Published in final edited form as:

J Econ Ageing. 2019 ; 14: . doi:10.1016/j.jeoa.2018.11.001.

Health Disparities and the Socioeconomic Gradient in Elderly Life-Cycle Consumption*

Ray Miller^{†,‡}, Neha Bairoliya^{§,‡}, David Canning[¶]

[‡]Harvard Center for Population and Development Studies, Cambridge, MA, 02138, United States

[§]The University of Tennessee, Knoxville, Department of Economics, Knoxville, TN, 37996, United States

[¶]Harvard School of Public Health, Boston, MA 02115, United States

Abstract

We quantify the importance of health disparities in explaining consumption differences at older ages by estimating a panel VAR model of elderly consumption, health, and mortality using data from the Health and Retirement Study. We use the estimated model and initial joint distribution of health and consumption to simulate elderly life-cycle paths and construct a measure of the net present value of expected remaining lifetime consumption at age sixty (NPVC). We first document a steep education gradient in elderly lifetime consumption. We then decompose the gradient in NPVC to quantify the effect of 1) differences in the health distribution at age sixty and 2) differential health and mortality transitions after age sixty. Our decomposition results suggest that roughly 10-12% of the education gradient in NPVC at age sixty could be closed by eliminating elderly health differences.

Keywords

consumption inequality; health; aging; education gradient; life-cycle

JEL classifications:

D12; D30; D63; I14

*Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Number P30AG024409 and by the National Institutes of Health (NIH, Grant No.: 5R01AG048037-02). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan.

[†]Corresponding author. Colorado State University, Department of Economics, Fort Collins, CO, 80523, United States, ray.miller@colostate.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1 Introduction

The turn of the century has witnessed not only a rise in economic inequality but also growing disparities in health and mortality.¹ While the correlations between health and economic circumstances have been widely studied, it is less clear to what extent variations in the former may translate into the latter. In this study, we examine the dynamic relationship between health and consumption over the elderly life course. In doing so, we shed light on the potential impact of policies that promote health equity on broader economic inequality among the elderly.

Adverse health shocks may influence individual or household consumption through a variety of mechanisms. These effects could be contemporaneous, for example through increased expenditures on medical care, reduced earnings, or decreased consumption of goods and services that are complements to good health. Or there could be dynamic effects that persist over time, for example through forced early retirement or perhaps even expedited death. In any case, these channels suggest that variations in health may explain, at least partly, the disparate consumption patterns observed across socioeconomic groups in the U.S. This relationship could be especially important at older ages where mortality rates are high and most of the variance in health across individuals is concentrated (Deaton and Paxson, 1998).

Understanding the extent to which health differences influence consumption inequality is a challenging task. Health is a multidimensional concept with many proposed subjective and objective measures (e.g. self-rated health, physical limitations, incidence of morbidities, etc). Different types of health events can impact consumption, both current and future, with varying intensities—a cardiac event may have a stronger and more persistent effect on consumption than hypertension. There is also the possibility of dynamic spillovers across health conditions—hypertension is a risk factor known to increase the likelihood of other adverse health outcomes including death (Ettehad et al., 2016). Finally, individuals with similar annual consumption but substantially different lifespans will have very different total lifetime consumption. Moreover, the life-cycle hypothesis suggests an unexpected health shock could potentially increase contemporaneous consumption if there is an associated decline in life expectancy. In these cases, cross-sectional associations between consumption and health would only be presenting a part of the bigger story.

In light of these potentially dynamic and multi-faceted influences, we take a lifecycle approach to better quantify the total contribution of health disparities to differential consumption patterns over elderly life in the U.S. Using a similar framework as Miller and Bairoliya (2018), we estimate a panel vector autoregressive (VAR) model to approximate the joint evolution of consumption, mortality, and a multivariate measure of health over the elderly life-cycle. We use longitudinal data from the Health and Retirement Study (HRS) supplemented with the Consumption and Activities Mail Survey (CAMS). Together, these provide a long and rich panel (1992-2014) for studying the joint distribution of health and consumption for the elderly. We use the estimated system to simulate and analyze potential

¹Refer to Heathcote et al. (2010); Cutler and Katz (1992); Katz et al. (1999); Attanasio et al. (2014) for rising economic inequality and Chetty et al. (2016); National Academies of Sciences, Engineering, and Medicine (2015); Meara et al. (2008); Case and Deaton (2017) for health and mortality.

outcome paths for a sub-sample of HRS respondents. Using the simulated paths for consumption and mortality, we construct a measure of the net present value of expected remaining lifetime consumption (NPVC) at age sixty. This measure, which can also be viewed as an age sixty wealth equivalent, provides a parsimonious setup for studying both the contemporaneous as well as dynamic effects of health in driving consumption differences.

In order to examine the influence of health disparities on consumption inequality, we study differences across educational attainment levels. We choose to focus on this socioeconomic dimension as substantial education gradients have been well documented in health and mortality (Conti et al., 2010; Pijoan-Mas and Ríos-Rull, 2014) as well as income (Houthakker, 1959; Morgan and David, 1963; Griliches and Mason, 1972). It is well known that income differences across education groups primarily stem from differential returns to education. While some of these income differences may further translate into health differences due to differential access to medical goods and services, studies show that differences in health outcomes are also due to differences in health behaviors (Cutler and Lleras-Muney, 2010), knowledge capital (Kenkel, 1991) and adherence to medication (Goldman and Smith, 2002). We do not focus on the reasons underlying the observed education gradients in health and consumption in this analysis, rather we ask how much of the education-consumption gradient can be attributed to health disparities across education groups. We further decompose how much of the gradient can be explained by differences in initial (age sixty) health distribution versus the differential evolution of health and mortality after age sixty across these groups.

Our main results are summarized as follows: we first document a substantial education gradient in cross-sectional consumption at age sixty—average consumption of college graduates is 2.5 times higher than high school dropouts. This ratio rises to 3.0 when examining NPVC—our estimate of elderly lifetime consumption. This rise is due to a positive correlation between consumption, health, and longevity. For example, we estimate that the average life expectancy of college graduates at age sixty is 5.3 years longer than high school dropouts. So not only are the college educated enjoying more average consumption in the cross-section, they enjoy the consumption for a longer and healthier period of time.

Our second set of results quantify the influence of health disparities on the education-consumption gradient. First, we find that conditional on health, differential survival rates across education groups actually explain very little of the gap in NPVC. The reason is that we do not find a strong education gradient in mortality rates after controlling for health. This finding is consistent with those reported by Pijoan-Mas and Ríos-Rull (2014). The authors also find that education, wealth, and income have very little impact on survival conditional on health. We do find, however, that the differential evolution of health can explain an important share of the estimated consumption disparities. For instance, eliminating differential elderly health transitions after sixty (mortality, morbidities, and subjective health) across education groups can close an estimated 4.7% of the educational gradient in NPVC between college graduates and high school dropouts at age sixty. However, this is a lower bound of the overall effect as part of the gradient is driven by initial age sixty health

conditions. When we give the initial health distribution of the college graduates to those with less than a high school education, the consumption gradient closes by 5.7%. Altogether, we estimate that the distribution of health at age sixty and subsequent health transitions combine to explain approximately 10.5% of the gap in NPVC between college graduates and high school dropouts at age sixty—and 11.5% of the gap between college graduates and high school graduates.

Our paper broadly contributes to a growing body of work studying the distribution of individual and household consumption (Krueger and Perri, 2003; Deaton and Paxson, 1994; Dynarski et al., 1997; Cutler et al., 1991) along several dimensions. We establish a significant education-consumption gradient at older ages and document that inequality across education groups accounts for a substantial share of the total inequality in elderly lifetime consumption. We also argue that cross-sectional consumption differences may underestimate the true nature of consumption disparities among the elderly and propose NPVC as an alternate measure for studying the distribution of consumption. Moreover, while previous studies have explored the role of income inequality in explaining consumption differences (Krueger and Perri, 2006; Cutler and Katz, 1992; Blundell and Preston, 1998; Aguiar and Bils, 2015), we provide an estimate of the impact of health. Finally, we add to the existing work looking at the importance of health disparities in explaining economic inequality (Shastri and Weil, 2003; Bloom et al., 2004). While these studies have looked at the importance of population health differences in explaining cross-country income differences, we are able to provide new insights by exploring the importance of health and mortality in explaining consumption inequality within the U.S.

Our paper is most closely related to two existing studies—Pijoan-Mas and Ríos-Rull (2014) and Miller and Bairoliya (2018). The former looks at the importance of initial health distribution versus differential health and mortality transitions at older ages in explaining differences in life expectancy at age fifty. We focus on analyzing the role of health in explaining consumption differences and also use a broader indicator of health, incorporating several morbidities and physical limitations. Miller and Bairoliya (2018) estimates the welfare distribution of the elderly and its change over time. We use a similar modeling approach in this paper to understand the relationship between health and consumption inequality at older ages. We estimate the education gradient in consumption and document the importance of educational differences in explaining overall consumption inequality. We also conduct decomposition exercises to quantify the impact of health on elderly consumption disparities.

The remainder of the paper is organized as follows. Section 2 describes our data and empirical methods used in this analysis. Section 3 discusses our estimation and simulation results and summarizes our findings from the decomposition experiments. Finally, section 4 provides concluding remarks.

2 Data and methods

2.1 Data

The Health and Retirement Study is a nationally representative, longitudinal panel survey of older Americans. It covers individuals over the age of fifty and their spouses on a biennial basis. The survey was first conducted in 1992 for the initial HRS cohort with multiple other birth cohorts added over subsequent waves. The study presently consists of five primary birth cohorts—the initial HRS cohort (born 1931-1941), AHEAD cohort (born before 1924), Children of Depression (born 1924-1930), War Babies (born 1942-1947), and Baby Boomers (born 1948-1959).² The HRS is a rich source of information on income, wealth, health, and demographics for the elderly. Household spending data is available from the Consumption and Activities Mail Survey (CAMS), which was sent to a random sub-sample of HRS participants during off years of the core survey, starting in 2001. We use the cleaned RAND HRS data file (v.P) for health and demographic variables from 1992 to 2014. We use data on fixed characteristics including education level, gender, race, and birth cohort in our health and mortality estimations. The education variable used in our analysis is based on highest level attained and includes three categories—less than a high school education (<HS), high school graduates (HS), and college graduates (College). We describe below our health and consumption measures in more detail.

