On The Rise of Health Spending and Longevity

Raquel Fonseca
RAND

Pierre-Carl Michaud
RAND, IZA and NETSPAR

Titus Galama
RAND

Arie Kapteyn
RAND, IZA and NETSPAR

April 2010

Abstract

We use a calibrated stochastic life-cycle model of endogenous health spending, asset accumulation and retirement to investigate the causes behind the increase in health spending and life expectancy over the period 1965-2005. We estimate that technological change and the increase in the generosity of health insurance on their own may explain independently 53% of the rise in health spending (insurance 29% and technology 24%) while income explains less than 10%. By simultaneously occurring over this period, these changes may have led to a “synergy” or interaction effect which helps explain an additional 37% increase in health spending. We estimate that technological change, taking the form of increased productivity at an annual rate of 1.8%, explains 59% of the rise in life expectancy at age 50 over this period while insurance and income explain less than 10%.

Keywords: demand for health, health spending, insurance, technological change, longevity.

JEL Codes: I10, I38, J26

---

1 Corresponding Author: Raquel Fonseca, 1776 Main Street, CA 90401-2138; fonseca@rand.org. We would like to thank Michael Hurd, Arthur van Soest, Peter Kooreman, Rob Alessie, Dana Goldman, Darius Lakdawalla, Eric French, John Jones, Luigi Pistaferri and seminar participants both at Marshall School of Business and Economic Department at USC, Tilburg, NETSPAR, Health and Macroeconomics Conference at UCSB, and the PSID Conference at the University of Michigan for their helpful comments on previous drafts of this paper. This research was supported by the National Institute on Aging, under grants Health and Economic Status in Older Populations P01AG008291, Economic and Health Determinants of Retirement Behavior P01AG022481 and Robust Integration Modeling of Health, Wealth, and Disability R01AG030824. Errors are our own.
1. Introduction

According to a recent report from the Congressional Budget Office (CBO, 2008), per capita health spending in 2005 was 6 times what it was in 1965. As a fraction of per capita income, health spending in the U.S. has grown from 4% to 16%. Newhouse (1992) concludes that more than half, perhaps even 65%, of the change is probably attributable to technological change, with more generous health insurance and the growth in income accounting for at most one third of the rise in health spending.

Over the same period, life expectancy has increased considerably, particularly among the elderly. In 2005, a 50 year old man could expect to live 28.9 years compared to 22.9 in 1965. Cutler, Deaton and Lleras-Muney (2006) argue that technological change is the leading explanation for the increase in longevity since evidence from the Health Insurance Experiment (HIE, Manning et al., 1987) and other recent studies on the SES-health gradient (Adams et al., 2003; Smith, 2007) suggest that neither growth in income nor the spread of health insurance can explain changes in life expectancy of that magnitude.

In a recent study, Hall and Jones (2007) show that one can generate the growth in health spending and the resulting change in life expectancy solely as a result of growth in income using a model of endogenous health spending (Grossman, 1972; Ehrlich and Chuma, 1990). Instead of modeling other changes simultaneously, Hall and Jones show that with reasonable parameter estimates and implied value of longevity gains, one can generate the rise in health spending and life expectancy, assuming it is an optimal response to the growth in income.

We follow a similar modeling approach, generating health spending and longevity growth as an optimal response to changing circumstances. We introduce in our framework
health insurance and technological change as explicit possible competing alternatives to income growth in an environment that features uncertainty and institutions (Social Security, Medicare, etc). These alternative explanations may not only be competing but may also be reinforcing. The hypothesis that important interactions may arise from these competing alternatives has long been put forward (e.g. Weisbrod, 1991). But few papers have attempted to estimate the size of such interaction effects, let alone calibrate a model that would generate such effects.

In this paper, we calibrate a stochastic life-cycle model of endogenous health spending, asset accumulation and retirement with detailed modeling of health insurance, Social Security and taxation using micro-data from 2000 to 2005. Our calibration approach involves fitting life-cycle profiles of assets, medical expenditures, retirement and mortality in the recent past. We also calibrate the parameters of the model such that behavioral responses to changes in health insurance and income are similar to those reported in the literature. We then go back in time to 1965. We perform counterfactual simulations where we first calibrate the necessary productivity gains such that we reach 1965 health spending levels, given 1965 income and health insurance system. We find that a very plausible rate of productivity growth of 1.8% is necessary to achieve the 1965 health spending levels. Once we obtain simulated profiles for 1965, we introduce sequentially each of the competing explanations for the rise in health spending to understand their relative importance. We find that far from competing, these are reinforcing forces: as much as 37% of the overall increase in health spending between 1965 and 2005 is due to the fact that these occurred simultaneously. This interaction effect is not a “residual”. Although the change in productivity in order to attain 1965 levels is a “residual”, the interaction between
these changes is a endogenous product of the model. Simply put, nor the price or income elasticity of medical expenditures is independent of economic circumstances, in particular the productivity of medical care. Our results suggest that the increase in health insurance copayment has been as important as technological change while growth in real income over this period explains at most 10% of the rise in spending. Moreover, we find that both income and health insurance copayments can explain up to 10% of the rise in longevity while productivity gains from technology account for up to 60% of the gain in life expectancy.

In section 2 we present the model we used to generate those results. Section 3 presents the data and the calibration. In section 4 we perform the counterfactual simulations. We conclude in section 5.

2. Model

We build a model where an agent makes decisions over the life-cycle regarding health spending, consumption and retirement. The agent faces health risk, job risk and insurance coverage risk (up to the age of Medicare eligibility). The life-cycle framework is closest to the one considered by French and Jones (2007) with three important differences. First, we endogenize health spending. Health spending can affect health, which in turn affects future health, utility and earnings directly. In that sense, our modeling of health spending is closest to the human capital framework of Grossman (1972). Second, an important difference is that we do not allow for uncertainty in earnings and medical expenditures given health, job and insurance status. Once medical expenditures are

---

2 Other studies investigating endogeneous health investment include Blau and Gilleskie (2008), Galama et al. (2008), Khwaja (2009), Yogo (2009), Bajari, Hong and Khwaja (2009), Halliday, He and Zhang (2009), Juergen and Tran (2007), Suen (2009)
endogenous, there is little reason to think of the realization of health spending as stochastic conditional on a given health insurance, job and health status. In fact, they are still stochastic, *ex ante*, as the individual does not know yet his health, job and health insurance status. Finally, another important difference is that we assume labor supply to be inelastic until the agent chooses the age of retirement and that retirement is an absorbing state. We do this because labor supply plays a secondary role in our analysis.  

2.1. Decisions and Preferences

The agent makes decisions at discrete ages \( t = t_0, \ldots, T \) and dies (if alive) with certainty at age \( T \). At each age \( t \), we assume the agent to make three decisions \( d_t = (c_t, m_t, r_t) \). She chooses non-medical expenditures \( c_t \), medical expenditures \( m_t \) and her retirement status \( r_t \). If she retires, \( r_t = 1 \), otherwise \( r_t = 0 \). For our purposes, we will assume retirement is an absorbing state. Retirement is possible after age 54. At age 70, we assume everyone to be retired (few in the data work beyond that age). At age \( t \), she has a health state \( h_t \) assumed to take 6 values \( \{0, 1, 2, 3, 4, 5\} \) where zero denotes death and values 1 to 5 denote the self-reported health status scale \( \{\text{poor, fair, good, very good, excellent}\} \). We assume preferences are represented by the following utility function

\[
u(c_t, r_t) = \frac{c_t^{1-\sigma}}{1-\sigma} + \psi_h h_{t,>2} + e^{\psi_e + \psi_r} r_t^e.
\]

3 Unlike Hall and Jones (2007), we do not model overlapping generations. The main reason is that aging is not likely to have played an important role in the rise of health spending over the last half century. For example, Newhouse (1992) estimates that it could explain at most 3% of the rise in health spending. We also abstract from general equilibrium effects, such as innovation, changes in medical norms induced by changes in income and insurance etc.
The parameter \( \sigma \) captures the degree of relative risk aversion, \( \psi_h \) measures the direct utility of good health where \( h_{t,>2} = I(h_t > 2) \). Finally \( \psi_0, \psi_1 \) control the direct utility of leisure in retirement, which may vary with age.

We assume she has a bequest motive. When she dies at age \( t \), she leaves a bequest in the form of accumulated assets, \( a_t \), which yield utility

\[
b(a_t) = \phi_b \log(a_t). \tag{2}
\]

The parameter \( \phi_b \) measures the strength of the bequest motive.

