Health Inequality over the Life-Cycle

Timothy J. Halliday*

University of Hawai'i at Mānoa and IZA

First Draft: August 10, 2009

October 2, 2009

Abstract

We investigate the evolution of health inequality over the life-course. Health is modeled as a latent variable that is determined by three factors: endowments, and permanent and transitory shocks. We employ Simulated Minimum Distance and the Panel Study of Income Dynamics to estimate the model. We estimate that permanent shocks account for under 10% of the total variation in health for the colleged educated, but between 35% and 70% of total health variability for people without college degrees. Consistent with this, we find

*Address: 2424 Maile Way; Saunders Hall 533; Honolulu, HI 96822. E-Mail: halliday@hawaii.edu. Tele: (808) 956-8615. I would like to thank the Maui High Performance Computing Center for computational time.

that health inequality moves substantially more slowly over the life-course for the college educated.

JEL Code: I1, C5

Key Words: health, dynamic panel data models, variance decomposition

1 Introduction

While the economics literature has devoted much effort to the modeling of income dynamics, the modeling of health dynamics has received relatively less attention. This is true despite health status becoming more prevalent in life-cycle models of consumer behavior (*e.g.* Yogo 2008; Palumbo 1999; Grossman 1972). Indeed, the literature on income dynamics is vast. Notable examples from it include Lillard and Willis (1978), Abowd and Card (1989), and Meghir and Pistaferri (2004). However, only recently have researchers started to explore the dynamics of health. Some studies, such as Adda, Banks and von Gaudecker (2007), Adams, et al. (2003) and Boersch-Supan, Heiss and Hurd (2005), investigate the joint dynamics of health and income using dynamic panel data techniques. These studies center largely on eliciting the causal pathways between health and economic status. Other studies, such as Contoyannis, Jones and Rice (2004) and Halliday (2007 and 2008), are more closely tied to the labor economics literature on income and employment dynamics

(*e.g.* Hyslop 1999). They focus exclusively on health status and emphasize the statistical properties of health dynamics by modeling health as a discrete variable and attempting to identify state dependence in the presence of unobserved heterogeneity. Nevertheless, despite recent progress, the literature on health dynamics is still very much a fledgling field.

We further this field by exploring health inequality - a topic that has largely been ignored in the literature. Particularly, we estimate a dynamic econometric model that allows us, not only to measure health inequality, but also to better understand its causes. To do this, we model health as a continuous latent variable which forms the basis of a survey respondent's self-reported health status (SRHS). Latent health depends on three factors: endowments that are inherited from childhood, and permanent and transitory shocks. Permanent shocks model events that leave a residue on a person's health such as the onset of chronic diseases (*e.g.* diabetes, arthritis, Parkinson's Disease, *etc.*) or irreversible injuries to cartilage and joints. Temporary shocks model events whose effects on health disappear after a period of time such as getting a cold, breaking a bone, etc. We estimate the model using a Simulated Minimum Distance (SMD) estimator and employing the Panel Study of Income Dynamics (PSID). Given our parameter estimates, we are able to decompose the variance of health into its three constituents. Moreover, because the model

incorporates a permanent shock, it can account for an important finding in the literature, namely, that health inequality, like consumption inequality, increases as cohorts age (see Deaton and Paxson 1994 and Deaton and Paxson 1998).

The balance of this paper is organized as follows. In Section 2, we discuss the data. In Section 3, we introduce our model of health dynamics and discuss the estimation procedure as well as some identification issues. In Section 4, we discuss our results. Finally, we conclude.

2 Data

We use a sample of Caucasian men aged 25 to 60 from the PSID. We use all waves from 1984 to 1997. We do not use data before 1984 because there is no information on SRHS available prior to 1984. We do not go past 1997 because the PSID was conducted every other year after 1997. Our main health measure is SRHS which is a categorical variable in which the respondent classifies their health into one of five categories: Excellent (SRHS = 1), Very Good (SRHS=2), Good (SRHS=3), Fair (SRHS=4), and Poor (SRHS=5). We map the SRHS variable into a binary variable. The top two categories are mapped into a one and the bottom three

categories are mapped into a zero.[1] We also use data on age and educational attainment. Descriptive statistics are reported in Table 1. A detailed discussion of our sample selection is provided in the appendix.

We defend our use of SRHS measures as follows. First, we are ultimately interested in a latent health index which, in turn, determines a person's assessment of their own health; SRHS is perfectly appropriate for this task. Moreover, there is substantial research that has shown that these measures of health correlate well with more objective health measures (Kaplan and Camacho 1983; Idler and Kasl 1995). Finally, many alternative health measures are not without flaws. For example, Baker, Stabile and Deri (2004) investigated the possibility of measurement errors in self-reported objective measures of health (such as those from the Health and Retirement Survey) by comparing them with medical records. They concluded that these measurement errors were often quite large and regrettably correlated with labor market activity.