2.1.1 Consumption—The CAMS collected household spending data on durables, nondurables, transportation, and housing. We use cleaned RAND 2015 CAMS data file (v. 2), which contains a constructed estimate of total household consumption derived from the available spending data.³ Broadly, household consumption was derived by estimating the per-period “usage” from consumer durables, automobiles, and housing expenditures using a similar method as proposed by Hurd and Rohwedder (2007).⁴ We construct our measure of individual consumption by subtracting household out-of-pocket health expenditures and dividing by household size.⁵ As consumption data is only available in between the core HRS waves, we merge each CAMS wave with the HRS core data from the previous wave.⁶

A major challenge to our analysis is the presence of missing consumption data. CAMS data is only available for approximately 20% of the sample for the years 2000-2012. However, closely related data is available for all individuals across all survey waves (1992-2014) including measures of wealth, income, and labor supply. In our benchmark analysis, we follow Miller and Bairoliya (2018) and use this additional data to perform multiple imputation of missing consumption data using the method for cross-sectional time-series data proposed by Honaker and King (2010) (see Appendix A in Miller and Bairoliya (2018)

²Baby Boomers are split into two groups by the HRS (early and mid) but we group them together as very few mid Baby Boomers have the lagged data required for estimation of our dynamic model.

³Data and consumption construction details available at <http://hrsonline.isr.umich.edu>.

⁴The annual service flow for durables was roughly estimated as $C \times p$ where C was the total cost and p the probability of purchase. Adjustments for interest payments, depreciation, and insurance costs were done for estimating transportation consumption. The consumption of housing was estimated as the sum of the rental equivalent of the owned house, property tax, homeowners insurance, and any additional rent payments.

⁵Health spending includes health insurance, medication, health services, and medial supplies. We use the CPI-U to convert all waves to 2010 dollars.

⁶This is the recommended procedure for use of the RAND CAMS data file and is also consistent with the time structure of our dynamic model.

for details). As a robustness, we check the sensitivity of our estimation results to the use of only non-imputed data.

2.1.2 Health and Mortality—We use a multivariate measure of morbidity including eight binary indicators for ever having been diagnosed by a doctor with the following health problems—(1) high blood pressure or hypertension; (2) diabetes or high blood sugar; (3) cancer or a malignant tumor of any kind except skin cancer; (4) chronic lung disease except asthma such as chronic bronchitis or emphysema; (5) heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems; (6) stroke or transient ischemic attack (TIA); (7) emotional, nervous, or psychiatric problems; and (8) arthritis or rheumatism. As a final measure of morbidity, we include an indicator for ever reported difficulty with any activity of daily living (ADL). Difficulty with ADLs are a commonly used health metric among the elderly and include activities such as walking across the room, bathing, and getting dressed.

As a general health measure we use self-rated health status reported on a five-point scale from poor (one) to excellent (five). Self-rated health has been shown to be predictive of mortality in the HRS and other datasets, even after controlling for other health conditions, health behavior, and socioeconomic characteristics (Idler and Benyamini, 1997; Stenholm et al., 2014). We hypothesize that people have private information about their health over and above disease diagnosis. This may reflect the severity of the disease or additional information on other health conditions. Finally, mortality data is taken from the Tracker file. Death dates in the survey are either reported by spouses or come from the exit interview.

2.2 Model

Our primary interest is in forecasting outcomes over the entire elderly life-cycle and in computing how projected forecasts change under different counterfactual scenarios. To this end, we model and estimate a system of dynamic equations to approximate the joint evolutionary process of consumption, health, and mortality over time using a similar modeling approach as Miller and Bairoliya (2018). This allows us to 1) examine complete expected elderly life-cycle consumption profiles across socioeconomic groups and 2) quantify the impact of health on consumption differences through counterfactual exercises.

Our basic model of health and consumption is illustrated in Figure 1. At the beginning of each time period, morbidity status is realized based on random shocks and exogenous characteristics (details below). Given these morbidity conditions (and other exogenous characteristics), general (self-rated) health evolves. Morbidities and general health then affect current period consumption. Note that we posit each of the morbidity states to contemporaneously influence consumption both directly and through changes in self-rated health. For example, heart disease may affect an individual's self-rated health status which in turn may lower contemporaneous consumption. However, heart disease may also influence consumption independently of changes in self-rated health.

We introduce a dynamic health effect by allowing morbidities and self-rated health to influence the probability of survival to the following period of life. Moreover, we expand these dynamic effects by also allowing current health and consumption to influence the

evolution of outcomes moving forward (conditional on survival). In a dynamic setting, our forecasting model can be conceptualized as a panel vector autoregression (VAR) of order p . The following subsections lay out the forecasting VAR model and identifying assumptions.

2.2.1 Panel VAR representation—While we allow for multiple lags in estimation of the model, the following VAR(1) demonstrates the key features of the framework. Let Y_{it} be a vector of outcomes for individual i at time t that includes log consumption c , self-rated health s , and our $n = 9$ morbidity states given by $n \times 1$ vector M . Conditional on survival, the outcomes evolve according to the structural VAR(1) model:

$$AY_{it} = BY_{it-1} + \epsilon_{it},$$

where ϵ is a vector of normally distributed shocks with mean zero. The shocks are assumed to be independent and identically distributed (*iid*) across individuals and time and independent across outcomes. The main diagonal terms of matrix A are scaled to one and we assume in our benchmark model that all parameters are homogeneous across individuals and time (e.g. $A_{it} = A \forall i, t$).

We estimate our model in three “blocks” of outcomes—the morbidity block consisting of n outcomes, the self-rated health block (one outcome), and the consumption block (one outcome). The unrestricted model can be written in block matrix form as:

$$\begin{array}{c} n \\ 1 \\ 1 \end{array} \begin{bmatrix} \overbrace{\quad n \quad} & \overbrace{\quad 1 \quad} & \overbrace{\quad 1 \quad} \\ -A_{11} & -A_{12} & -A_{13} \\ \cdots & \cdots & \cdots \\ -A_{n1} & 1 & -a_{23} \\ -A_{n1} & -a_{32} & 1 \end{bmatrix} \begin{bmatrix} M_{it} \\ s_{it} \\ c_{it} \end{bmatrix} = \begin{bmatrix} \overbrace{\quad n \quad} & \overbrace{\quad 1 \quad} & \overbrace{\quad 1 \quad} \\ B_{11} & B_{12} & B_{13} \\ \cdots & \cdots & \cdots \\ B_{n1} & b_{22} & b_{23} \\ B_{n1} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} M_{it-1} \\ s_{it-1} \\ c_{it-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,it} \\ \epsilon_{2,it} \\ \epsilon_{3,it} \end{bmatrix},$$

where $n \times n$ matrix A_{11} has main diagonal terms scaled to one.

As illustrated in Figure 1, the causal pathways we propose suggest a block recursive system. Specifically, we assume that $A_{12} = A_{13} = B_{13} = 0$ in the morbidity block and $a_{23} = b_{23} = 0$ in the self-rated health block. In other words, we assume consumption does not affect self-rated health or morbidities and self-rated health does not contemporaneously effect morbidities. Block triangulation of the system eliminates simultaneity across blocks and allows for block-by-block estimation.⁷

2.2.2 Exogenous characteristics—We also include a $k \times 1$ vector of exogenous individual characteristics X_{it} as predictors in our model. The VAR(1) model with exogenous regressors takes the following form:

$$AY_{it} = BY_{it-1} + CX_{it} + \epsilon_{it}. \quad (1)$$

⁷Note this produces the same results as the Cholesky decomposition of shocks from a reduced form VAR.

Exogenous characteristics include dummies for age, calendar year, education, gender, race, and birth cohort. We also include a time invariant individual unobserved endowment π in the consumption equation. The endowment π is modeled as a fixed effect with no restriction on the correlation with other model regressors. The unobserved individual effect helps maintain the appropriate variance in consumption across time by effectively acting as a person specific drift in the autoregressive process. The resulting exogenous effects then take the following form:

$$CX_{it} = n \left[\begin{array}{cccccc} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} & C_{17} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{21} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} & c_{27} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{31} & c_{32} & c_{33} & c_{34} & c_{35} & c_{36} & c_{37} \end{array} \right] \underbrace{\begin{bmatrix} Age_{it} \\ Year_t \\ Education_i \\ Gender_i \\ Race_i \\ Cohort_i \\ \pi_i \end{bmatrix}}_{k \times 1}.$$

$(n+2) \times k$

We exclude time invariant exogenous regressors (education, gender, race, birth cohort) from the consumption equation ($c_{33} = c_{34} = c_{35} = c_{36} = 0$) due to collinearity with the fixed effect. However, we include socioeconomic characteristics instead of individual fixed effects in the health equations because 1) morbidities are absorbing states and self-rated health is ordinal, each of which poses difficulties in estimating dynamic panel models with fixed effects; and 2) we are interested in how average health variations across observed socioeconomic groups influences life-cycle consumption. We also normalize $c_{37} = 1$ and set $C_{17} = c_{27} = 0$.

2.2.3 Consumption—The resulting consumption forecasting equation given in system (1) can be explicitly written as:

$$c_{it} = A_{31}M_{it} + a_{32}s_{it} + B_{31}M_{it-1} + b_{32}s_{it-1} + b_{33}c_{it-1} + c_{31}Age_{it} + c_{32}Year_t + \pi_i + \epsilon_{3,it}. \quad (2)$$

This is a standard linear dynamic panel data model with lagged dependent variable and individual level fixed effects (π). Given our block recursive system, this equation may be estimated independently of other blocks with all structural parameters identified including the variance of ϵ_3 . Note that including lags of health allows for differential effects over time. For example, a recent onset of heart disease may alter consumption more than if an individual has been living with a heart disease diagnosis for an extended period of time.