### 2.2. Assets

At each age, the agent can decide to forego consumption and invest instead in a composite asset \( a_t \) with non-stochastic annual return \( \tau \). We specify a life-time budget constraint following Hubbard, Skinner and Zeldes (1995). At each age, cash-on-hand, \( x_t \), is the sum of last-period assets \( a_t \) and total net income \( y_t \) (including asset income) while total expenditures are the sum of non-medical expenditures and out-of-pocket medical expenditures \( oop_t \) (more detail is provided in section 2.5). We assume assets must be non-

---

4 For simplicity, we assume being in excellent, very good and good health yields the same utility.

5 In models with an endogenous time horizon, the level of utility matters. The marginal expected utility from extending the horizon is larger with a larger base utility level. Hall and Jones (2007) use age specific constants to match the distribution of health spending over the life-cycle. In our model, the base level of utility for someone working and in bad health is set to zero such that if \( \sigma > 1 \), base utility is negative. If that were the case with probability one going forward individuals would let their health decay and never invest in health (if we allowed them to have negative health spending, they would choose to die). But there is always a positive probability that they will live to be in better health (health greater than 2) and will definitely retire at some point before age 70, hence potentially having positive utility flows. Hence, there is always a “reason” to live in the model. When workers retire, the value of living longer rises slightly as the marginal utility of leisure is positive and increases with age ( \( \psi_1 > 1 \)). We cap this term at age 70, the maximum retirement age. Hence, apart for the immediate period around the time of retirement, there are no age-dependent constant terms in our utility function that allows rationalizing the life-cycle profile of health spending.
negative at all ages (the agent cannot borrow, particularly against pension wealth). End of period assets $a_{t+1}$ are defined as

$$a_{t+1} = x_t + tr_t - c_t - oop_t, \quad a_{t+1} \geq 0 \text{ for all } t.$$  \hfill (3)

Government transfers are given by $tr_t = \max(c_{\min} - x_t, 0)$ where $c_{\min}$ is a minimum consumption level guaranteed by the existence of subsistence programs such as, temporary assistance for needy families (TANF), food stamps, housing benefit and Supplemental Security Income (SSI).

### 2.3. Income and Pensions

An agent’s total net income is given by

$$y_t = tax(y^e_t, y^s_t, y^o_t, \tau a_t)$$  \hfill (4)

where $tax()$ is a tax function that depends on annual earnings $y^e_t$, social security benefits $y^s_t$, other (spousal) income $y^o_t$ and asset income $\tau a_t$.

**Earnings** - Earnings are assumed deterministic, conditional on age, employment and health status. If the individual is employed, we specify the log earnings equation as

$$\log y^e_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \sum_{s=2}^{5} \alpha_s h_{t,s}$$  \hfill (5)

where $h_{t,s} = I(h_t = s)$. Note that future earnings are uncertain due to uncertainty in health and employment. If unemployed or retired, earnings are zero.

**Retirement Income** - Prior to age 62, individuals do not receive retirement income if they decide to retire. They finance consumption through their accumulated savings until
they reach the age of 62. After age 62, they are assumed to be eligible for Social Security at which point they can start receiving benefits if they retire. As for those who elect not to retire prior to 62, they can delay retiring/claiming Social Security until they reach the age of 70 at which point everyone is assumed to retire (and claim). Social Security benefits are a function of lifetime earnings and the age of retirement. The measure of lifetime earnings used is the average indexed monthly earnings (AIME) denoted by $ae_t$, which represents the average of the 35 highest years of earnings. Before retirement, we approximate the evolution of AIME using the function

$$ae_{t+1} = G(ae_t, y_t^e, t)$$

(6)

where $G()$ is an age-specific non-linear regression of next period $ae_{t+1}$ on current $ae_t$ and earnings, $y_t^e$. This function is estimated on Social Security earnings records from the Health and Retirement Study (see section 3.1.1). To calculate Social Security benefits at the normal or full retirement age (NRA), one simply translates the AIME into the Primary Insurance Amount (PIA) by applying a concave function which is piece-wise linear. If the individual retires prior to the NRA, his benefit is reduced by an actuarial adjustment factor (arf). The same holds if the individual retires beyond the NRA. We follow French (2005) and compute, at the time when someone claims, a new AIME such that

$$pia(ae_t^*) = arf(t)pia(ae_t)$$

(7)

holds. The new AIME is then kept constant until death. Social Security benefits are given by $y_t^* = pia(ae_t^*)$ once an individual has claimed.

---

6. Claiming Social Security benefits and retiring are the same in the model.
7. The approximation performs well in practice. The main advantage is that we avoid having to include the entire earnings history in the state-space (French 2005).
Other Income – Other income is primarily the spouse’s income. It is assumed to be a function of age and own earnings, $y_i^o = o(t, y^c_i)$.

2.4. Health

An individual’s future health status is modeled as a dynamic ordered response process, which depends on current health status, medical expenditure, through a health investment function, and age. We first discuss the health investment function and then discuss the dynamic health process.

Health Investment – There is little guidance from the literature on the form of the health investment function. As Ehrlich and Chuma (1990) argue, a function featuring diminishing returns is probably realistic. We use the following production function

$$I(m_i, h_i, t) = \theta(t, h_i) \frac{(1 + m_i)^{1-\theta_i} - 1}{1 - \theta_i}. \quad (8)$$

The elasticity of this function with respect to medical expenditures is positive but decreases with $m_i$ if $\theta_i > 1$. As Hall and Jones (2006) show, the rate at which the elasticity of the health production function to medical expenditures decreases is directly related to the responsiveness of the optimal investment to income changes. Intuitively, with constant life-cycle consumption and mortality rate, an additional dollar spent on health care is more valuable if the life extension it buys is valued more than the utility today from spending that dollar on consumption. The more concave the utility function is relative to the production function, the more valuable is the dollar spent on health. Hall and
Jones (2006) consider a constant elasticity production function. We allow the elasticity to vary.

The productivity term is given by \( \theta_0(h_t, t) = e^{\theta_{0w} + \theta_{0h} h_t + \theta_0 t} \). We allow the productivity term to increase in bad health (fair and poor) and to increase with age. We interpret the age coefficient in the productivity term as an interaction between depreciation and productivity. As the body depreciates, there is more that technology can do to “restore” health.

**Health Process** – Observed health is assumed to be a function of a latent health index, which evolves over time as follows:

\[
\begin{align*}
    h_{t+1}^* & = \gamma_0(t) + I(m_t, h_t, t) + \sum_{h=2}^{5} \gamma_h I(h_t = h) + \varepsilon_t \tag{10}
\end{align*}
\]

where \( \varepsilon_t \) is a standard normally distributed health shock and \( \gamma_0(t) \) is a 4th order polynomial in age. The latent health index is transformed into observed health status through the following rule,

\[
    h_t = h \text{ if } \delta_{h-1} \leq h_t^* < \delta_h, h = 0, \ldots, 5.
\]

where \( \{\delta_h\}_{h=0}^{5} \) is a set of thresholds. Because of the normality assumption about the error in (10), this defines a dynamic ordered probit process.

### 2.5. Employment and Health Insurance

Employment and health insurance are stochastic and exogenous to the agent. We assume that four states are possible: employed or unemployed and with or without health insurance. The states are defined as \( e_t = \{0, 1, 2, 3\} \) where zero is unemployed without insurance, 1 is unemployed with health insurance, 2 is employed without health insurance,
and 3 is employed with insurance. We assume the transition probabilities from one state to another are first-order Markov and depend only on age. Upon retirement, the agents who had insurance coverage on their job retain it while those without have to wait until they become Medicare eligible. If the individual has health insurance \((e_t = 1, 3)\), we use a standard health insurance contract with a deductible and co-insurance rate to transform total into out-pocket expenditures. Hence, we assume the availability of retiree health insurance for those with insurance on the job at the time of retirement.\(^8\)

Out-of-pocket medical expenditures when the agent is not Medicare eligible are given by

\[
oop(t, e_t) = \begin{cases} 
\min(m_t, \mu_{1t}) + \mu_{2t} \max(m_t - \mu_{1t}, 0) & \text{if } e_t = 1, 3 \\
m_t & \text{if } e_t = 0, 2
\end{cases}
\]

(11)

where \((\mu_{1t}, \mu_{2t})\) are the corresponding deductible and co-insurance rate respectively if covered at age \(t\). If Medicare eligible (age 65 and older), we use a similar insurance contract as when \(e_t = 1, 3\). Because we follow annual medical expenditures and not doctor visits and hospital stays it is impossible to incorporate payment rules under Medicare Part A other than through an approximation.\(^9\) But Medicare Part B coverage which covers outpatient procedures is very similar to the median employer provided health insurance contract. We provide more details in section 3.2.

