gender gap in mortality: how much is explained by behavior? Journal of Health Economics 54, 79–90.
Schünemann, J., Strulik, H., Trimborn, T., 2017b. Going from bad to worse: adaptation to poor health spending, longevity, and the value of life. Journal of Economic Behavior & Organization 140, 130–146.
Searle, S., Mitnitski, A., Gahbauer, E., Gill, T., Rockwood, K., 2008. A standard procedure for creating a frailty index. BMC Geriatrics 8, 1.
Stern, S., 1989. Measuring the effect of disability on labor force participation. The Journal of Human Resources, 361–395.
Storesletten, K., Telmer, C.I., Yaron, A., 2004. Consumption and risk sharing over the life cycle. Journal of Monetary Economics 51, 609–633.
Van Doorsaler, E., Gerdtham, U., 2002. Does inequality in self-assessed health predict inequality in survival by income? - Evidence from Swedish data. Social Science & Medicine 57, 1621–1629.
Woo, J., Goggins, W., Sham, A., Ho, S., 2005. Social determinants of frailty. Gerontology 51, 402–408.
Yang, Y., Lee, L.C., 2010. 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, 246–255.