2.2.4 Self-rated health—As self-rated health is not a continuous outcome but measured on a five point scale, forecasting of the measure is not a true linear VAR process. In contrast, we assume a continuous latent variable s^* underlies the observed outcome. The self-rated health model as defined in system (1) is then given by:

$$s_{it}^* = A_{21}M_{it} + B_{21}M_{it-1} + b_{22}s_{it-1} + [c_{21}, \dots, c_{27}]X_{it} + \epsilon_{2,it}, \quad (3)$$

with the observed health state defined as:

$$s_{it} = \delta \text{ if } \kappa_{\delta-1} < s_{it}^* < \kappa_{\delta} \text{ for } \delta = 1, \dots, 5$$

for cut-points $(\kappa_0, \dots, \kappa_5)$ with $\delta = 1$ representing the worst health state (poor) and $\delta = 5$ the best health state (excellent). Note that latent self-rated health is assumed to depend on the lagged value of the *observed* self-rated health category. We assume ϵ_2 is an *iid* shock with standard normal distribution. Thus the evolution of self-rated health follows an ordered probit structure.

2.2.5 Morbidities—Unlike consumption and self-rated health, block triangulation of the system does not allow direct identification of the structural parameters in the morbidity block as there are $n = 9$ separate outcomes. Instead the morbidity block is estimated as a reduced form VAR. The reduced form system is obtained by premultiplying the structural system block by the inverse of matrix $-A_{11}$:

$$M_{it} = -A_{11}^{-1}B_{11}M_{it-1} - A_{11}^{-1}B_{12}s_{it-1} - A_{11}^{-1}[C_{11}, \dots, C_{17}]X_{it} - A_{11}^{-1}\epsilon_{1,it}.$$

Denoting $-A_{11}^{-1}B_{11} = \hat{B}_1$, $-A_{11}^{-1}B_{12} = \hat{B}_2$, $-A_{11}^{-1}[C_{11}, \dots, C_{17}] = \hat{C}$ and $-A_{11}^{-1}\epsilon_{1,t} = e_t$ yields the following reduced form system:

$$M_{it} = \hat{B}_1M_{it-1} + \hat{B}_2s_{it-1} + \hat{C}X_{it} + e_{it}.$$

In the reduced form VAR all right hand side variables are predetermined at time t and morbidity states do not have direct contemporaneous effect on each other. However, the error terms e_t are composites of morbidity specific structural shocks and thus are potentially correlated across morbidity states (i.e. $cov(e_{it}, e'_{it}) \neq 0$).. This allows for contemporaneous correlation in the probability of morbidity states. For example, the onset of heart disease may be correlated with the onset of hypertension or stroke due to correlated contemporaneous shocks.

Contemporaneous morbidity shocks are assumed to follow a standard multivariate normal distribution with an $n \times n$ covariance matrix given by Σ . Note that this approach does not allow for identification of the variance in structural errors in vector ϵ_1 , but only of the variance in composite errors in vector e . Thus, while this approach is not sufficient for evaluating outcome responses to structural morbidity shocks, identification of composite errors is sufficient for forecasting outcomes as desired in our analysis.

As morbidity outcomes are binary, forecasting of the morbidity state vector is again not a linear VAR process. Similar to self-rated health, we assume a continuous latent variable m^* underlies each observed morbidity state such that:

$$m_{j,it} = 0 \text{ if } m_{j,it}^* \leq 0 \text{ } m_{j,it} = 1 \text{ if } m_{j,it}^* > 0.$$

We then have the following model:

$$\begin{bmatrix} m_{1,it}^* \\ \vdots \\ m_{n,it}^* \end{bmatrix} = \begin{bmatrix} \hat{b}_{11} & \cdots & \hat{b}_{1n} \\ \vdots & \ddots & \vdots \\ \hat{b}_{n1} & \cdots & \hat{b}_{nn} \end{bmatrix} \begin{bmatrix} m_{1,it-1} \\ \vdots \\ m_{n,it-1} \end{bmatrix} + \hat{B}_2 s_{it-1} + \hat{C} X_t + \begin{bmatrix} e_{1,it} \\ \vdots \\ e_{n,it} \end{bmatrix}. \quad (4)$$

Note that each latent morbidity variable is determined by lagged values of the other *observed* morbidity states. Given the assumed joint normality of the error structure, this morbidity block of equations is in the form of a multivariate probit model.

2.2.6 Higher order lags—Including additional outcome lags may be necessary to ensure there is no autocorrelation in the structural error terms of the system. The VAR(1) model extends easily to higher orders. For example, a VAR(2) version of our model (excluding exogenous variables X_{it} for exposition) takes the following form:

$$AY_{it} = BY_{it-1} + DY_{it-2} + \epsilon_{it}$$

and in block matrix form:

$$\begin{bmatrix} -A_{11} & 0 & 0 \\ -A_{21} & 1 & 0 \\ -A_{31} & -d_{32} & 1 \end{bmatrix} \begin{bmatrix} M_{it} \\ s_{it} \\ c_{it} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & 0 \\ B_{21} & b_{22} & 0 \\ \hat{B}_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} M_{it-1} \\ s_{it-1} \\ c_{it-1} \end{bmatrix} + \begin{bmatrix} D_{11} & D_{12} & 0 \\ D_{21} & d_{22} & 0 \\ \hat{D}_{31} & d_{32} & d_{33} \end{bmatrix} \begin{bmatrix} M_{it-2} \\ s_{it-2} \\ c_{it-2} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,it} \\ \epsilon_{2,it} \\ \epsilon_{3,it} \end{bmatrix}.$$

Here, for example, coefficient vector D_{31} allows the second lag of the morbidity state vector to directly affect current consumption. Note the same block triangulation of the system is assumed for additional outcome lags. Also note that it is not strictly required that the number of lags included be identical for each outcome. For example, excluding the second lag of self-rated health on consumption simply implies setting $d_{32} = 0$.

2.2.7 Survival—The final process to be modeled is survival from one period of life to the next. As all other outcomes are conditional on survival, mortality probabilities are estimated independently of the VAR system above. Conditional on being alive at time $t-1$, survival to the following period of life is given by:

$$\psi_{it} = I \left(\sum_{k=1}^K \left[\beta_k M_{it-k} + \gamma_k s_{it-k} \right] + \delta X_{it} + u_{it} > 0 \right), \quad (5)$$

where $I(\cdot)$ is an indicator function and $\psi = 1$ indicates survival, X a vector of observed individual characteristics (age, year, education, gender, race, and birth cohort), and u_{it} an *iid* random shock with standard normal distribution. The specification allows K lags of morbidity states and self-rated health to influence mortality probability.

2.3 Estimation

All individuals in the HRS born prior to 1960 and aged fifty and over at the time of their first survey are included in the estimation sample. This gives 35,882 unique individuals and 216,606 total individual-year observations. Following the biennial structure of the HRS, a model period corresponds to two calendar years and individuals are grouped in two-year age intervals. Appendix Table 5 shows descriptive statistics for the estimation sample by level of education. Incidence of each morbidity and self-rated health state was substantial among respondents, allowing for relatively precise estimates of their associations in the dynamic processes.

As there is no simultaneity across blocks in the system, we estimate the model block-by-block. The consumption block is comprised only of equation (2), which is a standard single equation linear dynamic panel data model with lagged dependent variables and individual level fixed effects. The equation is estimated via OLS. We use the bootstrap-based method of Everaert and Pozzi (2007) to correct for the so-called Nickell (1981) bias that is known to arise from OLS estimates of such models.⁸ Including a single period lag (two calendar years) of health on consumption and two lags (four years) for consumption on itself is sufficient to ensure that shocks are serially uncorrelated in the consumption equation.⁹

For consistency with the model of consumption, we use two lags of outcomes in all health and survival equations (i.e. we estimate a VAR(2) system and set $K = 2$ in the survival model). The ordered probit model of self-rated health (3) is estimated independently of other VAR blocks using maximum likelihood.¹⁰ The mortality equation (5) is estimated independently using a standard probit regression.

This leaves the morbidity block. The morbidity model (4) is structured as a multivariate probit with correlated shocks. We estimate this model using a chain of bivariate probit estimators as proposed by Mullahy (2016) due to the large number of outcomes and large number of observations in the HRS. With no additional assumptions, this approach allows for consistent estimation via maximum likelihood as opposed to relying on more computationally intensive simulation based methods. However, a potential estimation issue arises in the morbidity block because morbidity states are absorbing (e.g. *ever* been

⁸We implement the bootstrap with De Vos et al. (2015) Stata routine *xtbcfe*. We use the deterministic initialization as our benchmark where initial conditions are set equal to those observed. However, results are insensitive to the use of the burn-in initialization which assumes that initial conditions are in the infinite past and are drawn using the same bootstrap procedure used for bias correction (see De Vos et al. (2015) for details).

⁹Second lags of health outcomes were insignificant and noisy so we opted for parsimony by excluding them. This is equivalent to estimating the VAR(2) system with $D_{31} = d_{32} = 0$. First order autocorrelation was tested for consumption using the approach of Born and Breitung (2016) and implemented in Stata with Wursten et al. (2016). Under the null hypothesis of no autocorrelation, p-values were all greater than 0.613 regardless of imputed dataset used for the test.

¹⁰Note there is no incidental parameters or initial conditions problem in this case as there is no permanent unobserved heterogeneity or serial correlation in the self-rated health (or morbidity) model. The standard (ordered) probit estimator is consistent and provides asymptotically valid test statistics and standard errors.

diagnosed with heart disease). This means, for example, diagnosed heart disease at time t perfectly predicts heart disease at time $t + 1$ and we have quasi-complete separation. This implies the effective coefficients on the lagged dependent variables in the morbidity block are infinity (i.e. $\hat{b}_{11}, \hat{b}_{22}, \dots, \hat{b}_{nn} = \infty$, in system (4)). In a simple univariate probit model, the obvious solution is to condition on not being diagnosed with the morbidity at time t . However, estimation of the bivariate probit involves maximization of the joint likelihood function, so we estimate the model as is, without conditioning on time t morbidity status. While inclusion of all observations in the bivariate probit does not effect the likelihood or estimates of the remaining (non-infinite) coefficients, it is possible there could be numerical convergence problems. However, inclusion of all observations does not result in numerical instability in our case and likelihoods converge without issue. Moreover, conditioning the bivariate probit on, for example, not having been previously diagnosed with heart disease results in nearly identical estimates for parameters in the heart disease equation as the unconditional bivariate probit.