---

\(^8\) See French and Jones (2007) and Blau and Gilleskie (2008) for models where retiree health insurance is also stochastic. Both of these studies take insurance as given and beyond the control of the individual (except for the fact that the individual can change/quit a job with insurance). Marquis and Long (1995) analyze the price and income responsiveness of the demand for health insurance in the non-group market. They find a price (premium) elasticity of -0.3 to -0.4 and an income elasticity of 0.1. See Bajari, Hong and Khwaja (2009) for a study looking at health insurance decisions in a life-cycle framework.

\(^9\) See Blau and Gilleskie (2008) for a model where doctor visits rather than medical expenditures are modeled.
2.6. Solution of the Dynamic Problem

The agent maximizes the discounted sum of utility flows using a discount factor $\beta$. At each age, the agent can choose $d_t$ from a set of possible decisions $D(s_t)$ given the state space $s_t = (t, e_t, h_t, r_{t-1}, a_t, a_{e_t})$. The indirect utility of the agent at age $t$ is written as

$$v(s_t) = \max_{d_t \in D(s_t)} u(d_t) + \beta \left[ \sum_{h=0} \sum_e p_h(s_t, m_t) p_e(s_t) v(s_{t+1} | s_t, d_t) + p_0(s_t, m_t) h(a_{t+1}) \right]$$

where $p_h(s_t, m_t)$ denotes the probability of being in state $h_{t+1} = h$ at age $t+1$ and $p_e(s_t)$ denotes the probability of employment/health insurance status. Maximization in Equation (12) takes place subject to preferences and constraints (1)-(11).

The optimal solution: $[d^*(s_t), v^*(s_t)]_{i=0, \ldots, T}$ is obtained by backward recursion. We assume a terminal age equal to 105 and a starting age equal to 25. We assume everyone who is still alive at age 105 dies. The currency used throughout is thousands of $2004. We provide details on the solution method in the Technical appendix. We discretize continuous state and decision variables and search over multidimensional grids for the optimal solution.

2.7. Simulation

We simulate the life-cycle trajectories of 1,500 hypothetical individuals starting at age $t_0 = 25$ by first drawing initial conditions from the joint distribution of AIME, assets and health in the data (see section 3.3). We then apply the decision rules and update the state-space at each age until individuals reach age 105 (or die). We draw health and
employment/health insurance shocks from their respective distributions. We use linear interpolation for consumption, medical expenditure and retirement decisions.

3. Data and Calibration

We focus on describing the behavior of a relatively homogeneous group because the model allows for limited heterogeneity. We chose to calibrate the model so that it matches the behavior of non-hispanic white men without a college education. Many studies document preference heterogeneity across groups and growth in income has been very different for college graduates relative to the rest of the population. We could not find evidence that growth in health spending has been very different across socio-demographic groups.\(^{10}\)

We use three main sources of data. The first is the Panel Study of Income Dynamics (PSID), which we use for labor force and insurance status, health status, earnings and assets from age 25 till death. The PSID has information on medical expenditures at the household level. Second, as a more adequate source of data on individual medical expenditures, we used the Medical Expenditure Panel Survey (MEPS) for the years 2000 to 2003. Finally, we use the HRS for two purposes. First, we use Social Security earnings records of respondents to the Health and Retirement Study to estimate the relationship between AIME and earnings as described in equation (7). Second, the PSID seriously underestimates mortality in old age and has sparse cell counts over age 80. Hence, we use health status (which includes mortality) of respondents over age 50 from the HRS. Since cohort effects may exist, we select where possible data from the cohort born between 1936 to 1940 which is of retirement age 60 to 64 in the year 2000.

\(^{10}\) There is some evidence that growth in life expectancy may have been larger for higher educated individuals in the 1980s and 1990s. For example, Meera, Richards and Cutler (2008) report from the National Longitudinal Mortality Study (NLMS) life expectancy at age 25 has increased by 1.4 years from the 80s to the 90s for those with college education compared with 0.7 years for those with high school or less.
We use self-reported health as our measure of health. Although it is an imperfect measure of health, there are no other categorical measures that provide an all encompassing measure of health over the life-cycle. We reverse the original scale so that 1 equals poor health and 5 excellent health. Furthermore, we add the category 0 which is defined as death. Death is inferred from exit interviews in the HRS and PSID as well as a matching procedure to the National Death Records for HRS respondents.

3.1. Auxiliary Processes

3.1.1. Earnings, AIME and Other Income

We estimate the earnings profile from the PSID using waves from 1980 to 2005. We use the sample of respondents aged 21 to 70 with positive earnings. After deletion of extreme cases, this leaves records on 3714 respondents with on average 6.4 observations per respondent.\(^1\) We use log earnings as the dependent variable and include a quadratic in age, indicators for health states (2 to 5) and fixed effects as explanatory variables. After estimation, we use the average of the fixed effects for the 1940 cohort to trace out the earnings profile. Estimates are reported below along with standard errors (in parentheses).

\[
\log(y_t^{e}) = 0.809 + 0.099 t - 0.001 t^2 + 0.098 h_{1,2}^{t} + 0.14 h_{3,2} + 0.166 h_{4,2} + 0.158 h_{5,2} \\
(0.029) (0.009) (0.001) (0.035) (0.036) (0.037) (0.037)
\]

(13)

Our estimates show a substantial penalty from being in poor health \((h_i = 1)\) relative to other health states. Relative to being in poor health, being in fair health is associated with a 9.8% increase in earnings, being in good health 14% and being in

\(^{1}\) We drop observations with earnings larger than $200,000 per year and less than $5,000.
excellent health 15.8%. The estimated earnings profile peaks at age 46 with average earnings of just over $32,000.

The average indexed monthly earnings are the average of the highest 35 years of earnings where each year of earnings is indexed to age 60 wage levels using the National Wage Index. We use earnings histories to estimate that process. The process for AIME is estimated from Health and Retirement Study Data merged with Social Security earnings histories. We use data from those born between 1936 and 1940. We postulate the following age-specific log-log regression:

$$
\log ae_t = \xi_0 + \xi_1 \log y_t^e + \xi_2 \log y_{t-1} + \xi_3 \log ae_{t-1} + \nu_t
$$

where $\nu_t$ is a prediction error. We construct the AIME on the left-hand side using previous year’s computed AIME from earnings histories (35 highest years) and regress it on current earnings, and whether earnings are greater than zero. The R-squared from such regressions comes close to 0.99 for most ages. The coefficients for each age are displayed in the Technical Appendix.

We estimate the process for other income using the 1980 to 2005 waves of the PSID. Other income is restricted to spousal income. We estimate the following median regression

$$
Q_{0.5}(y_t^e | t, y_t^e) = -7.97 + 0.345t - 0.003t^2 + 0.076 y_t^e + 5.55 I(y_t^e > 0)
$$

which yields an age profile peaking at age 44 with a value of $11,510.

### 3.1.2. Employment and Insurance Transition Probabilities

We use the 1999 to 2005 waves of the PSID to estimate employment and insurance probabilities. We define non-employment as being unemployed and drop observations where respondents do not participate in the labor force. We define someone as insured
(prior to Medicare eligibility) if he is either covered by employer provided health insurance (own or through spouse) or private insurance. Figure 1a reports the fractions of employment and insurance status at each age. On average 17.9% of those aged 25 to 64 do not have health insurance. The fraction of uninsured workers steadily decreases with age until age 50 and then increases slightly. A small fraction, 3.2%, is unemployed. The fraction of unemployed decreases slightly with age. We estimate two-year transition probabilities as a function of age and the origin state using a multinomial logit model. These are estimated using the sample of respondents age 25-64. We compute annual transition probabilities by taking the (matrix) square root of the biennial transition matrix at each age.

3.1.3. Health Process

To estimate the health process, we ideally need panel data on both health and individual level medical expenditures. MEPS is a short panel (on average 2 years) that does not allow to estimate annual transition probabilities across health states reliably. On the other hand, the PSID does not have individual level data on medical expenditures (only at the household level). Finally, the PSID does not capture mortality well, particularly in old age. The HRS provides better coverage in old age but lacks data on the population younger than 50 and has very noisy medical expenditure data compared to the MEPS. Our strategy is to combine all three data sources for the purpose of calibrating the health process. The appendix gives details on how we calibrated the process by estimating the dynamic ordered probit process using maximum simulated likelihood.