3 Health Inequality

[1] This is not the standard partition of SRHS into a binary variable. However, when we used the standard partition in which the bottom two categories were mapped into zero and the remaining into unity, we had that at most 8% of the individual-time observations were classified as sick and, thus, there was substantially less transitioning into and out of health states which is crucial for our identification (see Section 3.2).

We begin by modeling individual i's latent health at age t. We postulate the following model for latent health (defined as $h^*_{i,t}$):

$$h^*_{i,t} = \delta_t + \gamma_i + u_{i,t} + \varepsilon_{i,t}. \tag{1}$$

There are three key terms in the model: endowments (denoted by γ_i), and a permanent and a transitory shock to health (denoted by $u_{i,t}$ and $\varepsilon_{i,t}$, respectively). In addition, to allow for a flexible treatment of the age profile of health, we include age effects (denoted by δ_t). The permanent component is modeled as a random walk with drift:

$$u_{i,t} = \eta + u_{i,t-1} + e_{i,t} \tag{2}$$

where $\eta < 0$ since health declines with age. We assume that the process begins at $t = 1$ and that $u_{i,0} = 0$. The process is observed until period T.

Each term in equation (1) has an interpretation. The endowments are individual specific and allow people to differ in their latent health. Endowments are personal characteristics that are formed early in life or inherited and affect a person's health. Their need in the health process is underscored by Halliday (2007 and 2008) whose estimations revealed a large role for unobserved heterogeneity in SRHS in the PSID. The term $e_{i,t}$ is a permanent shock to health and could represent events such as

onset of chronic illness or accidents that have lasting effects on a person's health. The term $\varepsilon_{i,t}$ models transitory shocks to health. Examples of these shocks could include mild bouts of illnesses such as the flu or broken bones. This specification of health dynamics is similar to many common specifications in the literature on income dynamics (*e.g.* Pistaferri and Meghir 2004; Abowd and Card 1989).

Stacking the permanent and transitory shocks as $e_i \equiv (e_{i,1}, ..., e_{i,T})$ and $\varepsilon_i \equiv (\varepsilon_{i,1}, ..., \varepsilon_{i,T})$, we assume that

$$\begin{pmatrix} \gamma_i \\ e_i \\ \varepsilon_i \end{pmatrix} \sim N \left(\mathbf{0}, \begin{bmatrix} \sigma_\gamma^2 & \mathbf{0}_{1,T} & \mathbf{0}_{1,T} \\ \mathbf{0}_{T,1} & \sigma_e^2 \mathbf{I}_T & \mathbf{0}_{T,T} \\ \mathbf{0}_{T,1} & \mathbf{0}_{T,T} & \mathbf{I}_T \end{bmatrix} \right). \tag{3}$$

This assumption implies that the permanent and transitory shocks are: (1) serially uncorrelated; (2) uncorrelated with each other; (3) homeskedastic. The parameter vector to be estimated is $\theta \equiv \left(\sigma_\gamma^2, \sigma_e^2, \eta, \delta_1, ..., \delta_T\right)$.[2]

We decompose the contributions of heterogeneity, and permanent and transitory shocks to health inequality as follows. Given our assumptions on the initial condi-

[2] In practice, we did not include an age dummy for every age in the estimations. We do this for two reasons. The first is that a complete set of age dummies would have been completely collinear with the age trend. The second is that the age effects prior to age 25 would not be identified due to a lack of data for these ages. Instead, we included a constant term and dummies for being older than ages 25 through 50 in 5 year increments.

tion, we write

$$h^*_{i,t} = \delta_t + \gamma_i + \eta t + \varepsilon_{i,t} + \sum_{s=0}^{t-1} e_{i,t-s} \tag{4}$$

which, in turn, implies

$$\sigma^2_{h^*_{i,t}} = \sigma^2_\gamma + t\sigma^2_e + 1. \tag{5}$$

At any point-in-time, health inequality depends on the variances of initial endowments and both types of shocks. This formula has several important implications. First, it tells us that inequality in latent health will increase as the cohort ages. This is a result that is consistent with empirical evidence on not only health inequality (Deaton and Paxson 1998), but also consumption inequality (Deaton and Paxson 1994; Primiceri and van Rens 2009; Storesletten, Telmer and Yaron 2004). Second, permanent shocks to health will explain an increasing portion of health inequality within a cohort.

3.1 Simulated Minimum Distance Estimation

Estimation of the parameter vector θ is hampered by the fact that the econometrician does not observe the individual's latent health stock. Rather, she observes

$$h_{i,t} = 1\left(h^*_{i,t} \geq 0\right). \tag{6}$$

We say that when $h_{i,t} = 1$ the person is healthy and that when $h_{i,t} = 0$ the person is unhealthy. To estimate θ, we employ an SMD procedure that is similar to Rust and Hall (2003) and is a special case of Simulated Generalized Method of Moments (SMM) discussed in McFadden (1989) and Stern (1997). This procedure matches simulated probabilities of health sequences to their counterparts in the data.[3]

