The welfare effects of nonlinear health dynamics*

Chiara Dal Bianco‡ Andrea Moro§

October 16, 2022

Abstract

We generate a continuous measure of health to estimate a non-parametric model of health dynamics, showing that adverse health shocks are highly persistent when suffered by people in poor health, a pattern that cannot be accounted for by canonical models. We incorporate this health measure into a life-cycle model of consumption, savings, and labor force participation. After estimating the model parameters, we simulate the effects of health shocks on economic outcomes. We find that bad health shocks have persistent adverse economic effects that are more dramatic for poor individuals starting in bad health. Bad health shocks also increase the dispersion of asset accumulation within this category of individuals. A canonical model of health dynamics would not uncover these effects.

Keywords: Health, Nonlinear dynamics, Life-cycle model, Inequality, Savings
JEL codes: I14, D15, J22 ,J26

*We thank Marco Bertoni and Guglielmo Weber for helpful suggestions and Stéphane Bonhomme for advice with estimating the nonlinear health process.

‡Department of Economics and Management, University of Padova, Italy. ORCID: 0000-0002-5449-0645. *E-mail*: chiara.dalbianco@unipd.it

§Department of Economics, Vanderbilt University, Nashville, TN, USA. ORCID: 0000-0001-5570-8151 *E-mail*: andrea@andreamoro.net

1 Introduction

Are severe health shocks more persistent than milder shocks? If so, what are their effects on individuals' economic outcomes? To answer these questions, we introduce a complex health measure and estimate its persistence using a non-parametric model with nonlinear dynamics. We incorporate this health measure into a life-cycle model of consumption, savings, and labor supply decisions to simulate the effects of various health shocks.

Several articles have studied the effects of health on life-cycle economic outcomes, helping our understanding of the complex interactions between health and economics.[1] However, most of these studies posit a binary or 3-valued health measure evolving according to a first-order Markov process.[2] Our analysis focuses on assessing the effects of especially severe and persistent health shocks, requiring a more detailed measure of health and complex dynamics.

We derive a continuous measure of health from several objective health indicators using a latent variable model in the spirit of Bound (1991).[3] We assume this health measure evolves according to a complex, nonlinear dynamic process with persistent and transitory components. Drawing from Arellano et al. (2017), we assume the former component evolves according to a non-parametric process with nonlinear persistence and age dependence. The persistence patterns resulting from this flexible approach could not be accounted for by a "canonical" first-order Markov process.

We incorporate this health measure into a standard life-cycle model where agents receive utility from consumption, leisure, and bequest of assets; they choose, every period, labor market participation, hours worked, and savings facing uncertainty over health, wages, and life expectancy. Health affects wages and life expectancy. Following the literature, health affects the time constraint, changing the tradeoff between consumption and leisure. The model also includes a fixed cost of work to fit the evidence on labor market participation and hours worked.

We estimate this model with a Simulated Method of Moments procedure using data from the English Longitudinal Study of the Ageing (ELSA), a panel of English households aged 50 and above. Focusing on English data simplifies our analysis, allowing us to exclude medical expenditure choices from the model (health care in the U.K. is publicly provided). We perform the estimation in two steps. First, we

[1] See, for example, French (2005), De Nardi et al. (2010) Capatina (2015), Low and Pistaferri (2015), Ameriks et al. (2020), Blundell et al. (2021) and the surveys De Nardi et al. (2016) and French and Jones (2017).

[2] Exceptions are De Nardi et al. (2017), modeling a *binary* measure of health with a second-order Markov process with age-dependent parameters, showing that health follows complex patterns of duration-dependence; Dal Bianco (2022), and Hosseini et al. (2021b) compute a multi-valued measure of health but model its dynamic with a canonical process.

[3] See Hosseini et al. (2021a) for an alternative approach based on a *frailty index*, computed as the cumulative sum of all adverse health indicators, which they compare with the dynamics of a self-reported health indicator.

estimate the parameters of the health, wage, and life expectancy processes, which we assume are exogenous to the life-cycle model. Next, we estimate the model parameters (the coefficient of risk aversion, the preferences for bequests, the costs of bad health, and of working) using the processes for wages, health, and life expectancy estimated in the first step. To highlight the importance of using a nonlinear process, we also estimate a model specification using a "canonical" health process whose persistent component evolves according to an AR(1) model.

The estimates of the nonlinear health process reveal that health indeed follows a complex, nonlinear dynamics. In particular, negative shocks are worse and much more persistent than good shocks when starting in bad health. When individuals start in good health instead, good shocks are only slightly more persistent. These asymmetries cannot be replicated with the canonical process by construction. They are also difficult to ascertain when health is a binary variable or when it assumes few discrete values because they depend on the arbitrary health cutoffs used implicitly to discretize the health measure.

The moments we match to estimate the model are the age profiles of accumulated assets, of labor force participation by health quantiles, and of hours worked. Our estimates of the life-cycle model match these moments very well both when using the nonlinear and the canonical model of health dynamics.

Using the estimated life-cycle model, we simulate counterfactual health shocks to understand the role and distributional consequences of bad health and its persistence. Specifically, we simulate individuals who start at the age of 51 with different levels of the persistent component of the health process (the 10th, 50th, and 90th percentile of its distribution). We impose to the persistent component of health a good, an intermediate, and a bad shock at the age of 52 (respectively, shocks that bring the persistent component of health to the 10th, 50th, and 90th percentile of its distribution conditional on its starting value).[4]

We find that the health asymmetries following different shocks in the nonlinear model produce considerable economic differences. Bad health shocks are worse and have more lasting effects when people start poor and in bad health than otherwise. For example, we find that for people starting at age 51 with £10,000 in assets and at the 10th percentile of the persistent component of health, a bad shock induces an accumulation for the survivors at age 85 of £33,000 less than after a median shock. Starting from the median of the persistent component of health, a bad shock induces an accumulation of only £14,000 less than after a median shock. Starting from £80,000 in assets, the same difference reduces to no more than £9,000 and depends little on the initial level of health. We also find that the canonical model does not predict these large differences between shocks for poor people.

Using the estimated value function, we compute that the willingness to pay to avoid a bad shock is about £20,000 for people starting with £10,000 in assets at the age of 51 and increases to £30000 for people starting with over £115,000 in assets: a

[4] Arellano et al. (2017) proposes a similar exercise when studying income dynamics.

bad shock is much costlier for the poor relative to their wealth.

We find that bad shocks generate higher dispersion in accumulated assets among people that start with poor wealth and health relative to people subject to better shocks. A decomposition exercise where we turn off the impact of health on the separate components of the model affected by health (wages, time endowment, and mortality) reveals that the effects of the three components are not additive (nonlinearities are important) and that the adverse effects of bad health are primarily due to its time cost.