3.2. Institutional Details
There are three important institutions represented in the model. First, there is the income
tax schedule. Second, the Social Security benefit calculation formula and finally the
Medicare and Non-Medicare health insurance out-of-pocket schedule.

We use the 2001 Tax function taking into account Federal taxes and Social Security &
Medicare taxes. We take the Social Security rules in effect for individuals born
between 1936 and 1940. For those younger than 65 and insured, we use data reported by
Blau and Gilleskie (2008) from the Health Insurance and Pension Provider Survey
(HIPPS). The median deductible is found to be $200 and the median co-insurance rate is
20%. The median premium is $480. We do not impose a stop-loss on the insurance
contract (maximum out-of-pocket). Since there is no direct cost to investing in health,
individuals in the model who desire to spend slightly more than the stop-loss would then
consume an infinite amount of care.

Establishing an overall co-insurance rate and deductible for Medicare is more
complicated. Only Medicare Part B, which covers outpatient treatments, has a common
deductible-coinsurance structure (part D was not in place for those respondents at the
time). There is no premium for part A and the cost-sharing schedule depends on hospital
stays, which we do not model. The premium for Medicare part B is $490 per year, the
deductible $200 and the co-insurance rate 0.2 (20%). Hence, we choose to simply use the

---

12 We allow for basic deduction in couples and the personal exemption. Prior to age 65, the deduction is
$7,600 and after 65 it is $9,650. We account for the partial tax treatment of Social Security benefits by taxing
only 50% of Social Security benefits.

13 Their Normal Retirement Age (NRA) is taken to be 65 years of age (in reality it was 65 for those born in
1936 and 65 and 6 months for those born in 1940). To calculate the monthly benefit, the AIME is
transformed into the Primary Insurance Amount (PIA). The PIA is a piece-wise linear concave function of
the AIME. The PIA is the benefit that would be paid if claimed at the NRA. If the individual claims prior to
the NRA, the PIA is reduced by 7.6% per year. Similarly, an individual can claim his benefit past the NRA in
which case a delayed retirement credit (DRC) is granted. The credit is 6.5% per year for those born in 1938
(6% for those born in 1936 and 7% for those born in 1940). Upon reaching 70, no DRC is applied.
same price schedule for Medicare and Non-Medicare insurance. Therefore, within the context of our model, there is very little difference in coverage prior to or after Medicare eligibility for those insured. The main difference is the uncertainty in insurance coverage.

3.3. Initial Conditions

We draw initial conditions for assets, AIME, insurance coverage and health status from the empirical distribution in the PSID data at age 25 (we use age 24 to 28 to boost the sample size and replicate observations using frequency weights).

3.4. Moments

To calibrate the model, we use four sets of moments. First, we use median assets by age from the PSID. Second, we use retirement hazard rates by age from the Health and Retirement Study. Third, we use average medical expenditures by age from the MEPS. Fourth, we use mortality rates from ages 50 to 100 computed from the Health and Retirement Study. We discuss the construction of those moments in the technical appendix. Figure 2 reports estimates of those moments by age.

---

14 We applied the insurance parameters to total medical expenditures and compared with actual out-of-pocket expenditures reported in MEPS. At the mean, the computed out-of-pocket expenditures are very similar to those from the data, both before and after the age of Medicare eligibility.

15 The AIME is not directly measured in the PSID. Since AIME is very close to earnings for those respondents (they have little labor market experience) we apply the following procedure. We first compute someone’s percentile in the earnings distribution at age 25. We then attribute the AIME of someone in the same percentile of the earnings distribution from the HRS restricted SSA earnings records. This imputation procedure preserves the correlation between earnings and other outcomes (assets, insurance coverage and health status).
3.5. Calibration

3.5.1 Preferences and Production Function

In principle, we can estimate all structural parameters of the model using a method of simulated moments approach. For this paper, we have instead used a calibration approach, which consists of choosing manually parameters such that the fit of the moments is satisfactory for our purposes. In choosing these parameters, we have borrowed partly from the literature on consumption and labor supply. We should note that no estimates of parameter values exist for a model that simultaneously features endogenous health spending, savings and retirement. Table 1 provides the values of the parameters we have chosen.

Estimates of risk aversion vary substantially across studies. We chose a value of 2.6 for risk aversion which is in the middle of the range of those estimated in the literature. We fix the consumption floor at $3000 following the argument that few of the social programs in place apply to married men (accounting for 85% of the sample) and that take-up of those problems is generally well below 50% (De Nardi, French and Jones, 2006).

The real rate of return is set to 3%. There is considerable heterogeneity in estimates of the discount rate. Some studies find very low values, even negative [-0.04, -0.02] (Hurd, 1989; French, 2005) while others find relatively larger values [0.04 to 0.07] (Gourinchas and Parker, 2002; Cagetti, 2003; De Nardi, French and Jones, 2006). We use a value of

---

16 Studies using linearized Euler equations and consumption data tend to estimate risk aversion parameters in the range 1.5 to 2.5 (Attanasio and Weber, 1995). Estimates based on asset data tend to be much larger. Cagetti (2003) estimates on PSID data risk aversion parameters ranging from 2.7 to 4.3. Recent estimates by French (2005), French and Jones (2007) and De Nardi, French and Jones (2006) suggest a risk aversion coefficient in the range 3-7. Each of these models is different and the estimate of risk aversion can be very sensitive to the amount of risk assumed, the sample used (elderly vs. young), the extent of liquidity constraints and the presence of a consumption floor. One of the reasons for the larger values in these recent studies is the role played by the consumption floor in depressing savings (Hubbard, Skinner and Zeldes, 1995). A larger value of risk aversion is needed to rationalize savings when there is a consumption floor.
0.05. We assume a relatively strong bequest motive in order to rationalize the high level of assets in old age (relative to annuity income). The value of $\theta_b$ is set to 0.05.

Next, we pin down retirement by choosing the marginal utility of leisure. We select values such that they capture the distribution of retirement ages and a median retirement age of 63.5 years.

Finally, we calibrate the health production function and disutility/utility of poor health status. We use the average medical expenditure profile, which rises almost linearly with age. This, along with the life-cycle path of health and mortality informs the productivity of medical care. We find that low values of productivity are necessary to rationalize why individuals spend so little on their health at earlier ages given the potentially important returns in terms of wages and consumption benefits. Without an age adjustment, the medical expenditure profile is much flatter, as we cannot rationalize why agents delay health investments until old age. Given that the health production function is concave, there is an incentive to smooth health investment over the life-cycle. We calibrate the age-related increase in productivity so that we match the medical expenditure profile. The increase in productivity due to poor/fair health is calibrated by making use of the differences in health expenditures by health status. Total medical expenditures are 8 times larger for those in poor health than those in excellent health in the MEPS. Without the health related productivity parameter, we can only generate a factor of 5.

We use a second piece of information to calibrate the health production function, in particular its curvature. We show in the technical appendix, using a simplified version of the Hall and Jones (2007) model, that both price and income elasticities depend on the curvature of the health production function. The intuition is that in a health production
model, the rate at which the marginal product of medical expenditure relative to the marginal utility of consumption increases as income or price increases is directly related to the value of life extension and therefore to the demand for health.

Hence, we calibrate the health production function parameter so that it replicates roughly, among those of working age, the price (co-insurance) and income elasticity found in the literature. The co-insurance elasticity tends to fall between -0.1 and -0.4 depending on the type of services (preventive vs. curative and in-patient vs. out-patient) (Ringel et al., 2000). Estimates from the Health Insurance Experiment (HIE) (Manning et al., 1987) yields an estimate between -0.17 and -0.22 which varies across co-insurance rates and is slightly larger for preventive care (-0.17 to -0.43). For calibration, we consider a change from a co-insurance rate of 0.2 to 0.5 and calculate the resulting change in medical expenditures.

As for the income elasticity, cross-sectional studies tend to obtain estimates in the range of 0.2 to 0.4 while studies using cross-country or time-series variation find an income elasticity close to unity. Cross-sectional studies face the problem of measurement error and reverse causality. One possible reason for the larger elasticity at the cross-national level is that such analyses omit differences in technology, which are correlated with income (higher income countries have better technology). There is yet another important difference between cross-sectional and aggregate cross-national studies resulting from the nature of income changes. If income shocks are transitory, the response is likely to be smaller than if the shock is permanent. A permanent income shock raises the value of living longer if the value of life is a normal good. Viscusi and Aldy (2003) find that the value of life increases with income when performing a meta-analysis of value of life
estimates. Hence, the permanent responses should be larger than transitory responses as the latter will not lead to a rise in the value of life.\textsuperscript{17} For these reasons, we expect in the model a permanent income elasticity of roughly unity while the transitory response should be in line with cross-sectional studies. Using this information, we find that the curvature of the health production function rationalizing these price and income response is close to \( \theta_1 = 2.1 \). The resulting elasticities are presented for various age groups in Table 2.