The moment conditions are formed as follows. First, letting N denote the sample size, we draw $N * S$ values from a standard Normal distribution. Next, letting A denote the terminal age, we draw $2 \times N \times S \times A$ values from a standard Normal distribution. We chose A to be 60. The initial condition was set at age 1. These draws remained fixed throughout the optimization. Third, we choose a parameter vector θ. With this parameter vector, we simulated $N \times S$ values of the heterogeneity and $N \times S \times A$ values for the permanent and transitory shocks. For each individual, this resulted in a total of S sequences $\left(h_{i,1}^{s}, ..., h_{i,60}^{s}\right)$. Fourth, we calculate the probabilities of simulated sequences of outcomes for individual i as

[3] We did not employ Simulated Maximum Likelihood (SML) as it would have required too many simulations. The reason for this is that our model is not Markovian in the sense that $P\left(h_{i,t}|h_{i,t-1}, ..., h_{i,1}\right)$ depends on the entire history of outcomes. As such, SML would have required matching vectors of outcomes from the simulations and the data of high dimensions - sometimes as a high as 14. Because a 14 dimensional vector of binary outcomes can take on a total of 16384 values, a large number of simulations would have been required to match all possible outcomes in the data.

follows

$$
\begin{aligned}
& P\left(h_{i,t}^{s}=d_{1}, h_{i,t+1}^{s}=d_{2}\right) \\
&= S^{-1}\sum_{s=1}^{S} 1\left(h_{i,t}^{s}=d_{1}, h_{i,t+1}^{s}=d_{2}\right) \\
&\equiv P_{i,t}^{S}\left(d_{1}, d_{2}; \theta\right)
\end{aligned}
\tag{7}
$$

These probabilities were calculated for $t = 25, 30, 35, 40, 45, 50$ and 55. Because there are 4 health sequences at each age, this yields a total of 28 moment conditions. Fifth, the empirical moments can formed as

$$
\begin{aligned}
& N^{-1}\sum_{i=1}^{N}\left(1\left(h_{i,t}=d_{1}, h_{i,t+1}=d_{2}\right)-P_{i,t}^{S}\left(d_{1}, d_{2}; \theta\right)\right) \\
&\equiv P_{t}^{D}\left(d_{1}, d_{2}\right)-P_{t}^{S}\left(d_{1}, d_{2}; \theta\right)
\end{aligned}
\tag{8}
$$

where $P_{t}^{D}\left(d_{1}, d_{2}\right) \equiv N^{-1}\sum_{i=1}^{N} 1\left(h_{i,t}=d_{1}, h_{i,t+1}=d_{2}\right)$ and $P_{t}^{S}(d_{1}, d_{2}; \theta) \equiv N^{-1}\sum_{i=1}^{N} P_{i,t}^{S}(d_{1}, d_{2}; \theta)$. In this sense, using the moments in equation (8) matches probabilities calculated directly from the data to their simulated counterparts.

We can now calculate the objective function. We note that each of these moments can be written as $m_{j}(\theta) \equiv N^{-1}\sum_{i=1}^{N} h_{j,i}(\theta)$ for $j = 1, ..., J$. If we stack them into the

$J \times 1$ vector $m(\theta) \equiv N^{-1}\sum_{i=1}^{N} h_i(\theta)$, we obtain the SMD objective function

$$Q(\theta) = m(\theta)' \Omega m(\theta) \tag{9}$$

where Ω is a $J \times J$ weighting matrix. We used an estimate of $E\left[h_i(\theta) h_i(\theta)'\right]^{-1}$ for the weighting matrix. GMM theory states that this weighting matrix yields the smallest asymptotic variance. Details concerning its calculation can be found in the appendix. As shown in McFadden (1989) and Wooldridge (2002), the asymptotic variance of our estimator is given by

$$AVar\left(\widehat{\theta}\right) = \left(\frac{1}{N} + \frac{1}{NS}\right)\left(E\left[\nabla_\theta h_i(\theta)\right] E\left[h_i(\theta) h_i(\theta)'\right]^{-1} E\left[\nabla_\theta h_i(\theta)'\right]\right)^{-1}. \tag{10}$$

Standard errors are based on this formula. In practice, we set S equal to 250 at the minimum and so our standard errors will be almost numerically equivalent to those from standard GMM. We use the Nelder-Mead algorithm to optimize the objective function.

3.2 Identification of σ_e^2

We now discuss identification of the variance of the permanent shock in the absence of a transitory shock. For reasons that we will discuss, allowing for the presence of

transitory shocks creates many complications that would be better dealt with in a separate theoretical paper. We consider the probability

$$P\left(h_{i,t}=1|h_{i,t-1}=0;\sigma_e^2\right)=P\left(e_{i,t}\geq -h_{i,t-1}^*|h_{i,t-1}^*<0;\sigma_e^2\right). \tag{11}$$

Next, we note the following

$$h_{i,t-1}^*\sim N\left(\delta_{t-1}+\eta\left(t-1\right),\sigma_\gamma^2+(t-1)\sigma_e^2+1\right).$$