Our results highlight the importance of accounting for a complex health measure, particularly when focusing on outcome inequality. First, we contribute to the literature by proposing a continuous measure of health and showing that its evolution follows a complex, nonlinear process. Second, we show that using this measure of health, adverse shocks are costlier, especially for poorer individuals, and generate higher outcome dispersion.

The rest of the paper proceeds as follows. The next section describes our data. Section 3 describes the construction of our health measure and the estimation of the health process. Section 4 describes the life-cycle model and its estimation. In Section 5 we simulate the effect of different health shocks on the level and the distribution of outcomes.

2 Data

We use data from the first seven waves (2002-2012) of the English Longitudinal Study of Ageing (ELSA), a biennial longitudinal survey representative of the English population aged 50 and above. ELSA contains detailed income and wealth information and several indicators of individuals' physical, mental, and cognitive health status.

We focus on male respondents. The final sample includes about 32,200 observations from 8,200 individuals. Labor supply participation decreases sharply with age, but a small fraction of respondents is still active after age 70. Hours worked and wages decrease substantially with age, especially after the state pension age (65 for men in the period we consider). Total wealth increases with age, as previously reported in the literature (see e.g. Blundell et al. (2016)).[5]

3 Health measurement and dynamics

We construct a continuous index of health using a latent variable model.[6] True health ψ_{it} is a linear function of a set of objective health indicators Z_{it} and an error term ξ:

$$\psi_{it} = Z'_{it}\alpha + \xi_{it}$$

[5] See External Appendix Table B.1 for more details.

[6] See Bound et al. (1999), Bound (1991), Bound et al. (2010), J. S. Butler and Pincus (2010)

We observe self-reported health, h^b_{it}, a binary indicator equal to 1 if the individual declares to be in good, very good, or excellent health, and 0 otherwise. We assume that h^b_{it} is determined by the latent variable h^*_{it}, equal to the sum of ψ_{it} and a reporting error μ_{it}:

$$\begin{aligned} h^*_{it} &= \psi_{it} + \mu_{it} \\ &= Z'_{it}\alpha + (\xi_{it} + \mu_{it}) \\ &= Z'_{it}\alpha + u_{it} \end{aligned} \tag{1}$$

The individual reports $h^b_{it} = 1$ if $h^*_{it} > 0$. We also assume that ψ_{it} and μ_{it} are uncorrelated. The composite error term u_{it} reflects differences in reporting behavior across individuals. The health index is equal to the (standardized) predicted value $h_{it} = \hat{h}^*_{it}$, according to (1). This procedure is analogous in spirit to using "objective" health information to instrument for the endogenous and error-ridden self-reported health measure (Bound et al. (1999)).

Z includes eyesight, hearing, mobility indicators, ADL (limitations with activities of daily living), IADL (limitations with instrumental activities of daily living), depression, heart diseases, other diseases (chronic lung disease, asthma, arthritis, ...), eye problems, incontinence, BMI and grip strength (an objective measure of muscular strength).

The resulting distribution of the health index is left-skewed, as shown in Figure 1(a).[7] The second and higher moments of the health shocks (defined as the difference in health between periods) change with age and also vary with the value of initial health (see External appendix figures B.1 and B.2 for details). This evidence suggests that health is a highly nonlinear process with some age dependence and justifies the adoption of a flexible specification.

Following Arellano et al. (2017), we define the health process as the sum of a persistent and a transitory component:[8]

$$h_{it} = \eta_{it} + \epsilon_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T \tag{2}$$

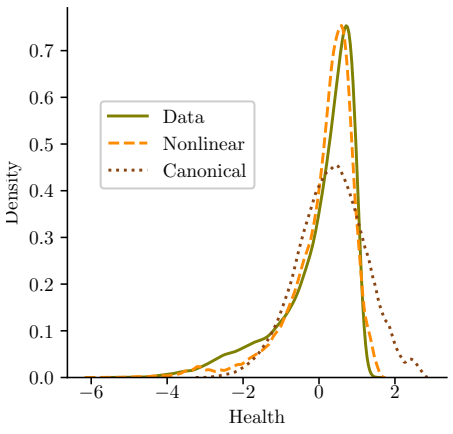
where η_{it} is the persistent component that follows a first-order Markov process, and ϵ_{it} is a transitory health shock with mean zero, independent of t and η. Note that from here on h_{it} denotes health residuals. We construct h_{it} as residuals from regressing the health index on a set of demographics, which includes a third-order polynomial in age, year of birth, education, and an indicator for having a partner. We include these controls to capture individual heterogeneity. Define $Q_t(\eta_{i,t-1}, \tau)$ the τ-th quantile of η_{it} given $\eta_{i,t-1}$, for $\tau \in (0, 1)$. The persistent component can be written as

$$\eta_{it} = Q_t(\eta_{i,t-1}, u_{it}) \qquad (u_{it}|\eta_{i,t-1}, \eta_{i,t-2,\ldots}) \sim \text{Uniform}(0, 1), \quad t > 1. \tag{3}$$

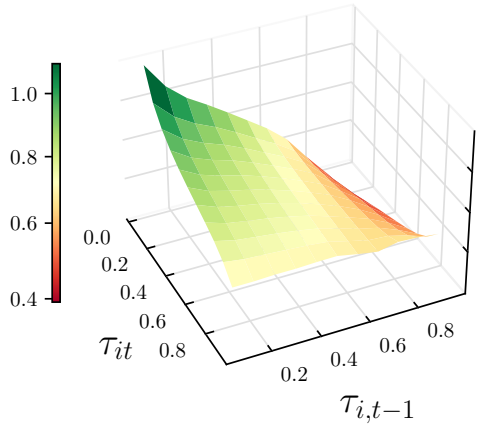
[7] Its shape is consistent with the one obtained by Hosseini et al. (2021a), defining health as the sum of adverse health indicators.

[8] Because our health data is biennial, for ease of notation in this section one time period is to be interpreted as two years. In External Appendix C we explain how we reconcile the biennial health process to the annual time span in the life-cycle model.