2.4 Simulations

After empirically estimating the parameters of the dynamic models, we use them to construct life-cycle profiles under a number of scenarios for a subset of sixty year-olds from the HRS. As our benchmark, we limit individuals in our simulation sample to the initial HRS cohort (born 1931-1941) as this was the first group to be included in the HRS and has the longest available panel of data.¹¹ Note that our model requires up to four years of lagged outcomes implying data is needed from age fifty-six as part of age sixty “initial” conditions. The HRS is structured such that individuals were initially surveyed at all ages between fifty and sixty implying this lagged data is not available for all individuals within the initial HRS cohort (e.g. some respondents were already sixty when first interviewed). Limiting our sample to those in the HRS cohort with the requisite lagged data leaves a simulation sample of 6,544 individuals.

Table 1 provides a summary of initial (age sixty) conditions in the simulation sample by level of education. Cross-sectional consumption at age sixty averaged \$17,870 for high school dropouts compared to \$45,490 for college graduates—a 2.5 fold difference. By most measures, health outcomes also demonstrated a steep educational gradient (with cancer being the clear exception). For example, almost 18% of high school dropouts reported heart disease but only 12.5% of college graduates. Similarly, 32.3% of college graduates reported they had never been diagnosed with any of the examined health conditions at age sixty with a corresponding number among high dropouts of only 18.8%. Mirroring these patterns, 15.1% of high school dropouts reported poor health but only 1.7% of college graduates.

Using age sixty data as initial ($t = 0$) conditions¹², we simulate the remaining life outcomes for each individual (i) as follows:

¹¹To be clear, individuals in the simulation sample were also included in the estimation sample. We conduct a sensitivity test in which the simulation sample is excluded from the estimation sample.

¹²Initial conditions also include unobserved endowments $\hat{\pi}$ estimated from model (2) using the prediction method of De Vos et al. (2015).

1. Survival shock u_{i1} is drawn and survival to time $t = 1$ (age 62) is determined according to equation (5). If individual survives, move to step two.
2. Morbidity shock vector e_{i1} is drawn from a standard multivariate normal distribution with estimated covariance matrix Σ . This shock vector along with the model outlined in equation (4) is used to compute simulated age 62 morbidity vector M_{i1} .
3. Given age 62 morbidities (M_{i1}), general health shock $\epsilon_{2,i1}$ is drawn and age 62 self-rated health (s_{i1}) is computed using equation (3).
4. Given age 62 self-rated health (s_{i1}) and morbidities (M_{i1}), consumption shock $\epsilon_{3,i1}$ is drawn to determine age 62 consumption (c_{i1}) using equation (2).
5. Steps 1-4 are repeated for $t = 2, 3, \dots$ until death or $t = 30$ (age 120).

The above process is repeated 5,000 for each individual and averaged across simulations to obtain measures of expected consumption, health, and mortality.

A comparison between the average simulated life-cycle profiles and those based on available data is shown by highest level of education in Figures 8-11 in appendix 4. Overall, the simulations match the available aggregated data well suggesting our life-cycle dynamics model provides a reasonable approximation of the underlying data generating processes. Note that by construction, the data and simulations are the same at age 60. However, using only age 60 data and the estimated model parameters, the simulations continue to match the data reasonably well even 18 years later, when the cohort is aged 78.¹³

2.5 NPVC

Our main outcome variable in this analysis is the net present value of expected remaining lifetime consumption starting at age sixty for each individual. It can be written as:

$$NPVC_i = E \sum_{t=60}^T \frac{\psi_{it} c_{it}}{(1+r)^{t-60}},$$

where the expectation operator is defined with respect to life-cycle health and consumption shocks. The benchmark risk-free interest rate (r) is set to 3%.

3 Results

3.1 Model estimates

Figure 2 provides select coefficient estimates from the panel VAR model while the full set of results are provided in Tables 6-8 in appendix 4. The first panel in Figure 2 provides the estimated average marginal effect of each morbidity condition on the contemporaneous probability of being in poor self-rated health. For example, cancer is associated with an

¹³The simulations match somewhat less well for the standard deviation of consumption for college graduates (bottom right panel of Figure 11). This is due to data variability at older ages driven by relatively few observations and sensitivity of standard deviation to high consumption outliers. Note that the simulated *average* consumption for college graduates still matches the data well, even at older ages.

increased probability of reporting the poor health state of 8.5 percentage points. In contrast, arthritis is associated with an increased probability of only 2.2 percentage points. As with cancer and arthritis, each of the other morbidities have a significant negative association with self-rated health.

As shown in the second and third panels of Figure 2, many morbidities have a direct negative effect on contemporaneous consumption and/or survival to the following period of life. A recent stroke, for example, is associated with an increased probability of death and a loss of consumption independent of its effect through lowering self-rated health. In contrast, arthritis does not have a direct statistically significant association with mortality or consumption. However, morbidities also effect consumption and survival indirectly through changes in self-rated health. Lower self-rated health is associated with a significant decrease in contemporaneous consumption and an increased probability of death. See, for example, the coefficients on self-reporting good health shown in the last row of Figure 2. Good health is associated with an increase in consumption of 0.08 log points relative to reporting poor health (the reference state). Moreover, reporting good instead of poor health is associated with a decreased probability of death of 6.9 percentage points.

While the first two panels in Figure 2 provide insights into the contemporaneous associations across morbidities, general health, and consumption in the model, these relationships continue to evolve dynamically throughout the system over the life-cycle. As an illustrative example, the final panel in Figure 2 shows the average marginal effect of each morbidity and self-rated health on the probability of having a stroke the following model period. Heart disease, for example, increases the probability of stroke by over half a percentage point. Figure 3 furthers this example by plotting the continued response to the onset of heart disease at age sixty-two on all model outcomes over time. For example, the onset of heart disease is associated with nearly a 25% increase in the probability of stroke and lung disease by the mid-seventies. Heart disease diagnosis is also associated with an immediate 74% spike in the probability of reporting the poor health state. This heightened probability of poor health fades with time but remains over 10% the remainder of the life-cycle (conditional on survival). Age-specific probability of death is also estimated to be over 40% higher by age seventy and remains more than 10% higher even into the nineties.

The final quadrant of Figure 3 plots the response of consumption (unconditional and conditional on survival) to the onset of heart disease. There is an immediate 1.6% decline in consumption conditional on survival, which recovers somewhat to 1.2% the following period. However, consumption then begins to decline steadily and remains 1.5-2.5% lower over the entirety of the remaining life-cycle. When examining expected unconditional consumption (i.e. imputing zero consumption for the dead state), levels continue to decline over the remaining life-cycle, reaching differences of more than 30% by the early nineties.

Turning to the socioeconomic differences in the estimation results, Figure 4 plots the estimated average marginal effects of educational attainment on health indicators. Relative to a college education, high school completion is associated with an increased probability of all health conditions except cancer, psychiatric problems, and arthritis. For example, relative to college graduates, completing only high school is associated with a marginal increase in

the probability of hypertension of one percentage point. Completing less than a high school education is associated with even larger marginal increases for most health outcomes (with heart disease and stroke as the exceptions). Moreover we find that education, conditional on morbidities, is associated with a higher level of self-rated health. For example, the average marginal increase in the probability of reporting poor health is 1.1 percentage points for high school graduates and 2.4 percentage points for high school dropouts.

3.2 Education gradient in health and consumption

In this analysis, we choose to focus on comparing differences in NPVC across education groups as previous studies have noted a strong education gradient in health, mortality, and income. Moreover, we find that consumption inequality across education groups accounts for as much as 28% of the total inequality in NPVC in our sample based on decomposition of the Theil index (see appendix Table 9). This is substantial, particularly in comparison to the share of variation explained by other observable characteristics like race (~ 4%) and sex (< 1%).

We focus on two measures of the education gradient in health and consumption—1) the gap in life expectancy or NPVC between high school dropouts and college graduates ($_{lhs}$); and 2) the same gaps between high school graduates and college graduates ($_{hs}$). These gradients in both life expectancy and NPVC are reported in Table 2. It is worth noting that the average life expectancy of college graduates at age sixty is about 5.3 years longer than high school dropouts and 2.3 years longer than high school graduates. Likewise, we find a strong education gradient in the lifetime consumption of the elderly with gaps equal to \$323,000 for high school graduates and \$512,000 for high school dropouts. In other terms, the average NPVC of college graduates is about 3.0 times higher than those with less than a high school education. This ratio is 20% larger than the 2.5 ratio observed in cross-sectional consumption at age sixty (refer to Table 1). The increase in disparities when comparing NPVC as opposed to cross-sectional consumption is driven by a positive correlation between consumption, health, and mortality. See, for example, Figure 5 which shows a strong positive association between life expectancy and cross-sectional annual consumption at age sixty. This indicates that life-cycle events unfolding after age sixty (differential mortality, health shocks, etc.) are important in driving consumption inequality among the elderly. Hence, analyses using only cross-sectional data will underestimate the magnitude of total elderly consumption inequality by a significant margin.

Figure 6 shows average simulated life-cycle profiles for select model outcomes by level of education. The first panel shows the average profile of consumption (conditional on survival). Annual consumption tends to fall over elderly life across all education groups. However, cross-sectional consumption inequality increases slightly with age—the average consumption of college graduates climbs from 2.5 times that of high school dropouts at age sixty to 2.7 times at age ninety. The second panel gives the cumulative mortality probability by education. For example, on average, sixty year-olds with less than a high school degree have an estimated 50% chance of surviving to age eighty, compared to a 70% chance among college graduates.