Now, using the Law of Iterated Expectations, we obtain that

$$P\left(h_{i,t}=1|h_{i,t-1}=0;\sigma_e^2\right)=\int_{h_{i,t-1}^*<0}\left[1-\Phi\left(\frac{-h_{i,t-1}^*}{\sqrt{\sigma_e^2+2}}\right)\right]dF\left(h_{i,t-1}^*;\sigma_e^2\right)=$$
$$\int_{h_{i,t-1}^*<0}\Phi\left(\frac{h_{i,t-1}^*}{\sqrt{\sigma_e^2+2}}\right)dF\left(h_{i,t-1}^*;\sigma_e^2\right).$$

The presence of the transitory shock would have complicated matters because it would have required working with the distribution of $e_{i,t}+\Delta\varepsilon_{i,t}$ conditional on $h_{i,t-1}^*$. This would have greatly complicated the expression inside the normal CDF as $\Delta\varepsilon_{i,t}$ is correlated with $h_{i,t-1}^*$. Next, consider two values for σ_e^2 given by σ_1^2 and σ_2^2 such

that $\sigma_1^2 > \sigma_2^2$. Then we will have that

$$
\begin{aligned}
& P\left(h_{i,t}=1|h_{i,t-1}=0;\sigma_1^2\right)-P\left(h_{i,t}=1|h_{i,t-1}=0;\sigma_2^2\right)= \\
& \int_{h_{i,t-1}^*<0}\left[\Phi\left(\frac{h_{i,t-1}^*}{\sqrt{\sigma_1^2+2}}\right)dF\left(h_{i,t-1}^*;\sigma_1^2\right)-\Phi\left(\frac{h_{i,t-1}^*}{\sqrt{\sigma_2^2+2}}\right)dF\left(h_{i,t-1}^*;\sigma_2^2\right)\right]= \\
& \int_{h_{i,t-1}^*<0}\left[\Phi\left(\frac{h_{i,t-1}^*}{\sqrt{\sigma_1^2+2}}\right)-\Phi\left(\frac{h_{i,t-1}^*}{\sqrt{\sigma_2^2+2}}\right)\right]dF\left(h_{i,t-1}^*;\sigma_1^2\right)+ \\
& \int_{h_{i,t-1}^*<0}\Phi\left(\frac{h_{i,t-1}^*}{\sqrt{\sigma_2^2+2}}\right)\left[dF\left(h_{i,t-1}^*;\sigma_1^2\right)-dF\left(h_{i,t-1}^*;\sigma_2^2\right)\right].
\end{aligned}
\tag{12}
$$

The first term after the second equality is positive since a CDF is monotonic increasing, $\sigma_1^2 > \sigma_2^2$ and $h_{i,t-1}^* < 0$. The second term is also positive if δ_{t-1} or $\eta(t-1)$ is sufficiently negative. To see this, we make two observations: the first is that $\Phi(x)$ is convex for $x < 0$ and the second is that $F\left(h_{i,t-1}^*;\sigma_1^2\right)$ is second order stochastically dominated by $F\left(h_{i,t-1}^*;\sigma_2^2\right)$. Also, we note that the second term is essentially the difference in the expectation of $\Phi(x)$ for $x < 0$ under two different distributions. The reason is that, for δ_{t-1} or $\eta(t-1)$ sufficiently negative, $\delta_{t-1}+\eta(t-1)$ will be sufficiently below zero and so that vast majority of values for $h_{i,t-1}^*$ will be negative. Hence, the integral in the second term is essentially over the entire support of $h_{i,t-1}^*$. Because a mean-preserving spread increases the expectation of a convex function, the second term is also positive. This result is analogous to a discussion in Meghir and Pistaferri (2004) of how higher income volatility reduces poverty persistence.

Thus, the key to identifying σ_e^2 is that $P\left(h_{i,t}=1|h_{i,t-1}=0;\sigma_e^2\right)$ is increasing in the variance of the permanent shock for sufficiently unhealthy people. Consequently, for these people, when we compare these probabilities from the simulations and the data, there will be a unique σ_e^2 that equates the two quantities.

4 Empirical Results

Table 2 reports the parameter estimates. We estimate the model for three groups: people with college degrees; people with more than twelve years of schooling, but no college degree; and people with at most twelve years of schooling. First, we note that the standard errors of all of the parameter estimates are small. This suggests that the objective function is not flat in the vicinity of its maximum and, so the parameters are all well identified. Second, we observe that the variance of the endowments is substantially higher for people with college degrees than those without; its estimate is 2.4278 for the college educated, 0.4285 for people with some education beyond high school and 0.7325 for people with at most 12 years of schooling. Third, we observe that the variance of the permanent shock is over ten times higher than for people without a college degree than it is for people with a degree. This is consistent with results in Case and Deaton (2005) who show that health declines more rapidly for the less educated. One interpretation that the authors provide for this is that people

with less education are more apt to be employed in blue collar occupations that take a larger toll on their bodies. One implication of this is that health inequality will rise more rapidly for people without college degrees.[4]

We now turn to the model's fit. Figures 1 through 3 plot the age profiles of being in good health from the simulations and from the data for our three sub-samples. Visual inspection of these figures suggests that the fit of the model is good. The third and second to last rows of Table 1 report the value of the object function evaluated at the maximum and the corresponding J-statistic (equal to the sample size multiplied by the objective function evaluated at the optimizer). The J-statistic is distrusted χ^2_{18}.[5] The simulated moments do come close to the moments in the data as evidenced by the reported values of $Q\left(\widehat{\theta}\right)$ between 0.0567 and 0.1434. Despite this, the tests for over-identification reject the null that all of our moment conditions are valid.