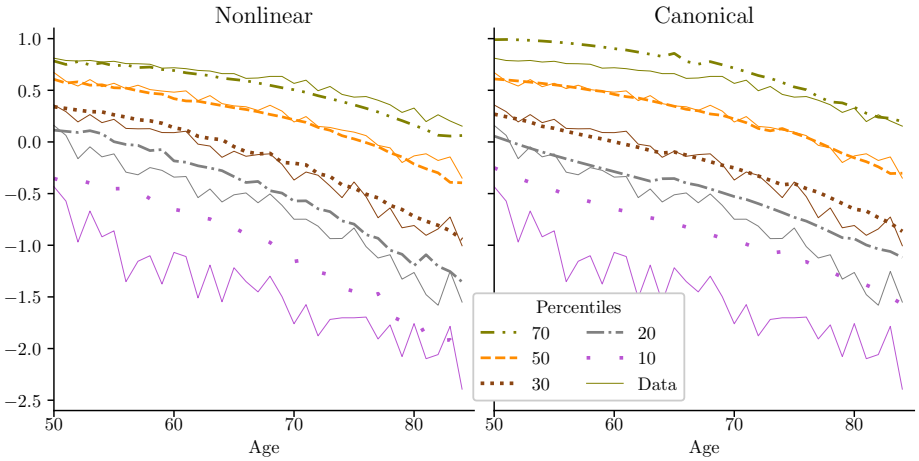
Figure 1: Estimates of the health processes



(a) Health index distribution



(b) Persistence, nonlinear model



(c) Health index, age profiles by health quantiles

Note: Panel (a) illustrates the distribution of the health index in the data (Equation (1), solid line), and in simulations using the nonlinear (dashed line) and canonical models (dotted line). Panel (b) reports estimates of (4) computed from simulations of the nonlinear model. Panel (c) illustrates average health conditional on survival at different quantiles of the health distribution. The thin solid lines represent values of the health index (Equation 1); the other lines represent averages from 15,000 simulated histories using the nonlinear model (left) and canonical model (right).

The $\eta's$ evolve over time non-parametrically allowing for nonlinear persistence. A special case of (3) is $\eta_{it} = \rho\eta_{i,t-1} + \nu_{it}$, an AR(1) process.[9], which we refer to as the "canonical" model, whereas we call the more general case the "nonlinear" model.

[9]See Blundell et al. (2021), Hosseini et al. (2021a) and Dal Bianco (2022).

Define a measure of the persistence of η when hit by a shock of rank τ as:

$$\rho_t(\eta_{i,t-1}, \tau) \equiv \frac{\partial Q_t(\eta_{i,t-1}, \tau)}{\partial \eta_{i,t-1}}. \tag{4}$$

In the canonical AR(1) model, ρ is constant, equal to the auto-regression parameter. The canonical model estimate of ρ is 0.953 (s.e. 0.016) (see External appendix C.2 for details).

To estimate the nonlinear process, we use the quantile-based panel data method developed by Arellano et al. (2017). Figure 1(b) reports the average value of ρ_t as a function of current and previous health deciles computed from simulations from the nonlinear model estimates.[10]. Health persistence is high when both the current and previous health shocks are low. Persistence is instead low when the current health shock is high and previous health is low. For individuals with low health, bad shocks are much more persistent than high shocks. Instead, for individuals starting with high health, persistence is slightly increasing.

Figure 1(c) illustrates that over the life cycle, the fit of the nonlinear model for the health index at different percentiles of the health distribution is only marginally better than the fit of the canonical model. However, in the life-cycle model we describe below the models' differences in persistence generate different economic outcomes.

4 The life-cycle model

Because of data limitations, we model the behavior of males who enter the model at age 50 ($t = 1$), may work until 70 but start receiving pension income at 65 (the state pension age in the UK in the period covered by our data). In each period t, the individual receives utility (U_t) from consumption (c_t) and leisure (l_t). When he dies, he values bequest according to a bequest function, b_t.

Subject to the time and budget constraints outlined below, the household head maximizes present discounted utility. Current utility depends on consumption and leisure according to the CRRA specification:

$$U(c_t, l_t) = \frac{1}{1-\nu}(c_t^{\gamma} l_t^{1-\gamma})^{1-\nu} \tag{5}$$

where ν is the relative risk aversion parameter and γ is the consumption weight.

Upon death, the agent also receives utility from bequests a according to:

$$b(a_t) = \phi_B \frac{(a_t + K)^{(1-\nu)\gamma}}{1-\nu}. \tag{6}$$

[10]The nonlinearity of the health persistence is well reproduced by the model, see External appendix Figure C.3(a)

Parameter ϕ_B measures the intensity of bequest motives, whereas K regulates the curvature of the bequest function and allows for a finite utility of zero bequests.

In each period, the agent chooses savings a_t, and, only up to the age of 70, labor supply s_t measured in hours worked. With a time endowment set to L, the agent faces a time cost of bad health which we model as a flexible function $\phi_h(\cdot)$ of health (see Appendix A for details). The time constraint is:

$$l_t = L - s_t - \phi_h(h_t) \tag{7}$$

Health is included in the model by discretizing the transitory and persistent components, and the transition matrices of the persistent component at each time t into N dimensional grids.[11] Health uncertainty affects utility through leisure, the wage offer (see below), and the survival probability.

We model log hourly wages (e) as the sum of a deterministic component, which depends on health and age, a stochastic one, specified as the sum of a persistent and a transitory shock. We assume that persistence in wages is captured by an AR(1) process and the transitory component captures measurement error only. The budget constraint is:

$$a_{t+1} = a_t + s_t \cdot e_t(1 - c_p \mathbb{1}(t < 15)) + r \cdot a_t - tax_t + tr_t - c_t - \phi_w(\mathbb{1}(s_t > 0, t)) \tag{8}$$

Where hours worked s_t are constrained to be zero after the age of 70, tax are taxes, r is the exogenous interest rate on accumulated assets and tr_t are transfers received in such a way that consumption c_t is greater or equal to the consumption floor $\underline{c}$, plus, starting from age 65, income from pension wealth p_t, annuitized at rate r_p. Finally, $\phi_w(\cdot)$ denotes a monetary cost of working.[12] Each agent is endowed with initial pension wealth p, to which he contributes a fraction c_p of his annual earnings.