The bottom two panels of Figure 6 show the probabilities of being in the worst (poor) and the best (excellent) self-rated health states over the life-cycle. The likelihood of reporting poor health increases substantially for all education groups between age sixty and ninety with a corresponding decline in the probability of remaining in excellent health. The college educated are least likely to be in poor health and most likely to be in excellent health throughout elderly life. Moreover, health deteriorates with age at a faster rate for the less educated. For instance, college graduates are 3.3 times more likely than high school dropouts to be in excellent health at age sixty but 5.3 times more likely by age eighty. Given these profiles, it seems plausible that health and mortality differences might be driving a substantial portion of the education gradient in NPVC of the elderly. We next turn our attention to a series of consumption gradient decomposition exercises to quantify the importance of these relationships.

3.3 Decomposition

We aim to answer the following key question in this analysis: how much of the education gradient in NPVC can be closed by eliminating health gaps among the elderly? Towards this goal, we conduct a number of experiments to estimate the impact of initial health differences at age sixty as well as the differential evolution of health across education groups after age sixty. In all our experiments, we eliminate disparities by assigning health initial conditions/transitions of college graduates to lower education groups.

Our main decomposition results are presented in Table 3. Columns (2) and (6) report the education gradient in average life expectancy and NPVC at age sixty for those with less than a high school degree (i.e. the gap between college graduates and high school dropouts). Columns (4) and (8) show analogous gradients for high school graduates (i.e. the gap between college and high school graduates). The first row repeats the baseline gradients shown in Table 2 for ease of comparison. The remaining rows summarize our decomposition exercises.

3.3.1 Health transitions—In our first set of experiments (rows two through five) we assign the health processes of college graduates to other education groups to understand how health evolution after sixty influences consumption differences. We begin by assigning only the mortality transitions (as shown in the survival model (5)) of the college graduates while holding all other transitions as in the baseline. As shown in the second row of Table 3, differences in the evolution of mortality, conditional on morbidities and self-rated health, explain very little of the consumption gradient. For high school dropouts, receiving the conditional mortality evolution of college graduates results in only a small increase in their life expectancy (0.3 years) and a decline in the consumption gradient of 0.4%.¹⁴ The analogous numbers for high school graduates are only moderately larger—a gain in life expectancy of 0.55 years and a 1.9% decline in the consumption gradient.

For both education groups, differences in the evolution of mortality, conditional on health, explains a smaller share of the observed gradient in both life expectancy and consumption,

¹⁴0.6% calculated as: $(5.12 - 5.10)/5.12 = 0.004$.

as compared to self-rated health and morbidity differentials. As shown in the third row of Table 3, assigning only the evolution of self-rated health, conditional on morbidities, decreases the gradient in life expectancy by 25.8% and consumption by 2.7% for high school dropouts—29.1% and 2.8% for high school graduates. Morbidity processes alone have a slightly smaller impact on the education gradient, closing the life expectancy and consumption gaps by 14.2% and 1.4% for high school dropouts and 19.4% and 1.9% for high school graduates.

Assigning all the health processes (mortality, self-rated health, and morbidities) of college graduates results in a 13.4% increase in life expectancy and 9.4% increase in NPVC of the high school dropouts. This implies that differences in health transitions at older ages alone can account for roughly 4.7% of the gap in NPVC estimated at age sixty. Further decomposition reveals that approximately 2.9% of the decline in the education-consumption gap is attributable to longer life expectancy alone with the remaining 1.8% arising due to the effect of living with improved health and fewer morbidities.¹⁵ Overall numbers are similar but somewhat higher for high school graduates, with the associated NPVC gradient closing by 6.8% when receiving all health transitions of college graduates.

In order to gain a sense of how differential health processes impact outcomes over the life-cycle, Figure 7 plots the average percentage change in select outcomes for high school dropouts when given all the health transitions of college graduates. There are large declines in the incidence of a majority of health conditions. For example, there is over a 25% decline in the probability of being diagnosed with lung disease by age eighty. Incidence of diabetes, stroke, and psychiatric problems all fall by more than 10% over a similar time frame. There are also significant and sustained improvements in self-rated health and mortality. For example, at the age of eighty, the probability of being in poor self-rated health is almost 50% lower while the probability of being dead is almost 20% lower. As a result of these health improvements, we see a steady increase in consumption over the life-cycle. Conditional on survival, consumption increases peak at about 6% by the late eighties. Accounting for mortality gains results in more than twice as high expected “unconditional” consumption by the mid-nineties.

3.3.2 Initial health conditions—We next turn to the role of initial health differences in explaining the estimated consumption gradient. In our previous experiments we only changed the evolution of health processes after age sixty while keeping the initial distribution of health the same for each education group. Due to a strong negative correlation observed between age sixty consumption and poor health, it is difficult to disentangle the effect of differential initial health conditions by simply assigning the initial health distribution of the college graduates to those in other education groups. We circumvent this issue by comparing sub-groups of individuals that have similar initial health states. Specifically, we first break the sample of individuals with less than high school education into N different sub-groups based on observable and unobservable characteristics—race, gender, and decile of the individual unobserved endowment π .¹⁶ We then compute the

¹⁵This decomposition entails assigning the counterfactual mortality profiles to high school dropouts while keeping their health and morbidity profiles as in the baseline.

average NPVC for each of these sub-groups conditional on initial (age sixty) self-rated health ($NPVC_{lhs}^{h,n}$). We use the health distribution of the college graduates with identical characteristics to compute a weighted average of the NPVC for each of the N subgroups of high school dropouts. Finally, we compute our counterfactual $NPVC_{lhs}^*$ by averaging over the distribution of the characteristics for the <HS group as follows:

$$NPVC_{lhs}^* = \sum_{n=1}^N \omega_{lhs}^n \sum_{h=1}^H NPVC_{lhs}^{h,n} * \omega_{coll}^{h,n}$$

where $\omega_{coll}^{h,n}$ is the share of college graduates in initial health state h within sub-group n and ω_{lhs}^n is the share of high school dropouts (unconditional on health) in sub-group n . We then use $NPVC_{lhs}^*$ and baseline NPVC of the college graduates to compute the new gradient. We conduct an analogous calculation for the life expectancy gradient.

As shown in Table 3, weighting by the initial health distribution of college graduates substantially lowers the estimated gradients in life expectancy. The life expectancy gap between college graduates and high school dropouts falls to 3.21 years—or about 40% smaller than the baseline gap. The NPVC gap also falls to \$483,000, suggesting that differences in age sixty health alone can explain about 5.7% of the education-consumption gradient. Results are similar but slightly smaller for high school graduates—the gradients in life expectancy and NPVC fall by 33.9% and 4.6% respectively.

The last row in Table 3 shows the same weighted average results when college health transitions are also assigned to all individuals. Closing the gaps in initial distribution and dynamic evolution of health shrinks the life expectancy gradient to less than a year and the consumption gradient to \$458,000 for high school dropouts.¹⁷ The life expectancy gap for high school graduates becomes slightly negative (primarily due to a larger share of females in the group) and the NPVC gap falls to \$286,000. This means that eliminating all health differences across education groups closes the gradient in NPVC at age sixty by an estimated 10.5% for high school dropouts and 11.5% for high school graduates.

3.4 Sensitivity

Table 4 presents the robustness of our decomposition results to a number of alternate modeling assumptions. The table reports the percent of the NPVC gradient for high school dropouts closed due to counterfactually eliminating all health differences with college graduates. The benchmark results are repeated in the first row, followed by the alternate scenarios. Results are broken into the share explained by health transitions, initial health conditions, and both transitions and initial conditions combined.

¹⁶Calculations are analogous for high school graduates. Due to lack of observations we group together the black and other race categories. We also decrease self-rated health by one for those with a college degree in the few sub-group/health state combinations that contain no high school dropouts.

¹⁷The remaining life expectancy gradient of 0.64 exists due to differential initial morbidity distributions (conditional on self-rated health) within education groups and due to selection effects across education groups on the basis of race and gender.

3.4.1 Non-imputed data—We first check the sensitivity of our results to using only non-imputed data for estimating all of our panel VAR model equations. The cost of this estimation strategy is a loss in precision and potential bias due to any systematic pattern in missing observations. We continue to use the imputed data for initial age sixty conditions for simulations to ensure there is enough data to conduct our decomposition exercises and so average educational gradient estimates remain representative of the larger population. The results from this analysis are given in the second row of Table 4. Relative to the benchmark of 10.5%, the total share of the education gradient closed falls somewhat to 8.4% when using only non-imputed data for model estimation. This reduction is due to small declines in both the share closed by health transitions and initial health conditions at age sixty.

3.4.2 Birth cohort—The third row of Table 4 shows results when simulations are run for War Babies and Baby Boomers (born 1942-59) as opposed to the benchmark initial HRS cohort (born 1931-41).¹⁸ Note that less longitudinal data is available for the younger cohorts and projections are required further into the future. Nonetheless, in the younger cohort, the total share of the gradient closed by eliminating health disparities is only slightly smaller (9.5%) than the older benchmark cohort. This decline is mostly driven by a fall in the share closed by eliminating health transitions after age sixty.

3.4.3 Interest rate—The next two rows show the sensitivity of our results to the choice of interest rate used in calculating NPVC. The percent of NPVC gap closed is somewhat lower when using a higher discount rate of 5% compared to our benchmark rate of 3%. This decline is primarily driven by a decline in the importance of health transitions as future realizations of outcomes are discounted more heavily. On the other hand, when future consumption is not discounted at all, health transitions close 1.3 percentage points more of the consumption gap as compared to the benchmark, driving the total gap closed up to 12.3%.