To give the reader a better idea of the relative importance of endowments *vis-a-vis* permanents and transitory shocks, we decompose the variance using equation (5) in Figures 4 through 5. Specifically, at a given age, we plot variances of the

[4] We checked the robustness of our estimates by re-optimizing the objective function using different simulations and starting from different initial parameter values. We obtained similar parameter estimates.

[5] We used 28 moment conditions and estimated 10 parameters. Note that the Table 2 does not report 6 of the age effects.

endowments, transitory shocks, and the cumulative sum of permanent shocks as a percentage of $\sigma^2_{h^*_t}$. The key difference in these figures is the role that permanent shocks play. For the college educated, permanent shocks account for well under 10% of the variation in health. In fact, for this demographic, permanent shocks matter less than temporary shocks! The biggest source of variation for this subgroup is endowments. In contrast, permanent shocks matter substantially more for people without college degrees. For people with some post-secondary education, permanent shocks account for between 45% and 70% of the variation in health status. For people with at most a high school diploma, they account for for between 35% and 60% of total variation. These figures suggest that cumulative sum of health events in adulthood takes a higher toll on those who are less educated.

5 Conclusion

In this paper, we investigated the evolution of health inequality over the life-course. We modeled health as a latent variable that is determined by endowments, and permanent and transitory shocks. This latent variable, in turn, determined a person's self-reported health. To estimate the model, we employed Simulated Minimum Distance and the Panel Study of Income Dynamics.

Estimation revealed that permanent shocks matter more for those with less education. For people with college degrees, permanent shocks account for under 10% of the variation in health. In contrast, for people without college degrees, they account for between 35% and 70% of total variation. Consistent with this, we find that health inequality rises more rapidly for the less educated. One interpretation of this finding is that events that occur in adulthood have relatively larger impacts on health for less educated people. One might expect this to be the case since people with less education are more apt to be employed as manual laborers.

There are several future avenues for research that builds upon the framework developed in this paper. First, we will estimate the effects of various health shocks on income. Second, we will incorporate a model of the joint dynamics of health and income into a life-cycle model of consumption to better understand the effects of health inequality on consumption inequality. Such an exercise will provide a structural mechanism linking the results in Deaton and Paxson (1994) and Deaton and Paxson (1998).

6 Appendix: Sample Selection

We first extracted all individuals from the 1984 to 1997 waves of the PSID who were ever heads of household. The initial sample size was 17,326 individuals. We then

dropped all women from the sample. This lowered the sample size to 11,415. Next, we dropped people with incomplete health data which brought the sample size to 11,387. After this, we kept only Caucasians resulting in sample size of 7,232. We then dropped people who were not in the panel continuously. This lowered the sample size to 6,905. Next, we kept only people between ages 25 and 60, inclusive. This brought the sample size to 5,669. Finally, we dropped the SEO giving us our final sample size of 4,910. Of these 4,910 people, 1,335 were college graduates; 865 had some education beyond high school, but no college degree; 2,308 had 12 or fewer years of schooling. There were an additional 102 people who were missing educational information.

7 Appendix: Calculation of $E\left[h_i(\theta)\, h_i(\theta)'\right]$

The vector $h_i(\theta)$ consists of J elements $h_{j,i}(\theta)$ that can be written as $1(y_{i,j} = d) - p_j$ where $p_j \equiv P(y_{i,j} = d)$. The variables $1(y_{i,j} = d)$ are indicators for events such as $\{h_{i,t} = d_1, h_{i,t+1} = d_2\}$. Note that $E[1(y_{i,j} = d)1(y_{i,k} = d)]$ for $j \neq k$ will either be zero, close to zero and/or hard to estimate accurately in the data. For example, we will have that

$$P\left(\{h_{i,t} = 1, h_{i,t+1} = 1\} \cap \{h_{i,t} = 0, h_{i,t+1} = 0\}\right) = 0 \tag{13}$$

since the two events are mutually exclusive. In addition, probabilities such as

$$P\left(\{h_{i,t}=d_1, h_{i,t+1}=d_2\}\cap\{h_{i,t+4+s}=d_1', h_{i,t+5+s}=d_2'\}\right) \text{ for } s\geq 1 \tag{14}$$

be very hard to estimate accurately in our data. For example, for $s=1$ and $t=30$, we would require people to be present in our data from ages 30 to 37; there are very few of these. For $s=11$, we would require people to be present from ages 30 to 46; as our panel is only 14 years long, there are no people who satisfy this criterion. Finally, many events given by (14) will have very small probabilities that are near zero. Accordingly, if we ignore the events given by (14) and stack the J probabilities, p_j , into a $J\times 1$ vector P, we then obtain that $E\left[h_i(\theta)\,h_i(\theta)'\right]\approx diag(P)-P*P'$ where $diag(P)$ is a diagonal matrix with jth, jth element p_j.