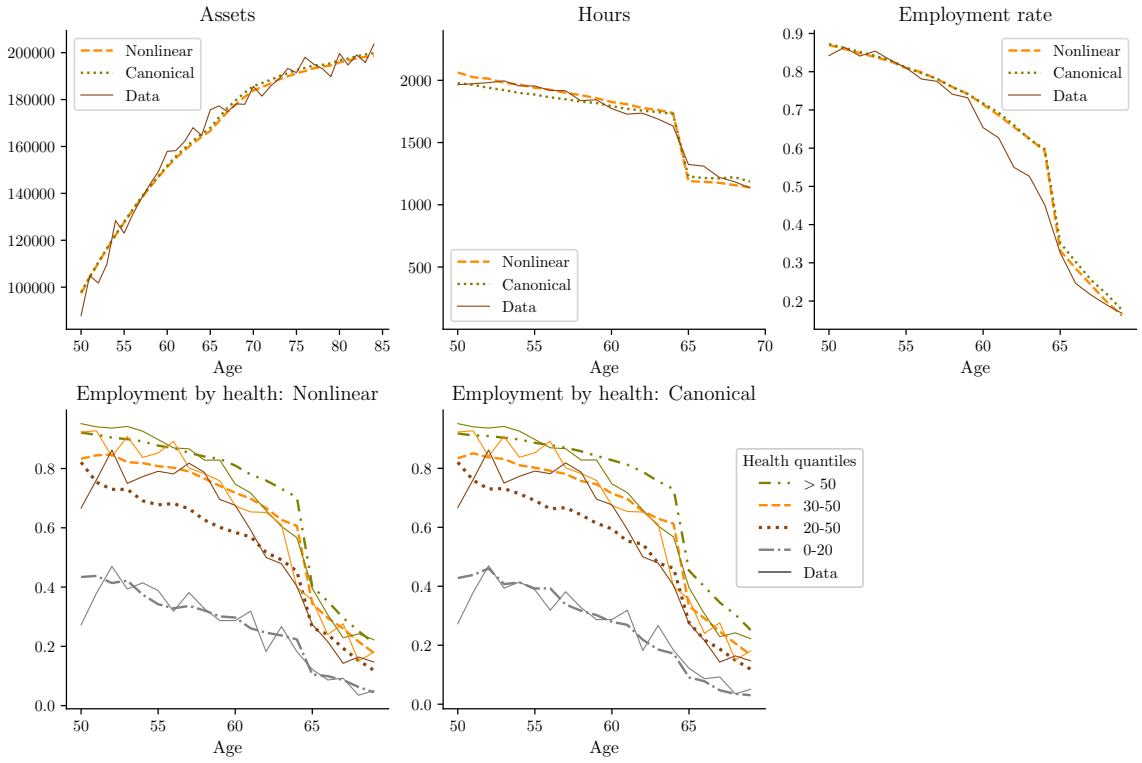
The probability of surviving to period $t+1$ conditional on being alive in period t is a function of age and health in period t, $\pi^{t+1} = \pi(h_t, t)$. We set the terminal age $T = 85$. Time is discounted at factor β. With these assumptions, preferences can be expressed using a recursive formulation with state-space be $X_t = \{a_t, e_t, h_t, p_t\}$, which we use to compute the numeric solution to the optimization problem. The expression for the Bellman equation is in Appendix A.

We estimate the life-cycle model with the nonlinear health dynamics and, for comparison, the canonical model dynamics. We use a Simulated Method of Moments, matching the age profiles of average assets, hours worked, and participation by four quantiles of the health distribution (using thresholds at percentiles 20, 30, and 50). The estimated parameters are in the Appendix. The fit of the nonlinear and the canonical model is very similar, despite the models producing different health persistence patterns, as shown in Section 3. Both match the targeted moments well (see Figure 2).

[11] De Nardi et al. (2020) use this procedure to incorporate the estimates of a nonlinear model of *earnings* dynamics into a life-cycle model. Additional details are in External Appendix C.

[12] We model the fixed cost of work as a monetary cost and not as a time cost (see French (2005)), to improve the fit of the hours worked age profile.

Figure 2: Estimation fit



Note: Age profiles of average assets, employment rate, and average hours worked conditional on being employed at the estimated parameters. Simulations of 15,000 histories, nonlinear and canonical models. The thin solid lines represent the age profiles in the data.

Our parsimonious specification, a linear function of the cost of work by age, cannot fully account for the steep decline in participation between ages 60 and 65. A psychic benefit from retiring starting at age 60 (justified by legislation allowing the withdrawal of private pension funds) may help improve matching these moments. We chose not to add this ad-hoc parameter since the model performs well overall.

5 The effects of health shocks

To assess the effect of a health shock and evaluate the impact of the nonlinear health dynamics, we simulate histories of individuals starting with the persistent component of the health process at percentiles 10, 50, and 90 of the health distribution at age 51 ($\tau_{init} = .1, .5, .9$), and we expose them to three shocks at age 52: a bad, medium, and good shock that bring them, respectively to the 10th 50th, and 90th percentiles of the health distribution conditional on τ_{init}. The resulting health dynamics of the good and bad shock, relative to the medium shock, are shown in Figure 3(a).

By construction, the canonical model results in symmetric health patterns after the good and bad shock, identical in magnitude with respect to τ_{init}.[13] The patterns resulting from the nonlinear health process, instead, are asymmetric and depend substantially on the initial value of the persistent component of the health process. In particular, the left panel of Figure 3(a) shows that a bad shock is worse and more persistent, relative to a median shock, when starting in bad health than when starting with medium or high health. Good shocks lead to higher levels of health when starting in bad health (relative to a median shock), but their magnitude is quantitatively smaller, in absolute value than bad shocks.

5.1 The nonlinear and canonical models produce different economic effects

These asymmetries have economic effects, as shown in Figure 3(b). When people start relatively poor, with £10,000 in assets at age 50, a bad shock induces an accumulation at the age of 85 of nearly £33,000 less than after a median shock. The difference in accumulation after a bad shock relative to a median shock is 43% this amount when starting at higher levels of health: £14,000 when $\tau_{init} = .5$ and less than £9,000 when $\tau_{init} = .9$. The gains after a good shock are smaller in magnitude, amounting at about £14,000, £3,000, and £10,000 respectively starting from $\tau_{init} = .1, .5$, and .9.

The losses from bad shocks are smaller and less asymmetric in the canonical model. They range from £16,000 to £24,000 at age 85. Even though the effect of shocks on health is theoretically identical across values of initial health, the effects of these shocks on asset accumulation are not the same because of the nonlinearities in the life-cycle model

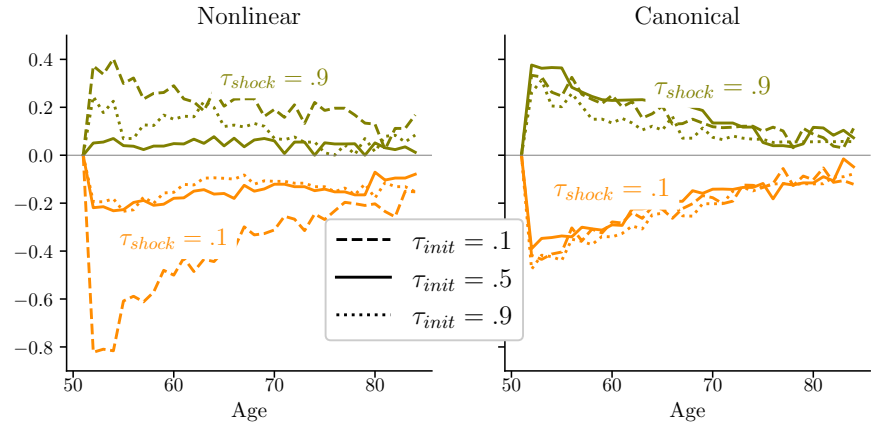
5.2 Bad health shocks are relatively worse for the poor

The economic effects of bad shocks are much weaker in absolute terms when people start off wealthier, as illustrated in Figure 3(c), who report results only for the nonlinear model with initial assets equal to £80,000, and £200,000. Differences in asset accumulation after receiving different shocks are smaller when people start with higher wealth. Patterns for participation and hours worked are illustrated in External Appendix figures F.4(a) and F.4(b).