3.4.4 Time trend—In our model estimation, we do not allow for any innovations in health or growth in consumption, due to macroeconomic changes, past the final wave of the survey (2014). Here, we test the sensitivity of our results to allowing time trends in the future. Specifically, for the models of self-rated health, morbidities, and mortality, we replace time dummies with a linear time trend that we assume continues past 2014. In the consumption model, replacing time dummies with a linear trend is empirically more problematic due to the timing of the great recession and subsequent recovery in relation to the survey data. So we keep time dummies in the consumption equation but assume that consumption grows at a rate of approximately 3% annually after 2014. As shown in Table 4, inclusion of linear time trends increases the share of the gradient closed to 11.8%, primarily by increasing the importance of future health transitions.

3.4.5 Parameter heterogeneity across education groups—Our benchmark model also assumes homogeneity of all parameters across education groups. We relax this assumption by estimating the model separately for each of the education groups. This

¹⁸War Babies and Baby Boomers are simulated together as a single cohort to ensure adequate sample size.

allows, for example, self-rated health to influence consumption differently for high school versus college graduates. The cost of this approach is a loss in precision as the model is estimated entirely independently on education specific sub-samples. As shown in Table 4, combined results are very similar to the benchmark when incorporating heterogeneity in parameters along the education dimension.

3.4.6 Out-of-sample projections—It is also of interest to test if our model is valid for out-of-sample projections. To this end, we exclude the HRS cohort from the estimation sample and re-estimate the model parameters.¹⁹ Simulations are then conducted for the HRS cohort which are completely out-of-sample. The out-of-sample simulations continue to match the data well suggesting our model may generalize to some extent (comparison figures available on request). As shown in the last row of Table 4, overall results are similar to our benchmark estimation strategy.

4 Conclusion

We estimated a panel VAR model using data from the Health and Retirement Study to understand the joint evolution of health and consumption at older ages. We used the estimated model and empirical joint distribution at age sixty to simulate life-cycle paths and construct a measure of the net present value of expected elderly consumption. We found an education gradient of \$512,000 in NPVC between the college educated and high school dropouts at age sixty. In other terms, sixty year-olds with a college degree could expect three times the remaining lifetime consumption of high school dropouts on average. We also estimated a life expectancy gradient of 5.3 years between these levels of education at age sixty. Gradients were smaller but still substantial between college and high school graduates—a 2.3 year gap in life expectancy and \$323,000 in NPVC.

Counterfactual experiments indicated that two-year mortality transitions, conditional on current health, do little to explain the education gradient in consumption among the elderly. However, the differential evolution of self-rated health and morbidities could each close in the range of 1-3% of the consumption gaps between the college educated and lower attainment groups. Combined, we find that health and mortality transitions could close 4-7% of the gradients whereas the initial health distribution could close 4-6%. Taken together, we estimate that eliminating all elderly health differences across education groups could close 10-12% of the education-consumption gradient.

Our study is not without limitations. First, it is important to note that this is primarily a descriptive analysis and we cannot make strong claims about causal inference. We cannot rule out selection into education categories and do not allow for any potential effect of elderly consumption on health. This type of reverse causality or any other source of endogeneity could potentially bias our estimates of the effect of health evolution on consumption, hence our decomposition results. Moreover, our counterfactual experiments should be interpreted as the potential decline in the consumption gradient in response to an unexpected improvement in the health processes of elderly with less than a college degree. If

¹⁹This approach requires dropping cohort dummies from the model.

counterfactual changes were anticipated early in life, there may also be endogenous changes in consumption patterns over the entire life-cycle. Finally, we do not allow for heterogeneities in medical innovations over time. For instance, we do not account for the relative increase in survival due to improvements in treatments for specific diseases like cancer. Another example would be that we continue to assume the onset of diabetes has a time invariant effect on latent self-rated health. If future medical advances drastically lower the impact of diabetes on self-rated health, this change would not be explicitly captured. Nonetheless, our study is an important first step towards understanding the role of health differences in explaining consumption inequality. Addressing some of the aforementioned empirical issues leaves room for important future research in this area.

Appendix: Additional tables and figures

Table 5:

Estimation sample descriptive statistics by highest level of education

	<HS	HS	College
Individuals	11,396	17,971	6,515
Observations	64,218	111,332	41,056
Age (mean)	67.93	64.55	62.90
Annual consumption (\$1000s, mean)	16.34	25.60	40.79
Hypertension (%)	55.08	48.71	41.45
Diabetes (%)	21.88	15.57	12.25
Cancer (%)	11.58	11.85	12.36
Lung disease (%)	12.66	7.98	4.17
Heart disease (%)	27.04	20.38	15.50
Stroke (%)	10.58	6.74	4.57
Psyche problem (%)	18.46	13.72	12.14
Arthritis (%)	58.46	51.30	40.31
Difficulty with ADLs (%)	36.32	21.26	13.07
No Morbidities (%)	13.44	19.32	26.99
Self-rated health (mean)			
Poor	16.39	6.63	3.19
Fair	29.56	16.83	9.20
Good	29.48	32.67	26.03
Very good	18.22	31.68	39.69
Excellent	6.35	12.18	21.89
Male (%)	44.09	40.61	53.44
Race (%)			
White	76.50	88.44	89.89
Black	15.74	8.21	5.52
Other	7.76	3.34	4.58
Birth cohort (%)			
AHEAD	23.17	13.09	8.39
CODA	14.76	10.85	9.42

	<HS	HS	College
HRS	31.08	29.18	24.58
War Babies	14.26	19.64	20.99
Baby Boomers	16.74	27.24	36.62

Notes: Mean and percentage estimates use base year sampling weights. Consumption is reported in real 2010 dollars.

Figure 8: Mean of life-cycle morbidity profiles by education

Notes: “Data” plots mean of all available data (inclusive of imputed missing values) in HRS by two-year age interval and cohort. “Simulated” plots mean of expected simulated outcome for each observation in the data (i.e. the expected outcome for each person-year observation in the data).

Figure 9: Mean of life-cycle morbidity profiles by education

Notes: “Data” plots mean of all available data (inclusive of imputed missing values) in HRS by two-year age interval and cohort. “Simulated” plots mean of expected simulated outcome for each observation in the data (i.e. the expected outcome for each person-year observation in the data).

Figure 10: Mean of life-cycle ADL limitation, self-rated health, and mortality profiles by education

Notes: “Data” plots mean of all available data (inclusive of imputed missing values) in HRS by two-year age interval and cohort. “Simulated” plots mean of expected simulated outcome for each observation in the data (i.e. the expected outcome for each person-year observation in the data).

Figure 11: Mean and standard deviation of consumption life-cycle profiles by education

Notes: Top panel plots mean and bottom panel standard deviation of consumption. “Data” plots mean (top panel) and standard deviation (bottom panel) of all available data (inclusive of imputed missing values) in HRS by two-year age interval and cohort. “Simulated” in top panel plots mean of expected simulated outcome for each observation in the data (i.e. the expected outcome for each person-year observation in the data). “Simulated” in the bottom panel plots mean of standard deviations of simulated outcome for each observation in the data (i.e. the mean of standard deviations calculated for each of the 5,000 simulation runs using only person-year observations that also appear in the data).

Table 6:

Model estimates for morbidities

Variable	Hypertension		Diabetes		Cancer		Lung disease		Heart disease		Stroke		Psych		Arthritis	
	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
Lag Hyper			0.267	0.034	-0.049	0.040	0.059	0.041	0.116	0.033	0.081	0.041	0.157	0.037	0.086	0.033
Lag Diab	0.268	0.051			0.073	0.047	0.003	0.052	0.034	0.044	0.058	0.053	-0.017	0.052	0.049	0.046
Lag Cancer	-0.031	0.052	0.002	0.054			0.049	0.060	-0.120	0.052	-0.026	0.059	-0.046	0.060	0.028	0.052
Lag Lung	0.061	0.060	0.043	0.059	0.106	0.059					0.002	0.064	0.132	0.062	0.148	0.065
Lag Heart	0.085	0.044	0.080	0.041	0.028	0.042	0.191	0.042	0.245	0.052	0.164	0.041	0.085	0.042	0.113	0.042
Lag Stroke	0.066	0.068	-0.020	0.064	-0.023	0.061	0.016	0.063	0.053	0.055			0.263	0.053	-0.045	0.063
Lag Psych	0.098	0.054	0.068	0.053	-0.073	0.060	0.088	0.057	0.109	0.049	0.155	0.055			0.248	0.055
Lag Arthritis	0.088	0.030	-0.007	0.034	-0.009	0.035	0.128	0.038	0.072	0.031	-0.007	0.037	0.110	0.036		
Lag ADL	0.070	0.033	0.026	0.036	-0.003	0.036	0.106	0.036	0.058	0.031	0.177	0.033	0.215	0.033	0.157	0.037
Lag Health 2	0.014	0.035	-0.022	0.033	-0.036	0.035	-0.081	0.033	-0.101	0.030	-0.126	0.033	-0.199	0.031	-0.088	0.038
Lag Health 3	0.011	0.036	-0.028	0.035	-0.051	0.037	-0.140	0.036	-0.158	0.032	-0.223	0.035	-0.274	0.034	-0.112	0.039
Lag Health 4	-0.026	0.038	-0.109	0.038	-0.077	0.039	-0.294	0.041	-0.244	0.034	-0.252	0.039	-0.356	0.039	-0.153	0.041
Lag Health 5 (best)	-0.109	0.043	-0.243	0.047	-0.115	0.046	-0.452	0.057	-0.289	0.041	-0.368	0.051	-0.465	0.052	-0.252	0.045
Lag2 Hyper			0.029	0.034	0.069	0.040	-0.072	0.041	0.048	0.033	0.063	0.041	-0.096	0.037	0.018	0.033
Lag2 Diab	-0.085	0.055			-0.081	0.051	-0.092	0.055	0.114	0.046	0.083	0.055	0.037	0.055	-0.045	0.049
Lag2 Cancer	0.006	0.056	0.002	0.058			0.052	0.064	0.138	0.056	0.007	0.063	0.070	0.064	0.037	0.056
Lag2 Lung	-0.130	0.066	-0.054	0.064	0.030	0.064			-0.087	0.058	0.052	0.069	-0.010	0.067	-0.048	0.072
Lag2 Heart	-0.029	0.047	0.005	0.043	-0.004	0.044	-0.058	0.044			-0.014	0.043	-0.052	0.045	-0.043	0.045
Lag2 Stroke	-0.027	0.076	0.021	0.070	0.024	0.067	-0.031	0.069	0.079	0.061			-0.191	0.059	0.046	0.070
Lag2 Psych	-0.044	0.058	-0.071	0.057	0.068	0.064	0.049	0.060	-0.041	0.052	-0.041	0.058			-0.113	0.060
Lag2 Arthre	-0.048	0.030	0.008	0.034	0.058	0.035	-0.014	0.036	0.031	0.030	-0.003	0.037	-0.021	0.035		
Lag2 ADL	-0.075	0.036	0.035	0.039	0.004	0.038	-0.017	0.037	0.021	0.032	-0.092	0.035	-0.068	0.034	-0.055	0.042
Lag2 Health 2	-0.039	0.037	-0.056	0.034	-0.047	0.037	-0.082	0.035	0.001	0.033	-0.074	0.035	-0.083	0.033	0.053	0.041
Lag2 Health 3	-0.040	0.038	-0.075	0.036	-0.013	0.038	-0.130	0.037	-0.008	0.034	-0.072	0.037	-0.144	0.036	0.085	0.043
Lag2 Health 4	-0.062	0.040	-0.121	0.039	-0.011	0.040	-0.190	0.041	-0.026	0.037	-0.060	0.041	-0.232	0.040	0.061	0.044
Lag2 Health 5	-0.075	0.043	-0.150	0.046	0.002	0.046	-0.293	0.053	-0.096	0.042	-0.097	0.050	-0.288	0.050	-0.017	0.048
CODA	-0.061	0.035	-0.029	0.039	-0.014	0.036	-0.031	0.041	-0.030	0.032	-0.003	0.036	0.030	0.039	-0.165	0.036