References

[1] Abowd JM, Card D. On the Covariance Structure of Earnings and Hours Changes. Econometrica 1989; 57: 411-45.

[2] Adams HP, Hurd MD, McFadden D, Merrill A, Ribeiro T. Healthy, Wealth and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status. Journal of Econometrics 2003; 112: 3-56.

[3] Adda J, Banks J, von Gaudecker HM. The Impact of Income Shocks on Health: Evidence from Cohort Data 2007; forthcoming Journal of the European Economic Association.

[4] Baker M, Stabile M, Deri C. What Do Self-Reported Objective Measures of Health Measure? Journal of Human Resources 2004; 39: 1067-1093.

[5] Börsch-Supan A, Heiss F, Hurd M. Healthy, Wealthy and Knowing Where to Live: Predicted Trajectories of Health, Wealth and Living Arrangements Among the Oldest Old 2003; NBER Working Paper No. 989.

[6] Case A and Deaton A. Broken Down by Work and Sex: How Our Health Declines. in Analyses in the Economics of Aging. ed. David Wise. 2005. University of Chicago Press: Chicago.

[7] Contoyannis P, Jones AM, Rice N. Simulation-Based Inference in Dynamic Panel Probit Models: An Application to Health. Empirical Economics 2004; 29: 49-77.

[8] Deaton A, Paxson CP. Intertemporal Choice and Inequality. Journal of Political Economy 1994; 102: 437-67.

[9] Deaton A, Paxson CP. Health, Income and Inequality over the Life-Cycle. in Frontiers of Aging. ed. David Wise. 1998. University of Chicago Press: Chicago.

[10] Grossman M. On the Concept of Health Capital and the Demand for Health. Journal of Political Economy 1972; 80: 223-255.

[11] Halliday T. Testing for State Dependence with Time-Variant Transition Probabilities. Econometric Reviews 2007; 26: 685-703.

[12] Halliday T. Heterogeneity, State Dependence and Health. Econometrics Journal 2008; 11: 499-516.

[13] Hyslop D. State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women. Econometrica 1999; 67: 1255-1294.

[14] Idler EL, Kasl SV. Self-Ratings of Health: Do They Also Predict Changes in Functional Ability? Journal of Gerontology 1995; 50: S344-S353.

[15] Kaplan GA, Camacho T. Perceived Health and Mortality: A 9 Year Follow-up of the Human Population Laboratory Cohort. American Journal of Epidemiology 1983; 177: 292.

[16] Lillard LA, Willis RJ. Dynamic Aspects of Earnings Mobility. Econometrica 1978; 46: 985-1012.

[17] Meghir C, Pistaferri L. Income Variance Dynamics and Heterogeneity. Econometrica 2004; 72: 1-32.

[18] McFadden D. A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration. Econometrica 1989; 57: 995-1026.

[19] Palumbo MG. Uncertain Medical Expenses and Precautionary Savings Near the End of the Life-Cycle. Review of Economic Studies 1999; 66: 395-421.

[20] Primiceri G, van Rens T. Heterogeneous Life-Cycle Profiles, Income Risk and Consumption Inequality. Journal of Monetary Economics 2009; 56: 20-39.

[21] Rust J, Hall G. Simulated Minimum Distance Estimation of a Model of Optimal Commodity Price Speculation with Endogenously Sampled Prices 2003; unpublished mimeo.

[22] Stern S. Simulation-Based Estimation. Journal of Economic Literature 1997; 35: 2006-2039.

[23] Storesletten K, Telmer CI, Yaron A. Consumption and Risk Sharing over the Life-Cycle. Journal of Monetary Economics 2004; 51: 609-633.

[24] Wooldridge J. Econometric Analysis of Cross-Sectional and Panel Data 2003; MIT Press: Cambridge, MA.

[25] Yogo M. Portfolio Choice in Retirement: Health Risk and the Demand for Annuities, Housing and Risky Assets 2008; unpublished mimeo.

Table 1: Descriptive Statistics

	Mean (Standard Deviation)
SRHS	2.122 (1.002)
Years of Education	13.262 (2.745)
Years of Education > 12	0.541 (0.498)
College Degree	0.345 (0.475)
Age	39.620 (9.379)

Table 2: Parameter Estimates

	College Degree	Years of Ed > 12, No Degree	Years if Ed <= 12
Constant	3.3918	3.5425	2.4141
	(0.0916)	(0.0780)	(0.0443)
σ_γ^2	2.4278	0.4285	0.7325
	(0.1632)	(0.0622)	(0.0377)
η	−0.0343	−0.0630	−0.0465
	(0.0017)	(0.0018)	(0.0011)
σ_e^2	0.0033	0.0482	0.0403
	(0.0004)	(0.0030)	(0.0020)
$Q\left(\widehat{\theta}\right)$	0.0567	0.14344	0.0688
J-Statistic	75.91	124.07	158.82
	[0.0000]	[0.0000]	[0.0000]
S	500	750	250
N	1335	865	2308