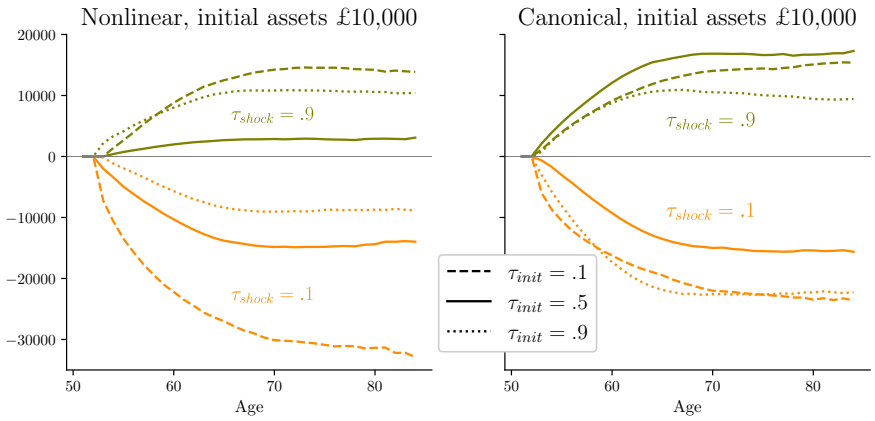
The more damaging effect of a bad health shock suffered by poorer individuals is better measured by the willingness to pay to avoid the health shock. We computed what wealth transfer at age 50 would make an individual indifferent between being subject to a bad shock ($\tau_{shock=0.1}$). For someone starting in relatively bad health ($\tau_{init} = .1$) and £10,000 in assets, such value is £20,000, and increases slightly for

[13] The discrepancies from this theoretical prediction in the right panel of Figure 3(a) (the three lines of each shock level should overlap) are due to approximation errors from the discretization of the health variable.

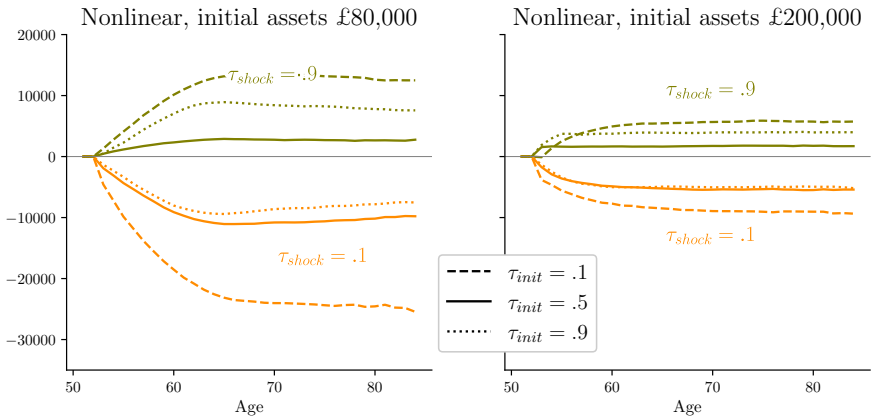
Figure 3: Effects of health shocks on health and assets by age



(a) Health after different health shocks



(b) Assets after health shocks for initially poor, nonlinear and canonical models



(c) Assets after health shocks for initially wealthy, nonlinear model

Note: The figures illustrate the age profiles of the difference in health (top two panels), or assets (bottom four panels) between individuals subject to a permanent component of health shock τ_{shock} ($\tau_{shock} = 0.9$ green lines and $\tau_{shock} = 0.1$ orange lines) and individuals subject to $\tau_{shock} = 0.5$, starting from different initial levels of the permanent component (τ_{init}). Averages over 5,000 simulated histories.

.

wealthier individuals, reaching £30,000 for individuals starting with over £115,000 in assets. As a fraction of initial assets, the willingness to pay is much higher for poorer individuals. The same value is never above £8,200 for people starting in good health (see Figure F.5 for details).

5.3 Sources of health effects

In the model, health affects economic outcomes through three channels: mortality, time available for leisure and work, and wages. To ascertain the relative importance of these channels we turn them off by assigning to all individuals the mortality, the time cost of bad health, and wages equal to the values these variables assume at a given quantile of the health distribution, conditional on age.

These effects generate both income effects (higher income and time available) and substitution effects (higher hourly wages), as well as dynamic effects from changes in mortality and uncertainty about future health. Removing the effect of health has very different effects on economic outcomes depending on the reference health level, as reported in Table 1(a).[14] When the effects of health through all three channels are removed, average assets are 8 percent higher if health were equal for everyone to the 75th percentile of its age-conditional distribution but 10 percent lower if health were set at the 25th percentile. Employment, hours, and income increase as well in the first scenario and decrease in the second.

The decomposition of these effects reveals that when people are assigned health at the 75th percentile, mortality alone accounts for 19% of the +8% total effect on assets at the time of death.[15]

This effect is driven by the need to finance a longer expected lifetime. The time cost is the factor contributing the most to all outcomes. When health increases, the reduced time cost shifts the static budget constraint outwards, inducing an increase in leisure, hours worked, and participation, making it less costly to save. The income effect induced by higher wages is different from the one induced by lower time cost because time available does not change.

As a consequence hours worked actually slightly decrease (the substitution effect dominates), and higher income and employment generate a more moderate effect on accumulated assets.

The table also reveals that the effect of the three components is not additive, a result of the model's nonlinearities. The effect on assets of mortality and time cost combined is less than the sum of the two effects: the increase in assets induced by the

[14] We report additional tables with decompositions of health levels set at the 25th, 50th, and 90th percentile in External appendix F.

[15] Capatina (2015), finds that shutting off mortality is the main driver of health effects when everyone is assigned to the high level of health of a 3-valued health variable. The main differences between her model to ours are: (1) health takes only three values that evolve with a Markov process without a transitory component; (2) the model is estimated using data from the U.S., (3) modeling the U.S. market, personal health expenditures are included as a choice variable.