### 3.5.2 Fit of Moments

In Figure 3, we compare moments of the data with those calculated on simulated data. Given the admittedly coarse calibration, the fit is relatively good. First, mortality is relatively close to the HRS data. In particular, we simulate a remaining life-expectancy of 27.1 years at age 50 while it is 26.6 in the HRS for this group. The progression of mortality is very well approximated before age 75 while we under-estimate mortality in the ages 75 to 90 years old. Health transitions are in general well replicated which is very important for this model (not shown here). For the model to match health transitions, it must also match medical expenditures. The simulated average medical expenditure profile tracks the data relatively well. We over-estimate expenditures in middle ages and under-estimate at older ages. The asset profile follows the hump-shape found in the data and the slow decline after retirement. However, we tend to over-estimate assets just before retirement. Finally, the simulated retirement pattern captures well the spike at age 62 and the increasing hazard beyond age 65. However, it misses the peak at age 65. One reason is that we have assumed

\textsuperscript{17} Dustmann and Windmeijer (2000) make a similar point but looking at time allocation jointly with medical investment. In that case the transitory wage response can be negative as workers spend less time investing in health when their wage rises temporarily while the positive income effect dominates for permanent wage changes. They find support for such effects in German panel data.
everyone with insurance at the time of retirement had access to retiree health insurance. Hence, those retiring at 62 and 63 do not have an incentive to delay until age 65.

4. Simulations

We use the model to investigate the relative contribution of various factors to the growth in health spending and life expectancy during the last 40 years. According to a recent report by the CBO (2008), per capita medical expenditures are roughly 500% larger today than they were in 1965 which represents an average growth rate of 3.6%.

Figure 4 shows, using data from Cutler and Meara (1998), average medical expenditures by age in 2004 relative to 1963 (adjusted to 2004 dollars using the CPI). The largest change is found for the elderly (65+). Over this period, there has been an increase in life-expectancy at age 50 of 6 years for men (from 22.9 to 28.9 years (6 years)). We investigate whether our model can replicate these changes as a result of growth in income, generosity of insurance and technological change.

Newhouse (1992) reviews the reasons behind the increase in health spending. The generosity of health insurance has increased over the last 40 years and the price elasticity of the demand for health is negative. The average co-insurance rate has fallen from an average of 0.6 in the 1960s to 0.2 in 2005 (Newhouse, 1992). On the income side, real earnings have increased by 27.3% in real terms. Social Security has become more generous, which has also raised the income of retirees by an average of 20% for medium and low earners (Diamond and Gruber, 1997).\(^{18}\) Given the positive income elasticity of

---

\(^{18}\) Other factors that could explain the rise include population structure (aging), physician induced demand and lagging factor productivity in the health sector. Newhouse (1992) reviews the evidence on each of those and finds little ground for the hypothesis that these explain a large share of what actually happened over the period.
demand, this should also explain a part of the rise in health spending as argued by Hall and Jones (2006). Finally, technological change has lead to large improvements in the treatment of several conditions (Deaton, Cutler and Lleras-Muney, 2006) and some have argued (Newhouse, 1992) that the bulk of the increase in life expectancy is due to technological change, which increases productivity. There is little evidence on the magnitude of productivity improvements. Cutler (2004) looks at cardio-vascular diseases, the leading cause of death among men over this period, and concludes that medical advancements explain 2/3 of the improvement in mortality over the last half century, the other third likely explained by the decrease in smoking among men. Lichtenberg (2003) provides evidence that a large part of the gain in life expectancy in recent decades could be attributed to pharmaceutical innovation.

We first set income and insurance parameters to levels corresponding with a 1965 baseline. This means reducing average earnings by 27%, reducing the generosity of social security benefits by 20% and increasing co-insurance rates to 0.6. This should not get us to the 1965 spending level since technological change also plays a role. We model technological change as a scalar that varies the productivity term in equation (8) as a function of calendar time $s$:

$$\theta_0(z, h, t) = z e^{\theta_0 + \theta_1 h_{z,c} + \theta_2 t}, \ z_s = e^{\theta_s s}$$

where $s = 0$ refers to 2005 ($z_0 = 1$). We calibrate the change in productivity so that it matches the level of medical expenditures in 1965. We find that going backward in time an annual productivity decline $\theta_z = -0.018$ roughly fits the data such that $z_{40} = 0.47$. Table 3 reports a number of outcomes in the 1965 and 2005 scenarios. The simulated 1965 scenario yields spending levels close to those that would be predicted if we used the ratio
of 2005 to 1963 spending levels (spending is 6 times higher in 2005) produced by CBO. The predicted average spending would by $697 while we obtain $740. The simulated average annual rate of increase in health spending is 3% in real terms, or 400% total. We simulate the share of net income devoted to health spending to be 15% in 2003 while it is 3% in 1965. Given minor differences in the definition of income (we use net income rather than gross income or GNP), this rate of growth is consistent with the data compiled by the CBO (4.5% in 1965 and 14.5% in 2005). Another important finding is that the age pattern of medical expenditures was much flatter in 1965 than in 2005, except at very young ages. Cutler and Meara (1998) report data on the age distributions of total medical expenditures from National Health Surveys between 1963 and 1987. Although we replicate the more rapid growth in old age (relative to the age 35-45), our simulation also generates high growth for those aged 25-34. Cutler and Meara (1998) report very little growth in that age group. We take this as suggestive evidence that technological change has been concentrated in the older population or on curative rather than preventive care. Nevertheless, our calibration of technological change is probably a good first-order approximation to the scale of productivity improvements over this period.

We simulate that life expectancy at age 50 has grown by 16% between 1965 and 2005. The relative change from 2005 to 1965 was 25.8% (for all men). Hence, this explains roughly 60% of the improvement in life expectancy between 1965 and 2005. Note that we do not model other changes in health, such as the reduction in smoking over this period. Hence, it is not surprising that we cannot explain the entire gap.

We also look at the average retirement age. Interestingly, the average retirement age does not change much. However, as we will see next, this is likely the result of the fact
that the effects of various changes (e.g. income, technology, health insurance copayment) on retirement cancel out.

Next we look at the relative contribution of each of these changes (insurance, income and technology) in explaining the rise in spending and life expectancy. We sequentially introduce changes starting from 1965. Results are presented in Table 4.

As can be seen from the second column of Table 4, increasing the generosity of health insurance increases medical expenditures substantially. The average rises from $742 to $1725 which represents a change of 132%. The effect is larger among younger individuals who we find are more price sensitive (Table 2). Increasing the generosity of health insurance raises the income share of medical expenditures from 3.1% to 7.2%. There is no change in the average age of retirement. An important finding is that increasing the generosity of health insurance has a very small effect on life expectancy; less than 0.3 years. This implies that the additional production of health has little health benefit and would not be consumed if individuals had to pay for it. This is consistent with the evidence from the HIE that generous insurance had little effect on health outcomes within the time frame of the study.

Increasing real earnings and the generosity of social security has surprisingly little effect on health spending. This is shown in column 3 of Table 4. Total expenditures increase by 44% from $742 to $1071. The average change in net income is 16.6%. Part of the reason why expenditures do not rise sharply is due to the responsiveness of the retirement age to the income change. We find that it reduces the retirement age by 2.4 years. This effect, along with the progressivity of income taxes, mitigates the income
effect. Furthermore, the ratio of medical expenditures to income rises only by 1 percentage point. Finally, there is a modest gain in life expectancy of 0.19 years.

In column 4, we increase the productivity of medical technology by 1.8% a year until 2005. This implies an improvement of 150% in productivity over the period. This increases total medical expenditures by about as much as caused by the change in insurance generosity. The magnitude of the change is similar across age groups, which is most likely the result of our assumption that productivity gains were distributed equally over the life-cycle and across health status. The average retirement age increases by 1 year, in most part due to the fact that individuals can now expect to live 3 more years due to technological improvements. Of the three changes (insurance, income and technology), technological change is the only one that is capable of generating appreciable gains in life expectancy. By itself, it explains roughly half of the increase in life expectancy from 1965 to 2005.