Figure 1: $P\left(h_{i,t}=1\right)$ - College Degree

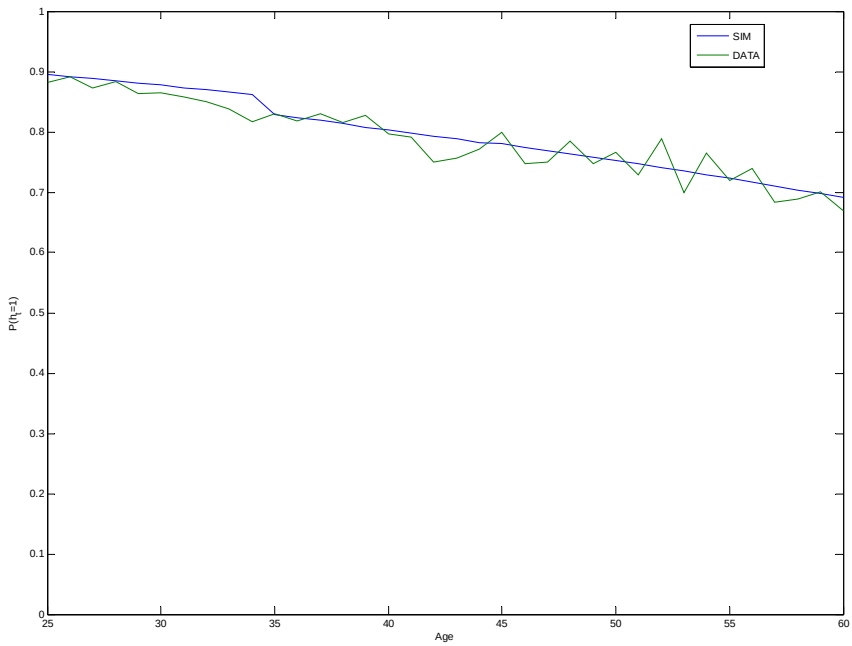

Figure 2: $P\left(h_{i,t}=1\right)$ - Years of Ed >1 2, No Degree

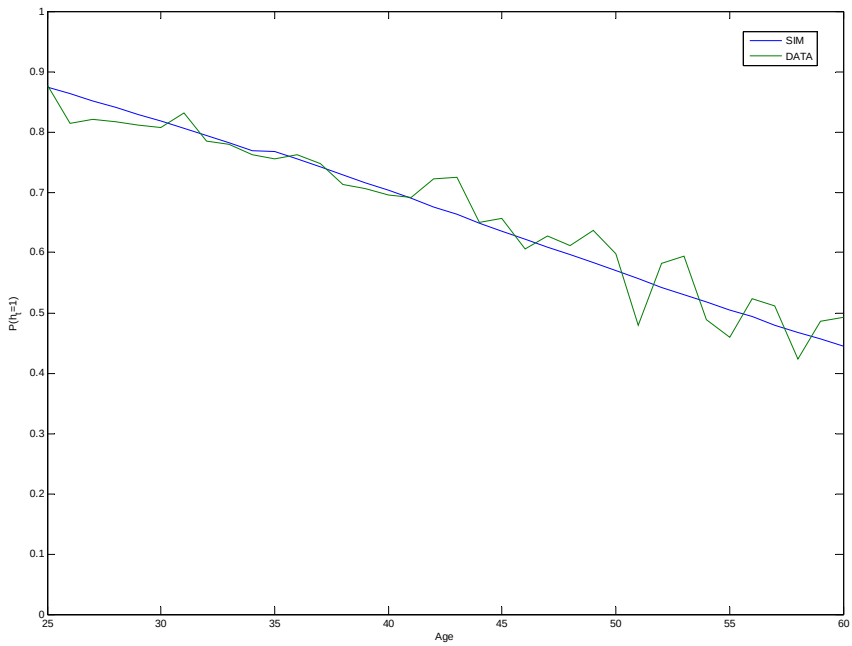

Figure 3: $P\left(h_{i,t}=1\right)$ - Years of Ed <=12