Table 1: Decomposition of the effects of health on outcomes

Effects removed	Assets	Income	Empl.	Hours
None (baseline)	158	11.2	0.65	1848
(a) Counterfactual health: 75th percentile				
All (relative to baseline)	+8.0%	+11.1%	+10.7%	+6.3%
Decomposition:				
Mortality	+0.9%	-0.3%	-0.2%	+0.2%
Time cost	+4.6	+8.2	+8.4	+7.0
Wages	+1.4	+2.3	+3.1	-1.2
Mortality+Time	+4.7	+7.1	+8.2	+6.3
Mortality+Wages	+2.4	+2.0	+2.9	-1.1
Time+Wages	+7.6	+12.2	+10.7	+7.0
(c) Counterfactual health: 25th percentile				
All (relative to baseline)	-10.3%	-9.9%	-1.0%	-11.2%

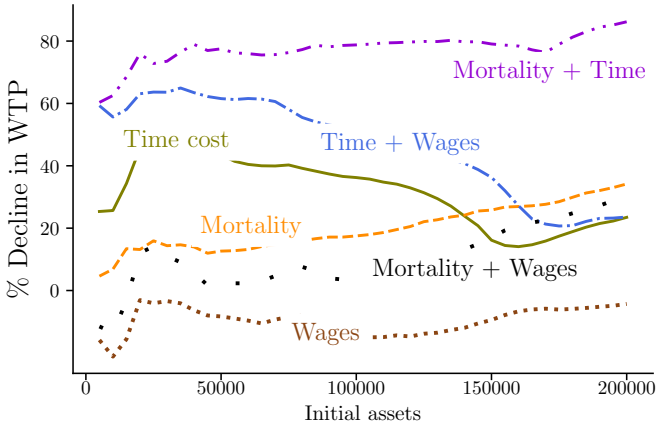
(a) Assets, Income, Employment and Hours

Effects removed	Assets		Cum. earnings at 65	
	80/20pct.	St. dev.	80/20pct.	Std. log
None (Baseline)	3.70	109	2.83	0.82
(a) Counterfactual health: 75th percentile				
All (relative to baseline)	-11%	+1%	-19%	-14%
Decomposition:				
Mortality	-4%	-1%	-2%	-2%
Time cost	-6	+2	-15	-11
Wages	-3	+0	-6	-5
Mortality+Time	-6	+0	-17	-11
Mortality+Wages	-6	-1	-8	-6
Time+Wages	-11	+2	-18	-13
(b) Counterfactual health: 25th percentile				
All (relative to baseline)	+18%	-3%	-1%	+1%

(b) Measures of inequality

Note: The table reports average outcomes in the baseline model simulations (first row), and percent changes from the baseline after assigning everyone either the mortality rate, time cost of health, or wage offer computed at the percentile of the age-specific health distribution indicated. Assets and (disposable) income in Panel (a) and standard deviation of assets in Panel (b) measured in £1000 averaged throughout the life cycle. Employment rate and hours averaged from age 50 to 69. Hours are annual, conditional on employment. Cumulated earnings in Panel (b) measured at age 65. See Table F.4 for the decomposition at the 25th and 50th percentile of health.

Figure 4: Components of willingness to pay to avoid a bad health shock



Note: The figure reports the percent decline in willingness to pay to avoid a bad health shock $\tau_{shock} = 0.1$ at age 52, for individuals starting from $\tau_{init} = 0.1$ with assets in the horizontal axis, after removing the effect of health on mortality, time available, wages, and their combinations. The effect of health is removed by computing the value the corresponding variable(s) assume at the 75th percentile of the health distribution conditional on age

reduced time cost is enough to fund the increased lifespan. When mortality does not change, instead, the effect of time cost and wages combined is *more* than the sum of the single effects when health increases, when health is set at the 75th percentile.

We perform the same decomposition on the willingness to pay to avoid a bad health shock. In Figure 4 we report the willingness to pay to avoid a bad shock (as a percent of the baseline willingness to pay) after removing the effect of health on each channel or their combination. To remove the effect of health, we assign individuals the value each variable would take at the 75th percentile of health (conditional on age) as in Table 1(a). We consider a health shock $\tau_{shock} = .1$ at age 52, starting from $\tau_{init} = 0.1$.

Interestingly, the results depend on the value of the initial assets of the individual. The time cost of bad health is again the most important factor. Removing this factor would decrease the willingness to pay by between 20 and 50%. The effect of health on mortality has a small impact only for the poorest individuals; for the highest-wealth individuals, removing this effect would decrease the willingness to pay by over 34%. This is because bequest motives become more important at a higher level of wealth and are negligible at low wealth. Removing the effects of health on mortality and time endowment combined reduces by 80% the willingness to pay to avoid the shock throughout most of the initial assets range we considered. Removing the effect of health on wages *increases* the willingness to pay to avoid a bad shock. This is because, in this experiment, we assign a majority of individuals a higher hourly wage than under the baseline (equal to the hourly wage offered to those at the 75th

percentile of health). Following a bad shock, participation and hours worked decline, hence compared to the baseline, a bad shock results in a higher income loss. The wages effect complements the effect of mortality since adding the effects of mortality and wages generates a willingness to pay closer to the baseline than mortality alone. Instead, a comparison of the dash-dotted blue line with the solid green line reveals that adding the effect of wages to the effect of the time endowment reduces the willingness to pay, meaning that the substitution effect dominates the income effect when the individual has more hours to allocate and a higher wage offer.

The results from the decompositions performed in this section suggest that policies aimed at reducing the time cost of bad health, such as incentivizing remote work or providing home-care support, would be especially effective in improving the welfare of those in bad health.

5.4 The effect of health on inequality

Our measure of health allows us to compute the effects of health on inequality at a granular level. The results from the previous section reveal the importance of the model's nonlinearities and interdependency between channels for assessing outcomes, accounting for a continuous measure of health changes results quantitatively and qualitatively. In this section, we show that different levels of health affect not only outcome levels but also their variation.

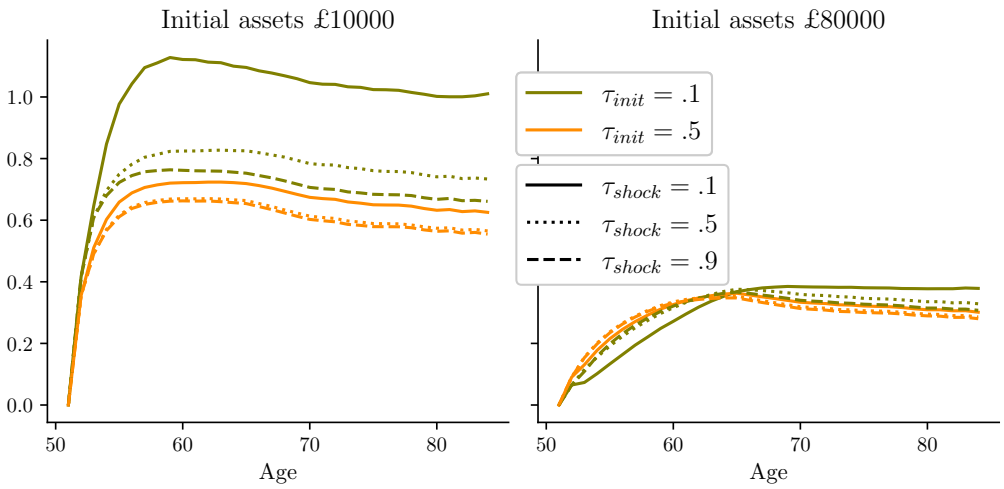
5.4.1 Removing the effects of health

Table 1(b) reports the effect of decomposing the effect of health on two measures of inequality. Removing all health effects by setting mortality, the time cost of health, and wages at the level they have when health is equal to the 75th percentile by age reduces the 80/20th percentile ratio by 0.42, or 11%. The standard deviation of assets increases slightly. The 80/20th percentile ratio of cumulated earnings at 65[16] decrease by 19% and the standard deviation of its log decreases by 14%. The decomposition reveals that the time cost of bad health is the most important contributor to inequality.