None of the scenarios are able to generate health expenditure levels of the magnitude witnessed in 2005, even if we sum their relative contribution. For example, insurance leads to a 132%, income a 44% and technology a 110% change in spending. This sums to 287%, i.e. well under the 464% that we reach when we implement all changes together. Hence, we must conclude that there are important interaction effects. This exercise would imply these interactions are responsible for a rise of 176% in health care spending over the last 40 years. Some of these interactions are intuitive. Improvements in technology will have a larger effect on medical consumption if the price of consumption is lower (insurance improving). Similarly, more income will lead to larger medical consumption if such consumption is more productive. What the model reveals is that an
increase in income at 1965 technology level does not have as large an effect on medical consumption as if that same change occurs with the 2005 technology. By occurring simultaneously, these changes create conditions for a “perfect storm” within the context of model when health investment is endogeneous to the state of technology, income and insurance.

In Figure 4, we show the sources of increase in health spending. Increased generosity of health insurance accounts on its own for 29% of the insurance in expenditures while income only for 10%. Technological change accounts for 24% of the change. But the largest component is the interaction effect, which amounts to 37% of the total change.

As for life expectancy, we report in Figure 5 the sources of the increase as we estimate them from the model. Technological change, taking the form of productivity gains accounts for the bulk of the increase, 59%. Income and insurance explain a very small share of the total increase (4% and 6% respectively). Interaction effects are smaller for life expectancy than they are for health spending, amounting to 6%. This leaves an unexplained residual category of 25% of the total increase in life expectancy. There are several plausible candidates to explain the remaining part; one is the reduction in smoking, the second is general improvements in living conditions not captured by the model (e.g. air quality, working conditions, etc).

5. Conclusion

We calibrate a stochastic life-cycle model of health spending, asset accumulation and retirement which allows us to look at the determinants of the rise in health care spending and longevity over the last 50 years. We find that with reasonable parameter values, the
model fits the data relatively well and is able in particular to replicate the income and price elasticity of the demand for medical care. We then perform a simulation where we attempt to go back to 1965 by simultaneously decreasing the generosity of health insurance and income to 1965 levels. We back out the rate of productivity gains necessary to achieve 1965 spending levels and longevity and find that a rate of 1.8% fits the data well.

We find that both insurance and technological change independently explain respectively 29% and 24% of the total change in health spending and that income explains roughly 10%. The remainder, 37%, is due to what we call synergy effects or interactions created by simultaneously increasing income, the generosity of health insurance and productivity. As for life expectancy, we find that the bulk of the gain is due to productivity growth while income growth and increased generosity of health insurance explain a relatively small fraction of the overall increase (less than 10%).

The relative importance of income in our study differs from Hall and Jones (2007) for two important reasons. First, Hall and Jones (2007) use real GNP per capita growth while we use real earnings growth, the former being much larger over this period. For workers in our sample, most of the growth in real GNP was not reflected in earnings growth. The average worker saw earnings growth of roughly 25% over the period we cover compared to growth in GNP per capita nearing 200% between 1965 and 2005. Second, we obtain a lower income elasticity of medical care. Hall and Jones require an income elasticity of medical care close to 4 in order to explain, solely using income growth, that medical expenditures rose by 800% while GNP per capita rose by roughly 200% between 1950 and 2000. Our estimate of the income elasticity (to a permanent change in income) is close to unity which is in line with the cross-national estimates from the literature and larger than
those from the micro-literature (between 0.2 and 0.4) (see Newhouse, 1992). This result is also consistent with recent evidence on the effect of income on health spending. For example, Acemoglu, Finkelstein and Notowidigdo (2009) find an income elasticity of health spending ranging from 0.7 to 1 using oil price shocks as an instrument for income at the local area level. Consistent with our findings, they argue that income is unlikely to be an important driver of the *income share* of health spending. Our findings rather suggest that insurance and technology are more likely to be important drivers.
References


Technical Appendix

Solution Method

Because some of the state variables are continuous (assets and AIME), we define a bi-dimensional equally-spaced grid over that space. We select 35 points for assets and 35 points for AIME. The maximum for assets is set to 300 and the minimum 0. For AIME, the minimum is zero and the maximum 6. Since decision rules are likely more non-linear at low values of assets and AIME, we set the grid in terms of equally spaced points using the square root of assets and AIME ($a_i^{1/2}$ and $ae_i^{1/2}$). We solve for value function at each point on that grid starting at terminal age $T$.

We have two continuous decision variables. We define a grid for each of these decisions (45 points for consumption and 45 for medical expenditures). We select bounds of 0 and 25 for medical expenditures (25 is the 99th percentile in the data) and define the grid for consumption in terms next period's assets using the boundary conditions along with cash on hand. For consumption, we center the grid around current cash on hand and define an equally spaced grid away from that point. We use interpolation to calculate next period's value function when the solution for consumption and medical expenditure does not fall on the grid given the current state and retirement choice. Since we use an equally-spaced grid, we use the bi-dimensional cubic spline approximation proposed by Habermann and Kindermann (2007). When retirement is an option, we compute the optimal solution for each retirement path and compare value functions to calculate $r^*(s_i)$.

The decision rules are generally insensitive to the number of grid points. We solve for decision rules using a program written in Ox from OxMetrics (with sub-programs written in C++) and the parallel message parsing interface MPICH on a 16 processor core cluster with 8 GB of RAM. A typical solution takes roughly 12 minutes.
AIME Approximation

The following table provides age-specific coefficients for the AIME regression. The average R-squared is 0.991. For ages above 60, the age 60 coefficients are used.

Table A.1 AIME Approximation Coefficients

<table>
<thead>
<tr>
<th>age</th>
<th>earnings&gt;0</th>
<th>log(earnings)</th>
<th>log(AIME(t-1))</th>
<th>Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>0.0848</td>
<td>0.1380</td>
<td>0.7425</td>
<td>-0.6538</td>
</tr>
<tr>
<td>27</td>
<td>0.0706</td>
<td>0.0793</td>
<td>0.8654</td>
<td>-0.2817</td>
</tr>
<tr>
<td>28</td>
<td>0.0713</td>
<td>0.0794</td>
<td>0.8360</td>
<td>-0.3334</td>
</tr>
<tr>
<td>29</td>
<td>0.0221</td>
<td>0.0768</td>
<td>0.8792</td>
<td>-0.2243</td>
</tr>
<tr>
<td>30</td>
<td>0.0511</td>
<td>0.0578</td>
<td>0.8719</td>
<td>-0.2093</td>
</tr>
<tr>
<td>31</td>
<td>0.0072</td>
<td>0.0541</td>
<td>0.9255</td>
<td>-0.1109</td>
</tr>
<tr>
<td>32</td>
<td>0.0075</td>
<td>0.0665</td>
<td>0.8884</td>
<td>-0.1772</td>
</tr>
<tr>
<td>33</td>
<td>0.0605</td>
<td>0.0447</td>
<td>0.9012</td>
<td>-0.1542</td>
</tr>
<tr>
<td>34</td>
<td>0.0104</td>
<td>0.0584</td>
<td>0.8906</td>
<td>-0.1521</td>
</tr>
<tr>
<td>35</td>
<td>0.0014</td>
<td>0.0485</td>
<td>0.9140</td>
<td>-0.1015</td>
</tr>
<tr>
<td>36</td>
<td>0.0588</td>
<td>0.0208</td>
<td>0.9410</td>
<td>-0.0589</td>
</tr>
<tr>
<td>37</td>
<td>0.0513</td>
<td>0.0260</td>
<td>0.9312</td>
<td>-0.0719</td>
</tr>
<tr>
<td>38</td>
<td>0.0269</td>
<td>0.0301</td>
<td>0.9430</td>
<td>-0.0575</td>
</tr>
<tr>
<td>39</td>
<td>0.0269</td>
<td>0.0208</td>
<td>0.9636</td>
<td>-0.0265</td>
</tr>
<tr>
<td>40</td>
<td>0.0166</td>
<td>0.0234</td>
<td>0.9645</td>
<td>-0.0259</td>
</tr>
<tr>
<td>41</td>
<td>0.0033</td>
<td>0.0228</td>
<td>0.9777</td>
<td>-0.0128</td>
</tr>
<tr>
<td>42</td>
<td>0.0259</td>
<td>0.0225</td>
<td>0.9610</td>
<td>-0.0323</td>
</tr>
<tr>
<td>43</td>
<td>0.0158</td>
<td>0.0181</td>
<td>0.9818</td>
<td>-0.0117</td>
</tr>
<tr>
<td>44</td>
<td>0.0048</td>
<td>0.0175</td>
<td>0.9871</td>
<td>-0.0013</td>
</tr>
<tr>
<td>45</td>
<td>0.0058</td>
<td>0.0146</td>
<td>0.9916</td>
<td>0.0048</td>
</tr>
<tr>
<td>46</td>
<td>0.0060</td>
<td>0.0165</td>
<td>0.9874</td>
<td>-0.0036</td>
</tr>
<tr>
<td>47</td>
<td>0.0033</td>
<td>0.0185</td>
<td>0.9807</td>
<td>-0.0046</td>
</tr>
<tr>
<td>48</td>
<td>0.0047</td>
<td>0.0170</td>
<td>0.9870</td>
<td>-0.0051</td>
</tr>
<tr>
<td>49</td>
<td>0.0059</td>
<td>0.0133</td>
<td>0.9912</td>
<td>0.0018</td>
</tr>
<tr>
<td>50</td>
<td>0.0056</td>
<td>0.0129</td>
<td>0.9931</td>
<td>0.0008</td>
</tr>
<tr>
<td>51</td>
<td>0.0040</td>
<td>0.0119</td>
<td>0.9956</td>
<td>0.0019</td>
</tr>
<tr>
<td>52</td>
<td>0.0000</td>
<td>0.0120</td>
<td>0.9964</td>
<td>0.0031</td>
</tr>
<tr>
<td>53</td>
<td>0.0163</td>
<td>0.0121</td>
<td>0.9757</td>
<td>-0.0022</td>
</tr>
<tr>
<td>54</td>
<td>0.0135</td>
<td>0.0189</td>
<td>0.9567</td>
<td>-0.0080</td>
</tr>
<tr>
<td>55</td>
<td>0.0060</td>
<td>0.0123</td>
<td>0.9840</td>
<td>-0.0016</td>
</tr>
<tr>
<td>56</td>
<td>0.0056</td>
<td>0.0104</td>
<td>0.9880</td>
<td>0.0004</td>
</tr>
<tr>
<td>57</td>
<td>0.0158</td>
<td>0.0076</td>
<td>0.9853</td>
<td>-0.0008</td>
</tr>
<tr>
<td>58</td>
<td>0.0134</td>
<td>0.0068</td>
<td>0.9875</td>
<td>0.0006</td>
</tr>
<tr>
<td>59</td>
<td>-0.0019</td>
<td>0.0144</td>
<td>0.9674</td>
<td>0.0065</td>
</tr>
<tr>
<td>60</td>
<td>-0.0056</td>
<td>0.0176</td>
<td>0.9578</td>
<td>0.0105</td>
</tr>
</tbody>
</table>
Calibration of Health Process