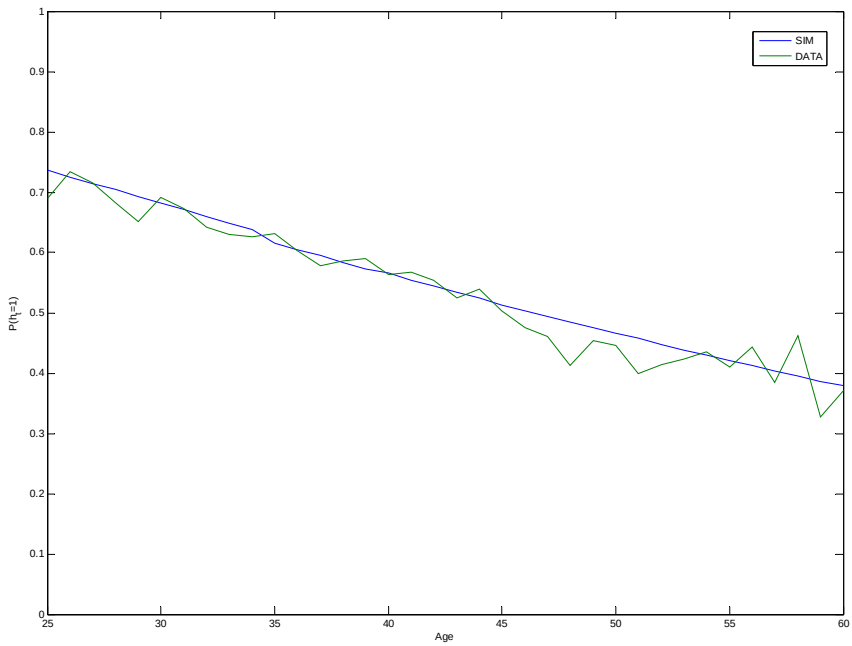

Figure 4: Variance Decomposition- College Degree

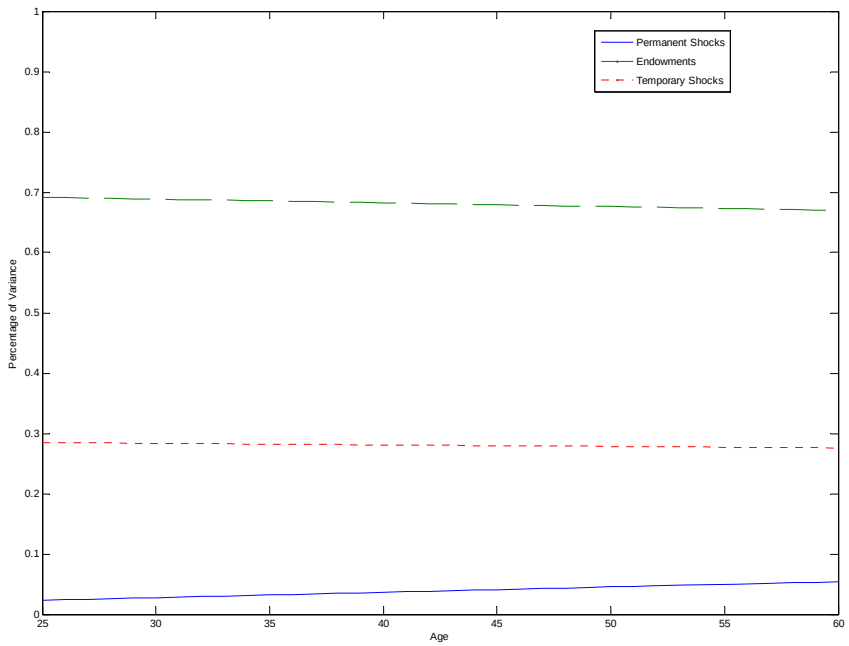

Figure 5: Variance Decomposition - Years of Ed >1 2, No Degree

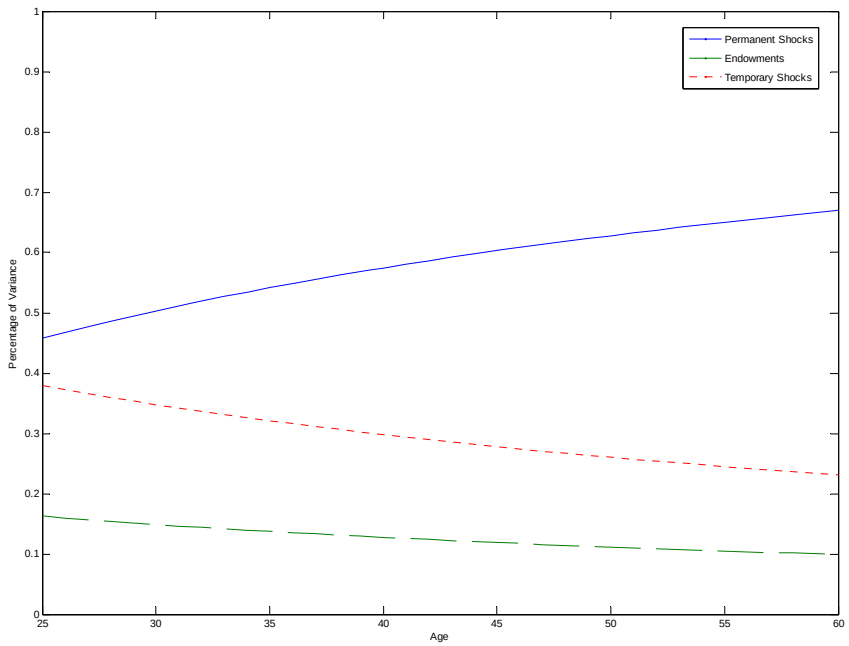

Figure 6: Variance Decomposition - Years of Ed <=12