Interestingly, indices of inequality increase or decrease only slightly if health is set at the 25th percentile of the distribution. The reason is that this experiment is equivalent to imposing a negative health shock on a majority of individuals, which, as highlighted earlier in Figure 3(b), affects the poor more than the rich. While setting everyone to the same percentile of health level contributes to reducing variation, the differential effect of the shock by wealth is larger in magnitude. Moreover, under this counterfactual, labor force participation decreases, contributing to increasing or moderating the decline in inequality: the fraction of individuals with zero cumulated earnings at 65 increases from 4.1% in the baseline to 6.3% when all the effects of

[16]Computed as in Guvenen et al. (2022) and Hosseini et al. (2021b).

Figure 5: Assets, coefficient of variation following a shock



The Figure reports the age profiles of the coefficients of variations of accumulated assets for different levels of τ_{init}, τ_{shock}, and initial assets. Average of 5,000 histories per simulation.

health are removed (this fraction decreases to 2.4% in the counterfactual health at the 75th percentile).

5.4.2 The effects of bad health shocks on inequality

The interpretation of the previous subsection's results is confirmed if we analyze the effect of the health shocks used to generate Figure 3(a) on inequality. Figure 5 displays the coefficient of variation of assets, by age, for people starting at age 51 with two initial levels of wealth, following permanent health shocks of three different intensities (marked by lines with different patterns) at the age of 52, and starting from a low or median level of the permanent component of the health shock (marked by lines of a different color).

The dispersion in assets across individuals is higher when starting from a lower initial level of the health shock and when starting with low wealth (solid green line in the left panel). For poorer individuals, dispersion increases dramatically after a bad shock, peaking above 1.4 before the age of 60, and remaining higher than 1.2. On the other hand, for individuals with the same initial wealth, dispersion is lower after a good shock than after a median shock. This result follows because while the variance of assets is approximately constant after the shock, average assets decrease after a bad shock. Differences in dispersion are negligible for initially wealthier individuals (right panel), and (not reported in the figure) are even smaller for people starting from an initial level of wealth of £200,000.

The persistence of bad shocks affects asset accumulation permanently, especially

so for individuals starting with a low level of wealth. This evidence helps explain the differential effects of health we uncovered previously when discussing the decomposition results of Table 1(b). Bad health shocks affect inequality differently by wealth, contributing to increasing variation when health is reduced to the 25th percentile for all. Good health shocks do not affect inequality as differently; hence when all effects of health are computed at the 75th percentile, the equalizing-health effects dominate, and indices of inequality are lower than the baseline.

5.4.3 The canonical model does not uncover the same differential effects on inequality

Notice, in the left panel of Figure 5, the asymmetric effects of good and bad shocks with respect to the median shock. Nonlinearities in the health process contribute to generating this asymmetry. In the nonlinear model, four years after a bad shock, the coefficient of variation gets over 35% larger than it would after a median shock and remains this high to the end of the life cycle. In the canonical model, variation increases, but not as dramatically, reaching about 25% of its level after a median shock (see External appendix Fig. F.6 for details).

For individuals starting off at age 50 with higher wealth, the two models perform similarly. These results show that for poor individuals, the canonical model cannot generate the same dispersion of outcomes generated by health shocks.

6 Conclusion

This paper shows that the health dynamic is highly nonlinear and that the quantile-based methodology developed by Arellano et al. (2017) can fit the health process' main features. Incorporating the estimated health process into a life-cycle model reveals substantial heterogeneity in the response to severe health shocks by initial health and wealth. A canonical model of health dynamics that fits the life-cycle profiles of assets, labor force participation, and hours worked equally well generates different economic effects of health shocks, especially for poor individuals subject to a bad shock. Our research reveals the importance of accounting for a complex health measure, especially when investigating the distributional effects of health shocks on poor individuals.

Our simulations highlight the time cost imposed by bad health as the most significant factor affecting sick individuals' welfare. This result suggests the importance of policies aimed at relieving individual time pressure, such as facilitating work from home or providing home care.

Because of computational limitations, we did not consider extensions of our model that may appear natural, such as the joint nonlinearity of the health, income, and mortality processes or the endogeneity of the health process. These are important factors that we leave to future research.

A Appendix: model and estimation details

Following De Nardi et al. (2020), after estimating the nonlinear health process using Arellano et al. (2017)'s procedure we simulate health histories and discretize the persistent and transitory components of health at each age and then compute the transition matrices for the persistent component. See the External appendix for additional details and descriptions of the estimation of the canonical health and income processes, mortality rates, and wealth profiles.

To define the Bellman Equation, let the state space be $X_t = \{a_t, e_t, h_t, p_t\}$. The dynamic programming problem for workers is to maximize over hours worked and next period's assets the Bellman equation:

$$\begin{aligned} V(X_t) &= \max_{a_{t+1}, s_t} \Bigg\{ U(c_t, l_t) + \beta \pi^{t+1} \iint_{e_{t+1}, h_{t+1}} V(X_{t+1}|X_t) dF(X_{t+1}|X_t) \\ &\quad + \beta(1-\pi^{t+1}) b(a_{t+1}) \Bigg\} \end{aligned} \tag{9}$$

s.t.:

$$a_{t+1} = a_t + s_t \cdot e_t \cdot (1 - c_p \mathbb{1}(t < 15)) + r \cdot a_t - tax_t + tr_t - c_t - \phi_w(\mathbb{1}(s_t > 0, t)) \tag{10}$$

$$\begin{aligned} l_t &= L - s_t - \phi_h(h_t) \\ tax_t &= f(s_t \cdot e_t \cdot (1 - c_p \mathbb{1}(t < 15)) + r_p \cdot p_t \mathbb{1}(t \geq 15), r \cdot a_t) \\ tr_t &= \max\left(0, \underline{c} - (a_t + r_p \cdot p_t \mathbb{1}(t \geq 15) + s_t \cdot e_t \cdot (1 - c_p \mathbb{1}(t < 15)) + r \cdot a_t - tax_t)\right) \\ a_t &\geq 0 \end{aligned}$$

Transfers tr_t are computed so that $c_t \geq \underline{c}, \forall t$. States h_t, e_t evolve according to the processes specified in Appendices C and D, with earnings $e_t = 0$ if $t \geq 21$ (after age 69). Assets a_t evolve according to constraint (10), and pension wealth $p_{t+1} = p_t + e_t(1 - c_p \mathbb{1}(t < 15))$.