Variable	Hypertension		Diabetes		Cancer		Lung disease		Heart disease		Stroke		Psych		Arthritis	
	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
HRS	-0.084	0.049	-0.024	0.053	-0.071	0.051	-0.096	0.058	-0.002	0.046	-0.036	0.052	0.037	0.055	-0.193	0.049
War Babies	-0.103	0.066	0.028	0.072	-0.061	0.071	-0.105	0.080	0.039	0.063	0.042	0.076	0.162	0.075	-0.087	0.066
Boomers	-0.153	0.081	0.032	0.088	-0.062	0.087	-0.145	0.100	0.071	0.078	0.031	0.094	0.280	0.093	-0.093	0.082
Black	0.197	0.022	0.094	0.021	-0.028	0.022	-0.152	0.026	-0.122	0.020	0.054	0.024	-0.189	0.025	0.006	0.020
Other race	0.036	0.032	0.214	0.034	-0.199	0.044	-0.139	0.045	-0.120	0.036	-0.136	0.049	-0.038	0.041	-0.056	0.033
Female	0.017	0.014	-0.108	0.016	-0.202	0.016	-0.075	0.018	-0.172	0.014	-0.067	0.017	0.128	0.017	0.153	0.014
<HS	0.131	0.021	0.166	0.024	-0.052	0.024	0.190	0.030	0.007	0.022	0.012	0.028	0.055	0.027	0.043	0.022
HS	0.067	0.018	0.070	0.021	-0.024	0.021	0.128	0.027	0.037	0.019	0.041	0.024	-0.001	0.024	0.025	0.018
Constant	-1.611	0.077	-2.044	0.084	-1.946	0.086	-2.007	0.096	-1.808	0.078	-2.357	0.101	-1.850	0.088	-1.266	0.079

Notes: Multivariate probit results with dependent variable across columns. Regressions also include dummies for age and year.

Table 7:

Model estimates ADLs, self-rated health, mortality, and consumption

Variable	ADLs		Self-rated health		Mortality		Consumption	
	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
Hyper	0.044	0.031	−0.273	0.015	0.083	0.027	−0.012	0.012
Diab	0.099	0.040	−0.277	0.019	0.122	0.033	−0.036	0.018
Cancer	0.029	0.045	−0.690	0.020	0.662	0.027	0.034	0.022
Lung	0.185	0.049	−0.478	0.024	0.412	0.033	−0.008	0.027
Heart	0.061	0.034	−0.496	0.017	0.178	0.025	−0.006	0.019
Stroke	0.385	0.047	−0.501	0.023	0.238	0.030	−0.073	0.020
Psych	0.344	0.044	−0.418	0.022	0.238	0.030	−0.088	0.017
Arthritis	0.191	0.026	−0.214	0.015	−0.019	0.024	0.013	0.014
ADL			−0.673	0.014	0.343	0.020	−0.070	0.018
Health 2	−0.234	0.031			−0.335	0.017	0.061	0.019
Health 3	−0.481	0.032			−0.545	0.019	0.082	0.021
Health 4	−0.652	0.034			−0.666	0.023	0.117	0.021
Health 5 (best)	−0.738	0.041			−0.657	0.033	0.147	0.025
Lag Hyper	−0.026	0.031	0.143	0.020	−0.034	0.027	−0.001	0.012
Lag Diab	0.009	0.042	0.102	0.026	0.060	0.035	0.012	0.016
Lag Cancer	−0.025	0.048	0.547	0.029	−0.453	0.030	−0.009	0.023
Lag Lung	−0.036	0.054	0.210	0.034	−0.103	0.036	0.000	0.021
Lag Heart	−0.057	0.036	0.297	0.023	−0.018	0.026	0.007	0.018
Lag Stroke	−0.178	0.054	0.363	0.033	−0.047	0.033	0.004	0.019
Lag Psych	−0.120	0.048	0.233	0.031	−0.135	0.033	0.037	0.019
Lag Arthre	0.068	0.026	0.119	0.018	−0.085	0.024	−0.011	0.013
Lag ADL			0.340	0.018	−0.116	0.020	0.013	0.017
Lag Health 2	−0.162	0.033	0.630	0.014	−0.038	0.019	0.016	0.011
Lag Health 3	−0.286	0.034	1.139	0.015	−0.064	0.021	0.015	0.012
Lag Health 4	−0.408	0.036	1.668	0.016	−0.103	0.024	0.015	0.011
Lag Health 5 (best)	−0.483	0.041	2.290	0.019	−0.137	0.032	0.020	0.015
CODA	0.040	0.029	0.019	0.015	−0.035	0.023		
HRS	0.026	0.042	0.022	0.021	−0.067	0.034		
War Babies	0.055	0.058	0.025	0.028	−0.131	0.051		
Boomers	0.131	0.072	−0.052	0.034	−0.099	0.062		
Black	0.090	0.019	−0.061	0.009	0.071	0.016		
Other race	0.034	0.033	−0.111	0.015	−0.065	0.031		
Female	0.014	0.013	0.031	0.006	−0.210	0.012		
<HS	0.138	0.020	−0.236	0.009	0.025	0.019		
HS	0.041	0.018	−0.112	0.008	0.044	0.017		
Constant	−0.955	0.070			−1.705	0.293		
Lag Con							0.173	0.005
Lag2 Con							0.085	0.005

Notes: Dependent variable across columns. Multivariate probit results reported for ADLs as dependent outcome. Standard (ordered) probit results reported for self-rated health and mortality as dependant outcomes. Linear dynamic panel estimates reported for consumption as outcome. All regressions also include dummies for age and year. Regression for self-rated health also includes second lag for all morbidities.

Table 8:

Morbidity shock covariance matrix (Σ)

	Hyper	Diabetes	Cancer	Lung	Heart	Stroke	Psych	Arthritis	ADLs
Hyper	1.00	0.26	0.04	0.08	0.28	0.29	0.14	0.09	0.09
Diabetes	0.26	1.00	0.07	0.05	0.09	0.14	0.06	0.03	0.08
Cancer	0.04	0.07	1.00	0.12	0.03	0.06	0.12	0.05	0.13
Lung	0.08	0.05	0.12	1.00	0.22	0.11	0.17	0.09	0.19
Heart	0.28	0.09	0.03	0.22	1.00	0.28	0.16	0.10	0.14
Stroke	0.29	0.14	0.06	0.11	0.28	1.00	0.22	0.10	0.40
Psych	0.14	0.06	0.12	0.17	0.16	0.22	1.00	0.15	0.29
Arthritis	0.09	0.03	0.05	0.09	0.10	0.10	0.15	1.00	0.26
ADLs	0.09	0.08	0.13	0.19	0.14	0.40	0.29	0.26	1.00

Table 9:

NPVC inequality decomposition

	Theil inequality index		
	Within group	Between group	Total
Education	0.18	0.07	0.25
Race	0.24	0.01	0.25
Gender	0.25	0.00	0.25

References

- Aguiar M and Bils M (2015). Has consumption inequality mirrored income inequality? *American Economic Review*, 105(9):2725–56.
- Attanasio O, Hurst E, and Pistaferri L (2014). The evolution of income, consumption, and leisure inequality in the United States, 1980–2010 In *Improving the Measurement of Consumer Expenditures*, pages 100–140. University of Chicago Press.
- Bloom DE, Canning D, and Sevilla J (2004). The effect of health on economic growth: a production function approach. *World development*, 32(1):1–13.
- Blundell R and Preston I (1998). Consumption inequality and income uncertainty. *The Quarterly Journal of Economics*, 113(2):603–640.
- Born B and Breitung J (2016). Testing for serial correlation in fixed-effects panel data models. *Econometric Reviews*, 35(7):1290–1316.
- Case A and Deaton A (2017). Mortality and morbidity in the 21st century. *Brookings papers on economic activity*, 2017:397. [PubMed: 29033460]
- Chetty R, Stepner M, Abraham S, Lin S, Scuderi B, Turner N, Bergeron A, and Cutler D (2016). The association between income and life expectancy in the United States, 2001–2014. *JAMA*, 315(16): 1750–1766. [PubMed: 27063997]
- Conti G, Heckman J, and Urzua S (2010). The education-health gradient. *American Economic Review*, 100(2):234–38. [PubMed: 24741117]
- Cutler DM and Katz LF (1992). Rising inequality? Changes in the distribution of income and consumption in the 1980s. NBER Working Paper No. 3964.