Self-reported health is defined in the same way in all three surveys. Hence, we resort to the following strategy:

- Estimate the distribution of medical expenditures as a function of self-reported health status and age in the MEPS. We assume the distribution to be a mixture of two distributions; the first is a discrete distribution representing the probability of positive medical expenditures, whereas the second distribution is lognormal conditional on expenditures being positive. We keep both the fraction of positives, mean and variance of the log normal distribution conditional on having positive expenditures for each age and health state.

- Pool PSID data for those age <50 (waves 1999 to 2005) and HRS data for those age >50 (waves 1992 to 2006). The distribution of self-reported health is remarkably similar at the cutoff ages of 50 as can be seen from Figure 1b.

- Estimate parameters of the health process (equation 10) other than those of the investment function ($\gamma$) by maximizing the following likelihood:

$$L(\gamma \mid \theta, h_{it}, t) = \prod_{i,t \in PSID} \int p_{h,PSID}(t, h_{it}, m_{it}; \theta) f_{MEPS}(m_{it} \mid h_{it}, t) dm_{it}$$

$$\times \prod_{i,t \in HRS} \int p_{h,HRS}(t, h_{it}, m_{it}; \theta) f_{MEPS}(m_{it} \mid h_{it}, t) dm_{it}$$

where the probabilities $p_{h,x}(t, h_{it}, m_{it}; \theta)$ are ordered probit probabilities for observing in dataset $x$ an individual at age $t + 1$ in state $h_{i,t+1} = h$ given age, current health status, medical expenditures and conditional on the productivity parameters $\theta$. The integrals are evaluated using simulation by drawing from the distribution of medical expenditures in the MEPS. We use 25 draws per respondent/wave. The resulting maximum simulated likelihood estimator $\hat{\gamma}(\theta)$ provides us with a calibration that fits the health process given a choice of productivity parameters. We adopt this conditional procedure because we calibrate $\theta$ using other moments of the data. 19 Based on diagnostic tests, it became apparent that the age-invariant assumption on the thresholds of the dynamic ordered model was too restrictive, particularly for mortality ($h_{it} = 0$). Hence, we allowed the threshold defining the upper bound for mortality to depend on a fourth order polynomial in age.

19 Trivially, at zero productivity, there are no simulated medical expenditures. Increasing the productivity will lift the simulated medical expenditure profile. But if the simulated profile is below the true profile, mortality will be over-estimated and vice versa if productivity is “too high”. Because the depreciation adjustment is with respect to the actual medical expenditure profile, it is not until the actual and simulated profile are close that mortality will be matched.
Moments used in Calibration

**Assets**

We estimate a median asset age-profile using the PSID following the methodology of French (2005). Assets include all real and financial assets minus debt. We express all monetary amounts in $2004 and discard observations with assets larger than one million dollars. We estimate a median regression model of the form

\[ Q_{0.5}(a_t | t, \text{cohort}, \text{hhsize}) = \alpha_a + \alpha_t + \lambda_{\text{cohort}} + \beta \text{hhsize} \]

where \( \alpha_t \) are age fixed effects. We also control for 5-year cohort fixed effects (\( \lambda_{\text{cohort}} \)) and household size (\( \text{hhsize}_a \)). Upon estimation we predict assets for household size of 3 born between 1936 and 1940.

**Retirement**

Retirement is a complex state to measure. First, some workers return to work after prolonged periods of inactivity in old age while not considering themselves retired. Second, some individuals work while claiming Social Security benefits. Since our model considers a simplified definition of retirement, it is not clear which concept should be used. We elected to use self-reports of respondents on whether they are retired in the HRS. We exclude as retired those who report being retired but are still at work. We use the retirement ages from the cohort born between 1931 and 1941 which has almost entirely retired by 2004. We find peaks in the retirement hazard at age 62 and 65 which are respectively the ages of early and normal retirement age under Social Security.

**Medical Expenditures**

We use median total medical expenditures from the 2000 to 2003 waves of the MEPS. These expenditures include in-patient and out-patient expenditures as well as drug expenditures. They do not however include nursing home expenditures. To be consistent with other monetary flows in the model, we express those in 2004 dollars using the CPI. Not surprisingly, medical expenditures increase steadily with age.

**Mortality**

We use the mortality profile from the Health and Retirement Study (older than 50) rather than from official life tables because we are looking at a specific sub-population (non-hispanic white men with high school education or less). We transform biannual rates into annual ones assuming a constant hazard within each age interval. Even with large samples, the mortality rates at older ages are quite volatile as can be seen from Figure 2. Smoothing these mortality rates (using a quadratic in log mortality rates) and computing a life-table yields remaining life expectancy of 47.3 years at age 25 and 26.9 years at age 50. We also report the mortality rates for men in 2002. White males without college education have slightly lower life expectancy than average male life expectancy. Life expectancy at age 25 according to the life table is 49.2 years and 27.6 at age 50.
Price and Income Elasticities

We demonstrate how the price and income elasticity of demand are related to the concavity of the health production function. We add a co-insurance rate (price) to the model used by Hall and Jones (2006). Let $1/h$ be the mortality rate and $h$ life expectancy (health status). Income $y$ can be allocated between consumption $c$ and medical expenditures $m$. Out-of-pocket medical expenditures are given by $\mu m$ where $\mu$ is the co-insurance rate. Health is produced through a production function $i(m)$ which is increasing and concave. The problem is given by

$$\max U(c,h) = \int_0^\infty e^{\lambda t} u(c) dt = hu(c)$$

s.t.

$$y = c + \mu m$$

$$h = i(m)$$

As in Hall and Jones, define $\eta_m = \frac{v(m)}{i(m)} m, \eta_u = \frac{u(c)}{u(c)} c$ and $\lambda = \frac{\eta_m}{\eta_m + \eta_u}$. The optimality condition is

$$s = \frac{m}{y} = \frac{1}{\mu} \lambda$$

The income elasticity of optimal medical expenditures $m^*$ is given by

$$\eta_y = 1 + \frac{y^2}{\mu m} \lambda_y$$

where $\lambda_y$ is the derivative of $\lambda$ with respect to $y$. The uncompensated price elasticity is

$$\eta_\mu = -1 + \frac{y}{m} \lambda_\mu.$$

Both $\lambda_\mu$ and $\lambda_y$ depend on the curvature of the utility and production functions since they involved second derivatives of $i(m)$ and $u(c)$. If both utility and production functions are of the constant elasticity type, the income and price elasticities are 1 and -1. Alternatively as in Hall and Jones, if $\lambda_y > 0$, for example if the production function exhibits constant elasticity while the utility function exhibits decreasing elasticity. In the case we consider both utility and production exhibit decreasing elasticity such that the curvature of the production function, taking the curvature of the utility function as given, is identified by calibrating the price and income elasticities to existing evidence on their magnitude (for example, the HIE).
Figures

Figure 1a Employment and Insurance Status

Notes: raw data from 1999 to 2005 waves of PSID.