The time cost of health $\phi_h(h_t)$ is a 5-knots spline defined between $\underline{h}$ and $\hat{h}$, the minimum and maximum value of health in our simulated sample, and with knots $\underline{h}, h_{20}, h_{30}, h_{50}, \hat{h}$, where the three intermediate knots are the 20th, 30th, and 50th percentile of the simulated health distribution (respectively, -.633, -.170, .340). We constrain $\phi_h(\hat{h}) = 0$. Table 2 reports the value of ϕ_h at the remaining four knots. The cost of work is a linear function

$$\phi_w(\mathbb{1}(s_t > 0, t)) = \begin{cases} \phi_w^1 + \phi_w^2 \cdot t, & \text{if } t < 65 \\ \phi_w^1 + \phi_w^2 \cdot t + \phi^3 \cdot s_t & \text{if } t >= 65 \text{ and } s_t > 1250 \end{cases}$$

The extra cost suffered by older workers working part-time is identified by the sudden drop in participation and hours worked occuring at the age of 65.

To estimate the model, we simulate 15,000 histories and minimize the sum of the squared difference between simulated moments and moments from data. We target

Table 2: Life-cycle model parameter estimates

Calibrated		
Time endowment	L	4880
Consumption floor	$\underline{c}$	1660
Risk aversion	ν	4
Pension annuity rate	p_r	0.0378
Pension contribution rate	c_w	0.06
Discount factor	β	0.9756

Estimated		Nonlinear	Canonical
Consumption weight	γ	0.378	0.379
Bequest	ϕ_B	0.042	0.044
	K	533219	533219
Cost of health	$\phi_h(h_{50})$	1401.9	1402.0
	$\phi_h(h_{30})$	1409.0	1412.1
	$\phi_h(h_{20})$	2312.5	2403.5
	$\phi_h(\underline{h})$	4879.0	4878.5
Cost of work	ϕ_w^1	3585.0	3597.9
	ϕ_w^2	32.8	32.9
	ϕ_w^3	2.8	2.9

135 averages: accumulated assets by age (35 moments), hours worked if working, by age (20 moments), and labor force participation by age and 4 health quantiles (below the 20th percentile, between the 20th and the 30th percentiles, between the 30th and the 50th percentiles, and above the 50th percentile) (80 moments).

We discretize the state space of assets, state pension, wages, and transitory and permanent components of health, respectively using 30, 6, 5, 5, and 19 (24 in the canonical model) gridpoints. To minimize the loss function we adopt a combination of simulated annealing and downhill simplex procedures, starting from several initial values of the parameter vector, and stopping when the routine converges to a parameter vector equal to the initial starting vector. In our loss minimization routine we also constrained $K > 0$ and $\phi_h(\cdot) < L$. Table 2 reports the estimated parameter values.

References

Ameriks, John, Joseph Briggs, Andrew Caplin, Matthew D. Shapiro, and Christopher Tonetti, "Long-Term-Care Utility and Late-in-Life Saving," *Journal of Political Economy*, 2020, *128* (6), 2375–2451. (Cited on page 2)

Arellano, Manuel, Richard Blundell, and Stéphane Bonhomme, "Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework," *Econometrica*, 2017, *85* (3), 693–734. (Cited on pages 2, 3, 5, 7, 17, and 18)

Blundell, Richard, Jack Britton, Monica Costa Dias, and Eric French, "The impact of health on labor supply near retirement," *Journal of Human Resources*, 2021, pp. 1217–9240R4. (Cited on pages 2 and 6)

Blundell, Richard W, Jack Britton, Monica Costa Dias, and Eric French, "The dynamic effects of health on the employment of older workers," Michigan Retirement Research Center Research Paper 2016. (Cited on page 4)

Bound, John, "Self-reported versus objective measures of health in retirement models," *Journal of Human Resources*, 1991, *26* (1), 106–138. (Cited on pages 2 and 4)

_ , Michael Schoenbaum, Todd R. Stinebrickner, and Timothy Waidmann, "The dynamic effects of health on the labor force transitions of older workers," *Labour Economics*, 1999, *6*, 179–202. (Cited on pages 4 and 5)

_ , Todd Stinebrickner, and Timothy Waidmann, "Health, economic resources and the work decisions of older men," *Journal of Econometrics*, 2010, *156* (1), 106–129. (Cited on page 4)

Burkhauser, Jean M. Mitchell J. S. Butler Richard V. and Theodore P. Pincus, "Measurement error in self-reported health variables.," *Review of Economics and Statistics*, 2010, *69* (4), 644–650. (Cited on page 4)

Capatina, Elena, "Life-cycle effects of health risk," *Journal of Monetary Economics*, 2015, *74*, 67–88. (Cited on pages 2 and 12)

Dal Bianco, Chiara, "Disability Insurance and the Effects of Return-to-work Policies," Working paper, Università di Padova 2022. (Cited on pages 2 and 6)

De Nardi, Mariacristina, Eric French, and John B Jones, "Why do the elderly save? The role of medical expenses," *Journal of Political Economy*, 2010, *118* (1), 39–75. (Cited on page 2)

_ , Svetlana Pashchenko, and Ponpoje Porapakkarm, "The lifetime costs of bad health," National Bureau of Economic Research working paper #w23963 2017. (Cited on page 2)

French, Eric, "The effects of health, wealth, and wages on labour supply and retirement behaviour," *The Review of Economic Studies*, 2005, *72* (2), 395–427. (Cited on pages 2 and 8)

_ and John Bailey Jones, "Health, health insurance, and retirement: a survey,"

Annual Review of Economics, 2017, *9*, 383–409. (Cited on page 2)

Guvenen, Fatih, Greg Kaplan, Jae Song, and Justin Weidner, "Lifetime Earnings in the United States over Six Decades," *American Economic Journal: Applied Economics*, October 2022, *14* (4), 446–79. (Cited on page 15)

Hosseini, Roozbeh, Karen A. Kopecky, and Kai Zhao, "The evolution of health over the life cycle," *Review of Economic Dynamics*, 2021. (Cited on pages 2, 5, and 6)

_ , Karen A Kopecky, and Kai Zhao, "How important is health inequality for lifetime earnings inequality?," *FRB Atlanta Working Paper*, 2021. (Cited on pages 2 and 15)

Low, Hamish and Luigi Pistaferri, "Disability insurance and the dynamics of the incentive insurance trade-off," *American Economic Review*, 2015, *105* (10), 2986–3029. (Cited on page 2)

Nardi, Mariacristina De, Eric French, and John Bailey Jones, "Savings after retirement: A survey," *Annual Review of Economics*, 2016, *8*, 177–204. (Cited on page 2)

_ , Giulio Fella, and Gonzalo Paz-Pardo, "Nonlinear household earnings dynamics, self-insurance, and welfare," *Journal of the European Economic Association*, 2020, *18* (2), 890–926. (Cited on pages 8 and 18)