- Cutler DM, Katz LF, Card D, and Hall RE (1991). Macroeconomic performance and the disadvantaged. *Brookings papers on economic activity*, 1991(2):1–74.
- Cutler DM and Lleras-Muney A (2010). Understanding differences in health behaviors by education. *Journal of health economics*, 29(1):1–28. [PubMed: 19963292]
- De Vos I, Everaert G, Ruysen I, et al. (2015). Bootstrap-based bias correction and inference for dynamic panels with fixed effects. *Stata Journal*, 15(4):986–1018(33).
- Deaton A and Paxson C (1994). Intertemporal choice and inequality. *Journal of political economy*, 102(3):437–467.
- Deaton AS and Paxson CH (1998). Aging and inequality in income and health. *The American Economic Review*, 88(2):248–253.
- Dynarski S, Gruber J, Moffitt RA, and Burtless G (1997). Can families smooth variable earnings? *Brookings papers on economic activity*, 1997(1):229–303.
- Ettehad D, Emdin CA, Kiran A, Anderson SG, Callender T, Emberson J, Chalmers J, Rodgers A, and Rahimi K (2016). Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. *The Lancet*, 387(10022):957–967.
- Everaert G and Pozzi L (2007). Bootstrap-based bias correction for dynamic panels. *Journal of Economic Dynamics and Control*, 31(4):1160–1184.
- Goldman DP and Smith JP (2002). Can patient self-management help explain the ses health gradient? *Proceedings of the National Academy of Sciences*, 99(16):10929–10934.
- Griliches Z and Mason WM (1972). Education, income, and ability. *Journal of political Economy*, 80(3, Part 2):S74–S103.
- Heathcote J, Perri F, and Violante GL (2010). Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006. *Review of Economic Dynamics*, 13(1):15–51.
- Honaker J and King G (2010). What to do about missing values in time-series cross-section data. *American Journal of Political Science*, 54(2):561–581.
- Houthakker HS (1959). Education and income. *The Review of Economics and Statistics*, pages 24–28.
- Hurd MD and Rohwedder S (2007). Economic well-being at older ages: Income and consumption-based poverty measures in the HRS. RAND Corporation Working Paper WR-410.
- Idler EL and Benyamini Y (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior*, pages 21–37. [PubMed: 9097506]
- Katz LF et al. (1999). Changes in the wage structure and earnings inequality. *Handbook of Labor Economics*, 3:1463–1555.
- Kenkel DS (1991). Health behavior, health knowledge, and schooling. *Journal of Political Economy*, 99(2):287–305.
- Krueger D and Perri F (2003). On the welfare consequences of the increase in inequality in the united states. *NBER macroeconomics annual*, 18:83–121.
- Krueger D and Perri F (2006). Does income inequality lead to consumption inequality? evidence and theory. *The Review of Economic Studies*, 73(1):163–193.
- Meara ER, Richards S, and Cutler DM (2008). The gap gets bigger: changes in mortality and life expectancy, by education, 1981–2000. *Health Affairs*, 27(2):350–360. [PubMed: 18332489]
- Miller R and Bairoliya N (2018). Health, longevity, and welfare inequality of the elderly. PGDA Working Paper No. 159.
- Morgan J and David M (1963). Education and income. *The Quarterly Journal of Economics*, 77(3): 423–437.
- Mullahy J (2016). Estimation of multivariate probit models via bivariate probit. *Stata Journal*, 16(1): 37–51.
- National Academies of Sciences, Engineering, and Medicine (2015). The growing gap in life expectancy by income: Implications for federal programs and policy responses. National Academies Press, Washington, DC.
- Nickell S (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6):1417–1426.
- Pijoan-Mas J and Ríos-Rull J-V (2014). Heterogeneity in expected longevity. *Demography*, 51(6): 2075–2102. [PubMed: 25391225]

- Shastry GK and Weil DN (2003). How much of cross-country income variation is explained by health? *Journal of the European Economic Association*, 1(2-3):387–396.
- Stenholm S, Pentti J, Kawachi I, Westerlund H, Kivimäki M, and Vahtera J (2014). Self-rated health in the last 12 years of life compared to matched surviving controls: the health and retirement study. *PloS one*, 9(9):e107879. [PubMed: 25237814]
- Wursten J et al. (2016). Xtpptest: Stata module to perform Born & Breitung bias-corrected LM-based test for serial correlation. *Statistical Software Components*.

Figure 1:
Contemporaneous model

Figure 2: Select estimation results for self-rated health, consumption, mortality, and stroke outcomes

Notes: Dependent variables across columns. Average marginal effects on the probability of an outcome reported for probit results—poor health, mortality, and stroke. Contemporaneous associations reported for poor health and consumption as dependent variables. Lagged associations reported for mortality and stroke. Good health coefficients use poor health state as reference group. Spikes indicate 95% confidence intervals.

Figure 3: Response to incidence of heart disease at age sixty-two

Notes: Results plot percentage difference in expected outcomes with the exogenous onset of heart disease at age sixty-two relative to remaining without heart disease at sixty-two.

Sample includes all individuals in the simulation sample without heart disease at age sixty.

Expected health outcomes are conditional on survival.

Figure 4: Average marginal effect of education on health probabilities

Notes: Dependent variables across rows. Less than high school education is the reference group. Spikes indicate 95% confidence intervals.

Figure 5:
Average life expectancy by decile of age sixty annual consumption

Figure 6: Average life-cycle profiles by education groups

Notes: Consumption profiles are expected values conditional on survival.

Figure 7: Change in expected outcomes for <HS group when given the health transitions of college graduates

Notes: Results plot percentage difference in expected outcomes with the health transitions of college graduates relative to less than high school education. Sample includes all individuals in the simulation sample with less than high school education. Outcomes other than unconditional consumption are conditional on survival.

Table 1:

Simulation sample age sixty descriptive statistics by highest level of education

	<HS	HS	College
Individuals	1,971	3,422	1,151
Consumption (\$1000s, mean)	17.87	28.80	45.49
Hypertension (%)	44.53	39.41	36.13
Diabetes (%)	16.57	10.52	9.55
Cancer (%)	7.23	8.16	6.64
Lung disease (%)	11.19	6.42	2.19
Heart disease (%)	17.96	13.53	12.53
Stroke (%)	4.77	3.40	2.19
Psyche problem (%)	15.60	8.56	5.29
Arthritis (%)	54.61	46.31	35.85
Difficulty with ADLs (%)	26.67	12.79	7.14
No Morbidities (%)	18.83	25.84	32.35
Self-rated health (mean)			
Poor	15.16	5.11	1.70
Fair	26.76	13.70	5.80
Good	30.91	29.94	26.01
Very good	19.34	33.98	41.00
Excellent	7.82	17.26	25.48
Male (%)	44.39	42.28	58.49
Race (%)			
White	79.34	89.83	90.98
Black	16.08	7.97	5.00
Other	4.58	2.20	4.02

Notes: Mean and percentage estimates use base year sampling weights. Annual consumption reported in real 2010 dollars.

Table 2:

Education gradients at age sixty

	<HS (1)	HS (2)	College (3)	<i>lhs</i> (4)	<i>hs</i> (5)
Life expectancy	18.44	21.47	23.74	5.30	2.27
Consumption (NPVC)	2.55	4.44	7.67	5.12	3.23

Notes: Consumption NPVC indicates net present value of remaining lifetime consumption, reported in 100,000s of 2010 dollars. *lhs* reports difference between column (3) and column (1). *hs* reports difference between column (3) and column (2).

Table 3:

Decomposition of education gradient in consumption at age sixty

	Life expectancy				NPVC			
	<HS (1)	<i>lhs</i> (2)	HS (3)	<i>hs</i> (4)	<HS (5)	<i>lhs</i> (6)	HS (7)	<i>hs</i> (8)
<i>Baseline</i>	18.44	5.30	21.47	2.27	2.55	5.12	4.44	3.23
<i>Health transitions</i> [†]								
Mortality	18.74	5.00	22.02	1.72	2.57	5.10	4.50	3.17
Self-rated health	19.81	3.93	22.13	1.61	2.69	4.98	4.53	3.14
Morbidities	19.19	4.55	21.91	1.83	2.62	5.05	4.50	3.17
All health	20.92	2.82	23.16	0.58	2.79	4.88	4.66	3.01
<i>Initial health</i> [‡]	20.53	3.21	22.24	1.50	2.84	4.83	4.59	3.08
<i>Transitions + initial</i> [*]	23.10	0.64	23.92	-0.18	3.09	4.58	4.81	2.86

Notes: NPVC indicates net present value of remaining lifetime consumption, reported in 100,000s of 2010 dollars. Columns (2) and (6) report the education gradient in life expectancy and expected consumption at age sixty for <HS education group. Columns (4) and (8) report analogous gradients for high school graduates.

[†] Reports average outcomes by education when all are given the particular health transition of the college group.

[‡] Reports average outcomes by education after re-weighting to match the distribution across self-rated health states of college graduates at age sixty.

^{*} Reports average outcomes by education when all are given all health transitions and initial health distribution of the college group as defined above.

Table 4:

Sensitivity results: percent of consumption gradient closed for <HS sample

	Health transitions	Initial conditions	Both
Benchmark	4.7	5.7	10.5
Non-imputed data only	4.0	4.2	8.4
War Babies / Baby Boomers	3.9	5.4	9.5
5% interest rate: $r = 0.05$	4.0	5.7	9.7
No discounting: $r = 0$	6.0	6.0	12.3
Linear time trends	5.6	5.8	11.8
Education heterogeneity	3.0	5.3	10.6
Out-of-sample	4.6	6.0	10.8

Notes: Results shown are the percent of NPVC gradient closed by eliminating differences in health transitions after age sixty (column one), in initial health transitions (column two), both transitions and initial conditions combined (column three).