Figure 1b Health Status

Figure 2 Moments for Calibration

Notes: The figures present the moments from the data used to calibrate parameters. See the discussion in section 4.4.
Figure 3 Goodness-Fit of Model Simulations

Notes: Simulated and data profiles based on parameters presented in Table 1.
Figure 4 Increase in Health Spending

Share of Increase in Health Spending 1965-2005

Notes: Based on simulation output from Table 4. Insurance refers to the induced change brought by an increase in the generosity of health insurance keeping income and productivity at the 1965 level. Income refers to the induced change brought by the increase in real earnings and the generosity of Social Security keeping insurance and productivity at the 1965 level. Productivity refers to the induced change brought by improvements in productivity keeping income and insurance at the 1965 level. Finally, interaction refers to the residual induced change brought by the interaction of income, insurance and productivity change when occurring simultaneously.
Figure 5 Increase in Life Expectancy

Share of Increase in Life Expectancy at age 50, 1965-2005

Notes: Based on simulation output from Table 4. Insurance refers to the induced change brought up by an increase in the generosity of health insurance keeping income and productivity at the 1965 level. Income refers to the induced change brought by the increase in real earnings and the generosity of Social Security keeping insurance and productivity at the 1965 level. Productivity refers to the induced change brought by improvements in productivity keeping income and insurance at the 1965 level. Finally, interaction refers to the residual induced change brought by the interaction of income, insurance and productivity change when occurring simultaneously. Since the simulated change in life expectancy is less than the observed relative share. There is another residual category which captures the unexplained component of the increase in life expectancy in the data.
## Tables

### Table 1 Structural Parameters Chosen by Calibration

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>Coefficient of relative risk aversion</td>
<td>2.6</td>
</tr>
<tr>
<td>$\psi_n$</td>
<td>Utility benefit if health status &gt;2</td>
<td>0.035</td>
</tr>
<tr>
<td>$\psi_0$</td>
<td>Baseline log utility benefit of retirement</td>
<td>4.65e-03</td>
</tr>
<tr>
<td>$\psi_a$</td>
<td>Incremental log utility benefit of retirement by age</td>
<td>6.15e-04</td>
</tr>
<tr>
<td>$\theta_b$</td>
<td>Strength of bequest motive</td>
<td>0.05</td>
</tr>
<tr>
<td>$\theta_{00}$</td>
<td>Baseline log productivity (age 25, good health)</td>
<td>-3.352</td>
</tr>
<tr>
<td>$\theta_{0h}$</td>
<td>% increment in productivity fair/poor health</td>
<td>1</td>
</tr>
<tr>
<td>$\theta_{0t}$</td>
<td>% change in productivity with age</td>
<td>0.02</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>Concavity of production function</td>
<td>2.1</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Real interest rate</td>
<td>0.03</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Discount factor</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Notes: See discussion in section 4.5 for the justification.
Table 2 Price and Income Effects on Medical Expenditures in 2005 Baseline Scenario

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Co-Pay (0.2 to 0.5)</th>
<th>Permanent Income (+10%)</th>
<th>Transitory Income (+10%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-35</td>
<td>-0.345</td>
<td>1.564</td>
<td>0.631</td>
</tr>
<tr>
<td>35-45</td>
<td>-0.281</td>
<td>0.888</td>
<td>0.576</td>
</tr>
<tr>
<td>45-55</td>
<td>-0.267</td>
<td>0.909</td>
<td>0.583</td>
</tr>
</tbody>
</table>

Notes: Simulation based on calibration in Table 1. The first two columns consider permanent changes. In the first column, we increase the co-payment from 0.2 to 0.5 across all ages for those insured and look at the resulting change in total medical expenditures. In the second column, we increase earnings by 10% over the entire working life. The reported elasticity is the average elasticity within the age group. In the last column, we increase earnings at the mid point in the age interval (30 for the first, 40 for the second and 50 for the last age group).
<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Simulation results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1965</td>
</tr>
<tr>
<td>Total Medical Expenditures ($k)</td>
<td></td>
</tr>
<tr>
<td>25-35</td>
<td>0.131</td>
</tr>
<tr>
<td>35-45</td>
<td>0.563</td>
</tr>
<tr>
<td>45-55</td>
<td>1.020</td>
</tr>
<tr>
<td>55-65</td>
<td>1.172</td>
</tr>
<tr>
<td>65-75</td>
<td>1.039</td>
</tr>
<tr>
<td>75-85</td>
<td>0.740</td>
</tr>
<tr>
<td>85+</td>
<td>0.945</td>
</tr>
<tr>
<td>Total ($k)</td>
<td>0.742</td>
</tr>
<tr>
<td>net income ($k)</td>
<td>24.030</td>
</tr>
<tr>
<td>share of net income</td>
<td>0.031</td>
</tr>
<tr>
<td>Average retirement age (years)</td>
<td>64.2</td>
</tr>
<tr>
<td>Life expectancy at age 50 (years)</td>
<td>23.88</td>
</tr>
</tbody>
</table>

**Notes:** simulated results in 2005 and 1965 using the insurance, income and technology changes. Amounts in thousands of $US 2004.
### Table 4 Simulated Incremental Changes Starting from 1965 Scenario

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>1965</th>
<th>increase generosity insurance</th>
<th>increase real earnings and social security generosity</th>
<th>increase productivity</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. Medical Expenditures($k)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 25-35</td>
<td>0.131</td>
<td>0.397</td>
<td>0.235</td>
<td>0.398</td>
<td>1.376</td>
</tr>
<tr>
<td>35-45</td>
<td>0.563</td>
<td>1.355</td>
<td>0.816</td>
<td>1.201</td>
<td>3.210</td>
</tr>
<tr>
<td>45-55</td>
<td>1.020</td>
<td>2.265</td>
<td>1.418</td>
<td>1.941</td>
<td>4.535</td>
</tr>
<tr>
<td>55-65</td>
<td>1.172</td>
<td>2.538</td>
<td>1.614</td>
<td>2.233</td>
<td>5.363</td>
</tr>
<tr>
<td>65-75</td>
<td>1.039</td>
<td>2.494</td>
<td>1.530</td>
<td>2.198</td>
<td>6.149</td>
</tr>
<tr>
<td>75-85</td>
<td>0.740</td>
<td>1.925</td>
<td>1.197</td>
<td>1.887</td>
<td>6.052</td>
</tr>
<tr>
<td>85+</td>
<td>0.945</td>
<td>2.378</td>
<td>1.620</td>
<td>2.302</td>
<td>7.585</td>
</tr>
<tr>
<td>Avg Medical Expenditures($k)</td>
<td>0.742</td>
<td>1.725</td>
<td>1.071</td>
<td>1.563</td>
<td>4.183</td>
</tr>
<tr>
<td>Net income ($k)</td>
<td>24.030</td>
<td>23.991</td>
<td>27.985</td>
<td>23.760</td>
<td>28.221</td>
</tr>
<tr>
<td>Share of net income</td>
<td>0.031</td>
<td>0.072</td>
<td>0.038</td>
<td>0.066</td>
<td>0.148</td>
</tr>
<tr>
<td>Average retirement age (years)</td>
<td>64.2</td>
<td>64.3</td>
<td>61.9</td>
<td>65.1</td>
<td>63.9</td>
</tr>
<tr>
<td>Life expectancy at age 50 (years)</td>
<td>23.88</td>
<td>24.16</td>
<td>24.07</td>
<td>26.86</td>
<td>27.65</td>
</tr>
</tbody>
</table>

**Notes:** Simulated outcomes. First column represents scenario in 1965. Column 2 to 4 introduce changes one at a time starting from 1965. Finally, the last column presents estimates in 2005 (when all changes are introduced at once). All amounts in thousands $US 2004.