Heterogeneity, State Dependence and Health

by Timothy J. Halliday, University of Hawaii-Manoa and John A. Burns School of Medicine

Working Paper No. 05-03

February 2, 2005

Abstract

In this paper, we use longitudinal data on Self-Reported Health Status from the Panel Study of Income Dynamics to estimate a model of the evolution of health over the life-cycle. The model allows for two sources of persistence in health: unobserved heterogeneity, which models an individual's (unobserved) ability to cope with health shocks, and state dependence, which models the extent to which the ability to cope with health shocks depends on health status. We allow for flexibility in both sources of persistence. Estimation indicates that heterogeneity is an important determinant of health suggesting that a person's health today has important antecedents earlier on in life. We also find evidence of state dependence. However, its magnitude depends crucially on the individual's age and unobserved heterogeneity. The relative contributions of heterogeneity and state dependence that we uncover have different implications for how health policy should be conducted.

JEL Classification: I1, C5
Keywords: health, dynamic panel data models, gradient.

Corresponding Author: Timothy J. Halliday, Department of Economics, University of Hawaii-Manoa, 2424 Maile Way, Saunders Hall Room 533, Honolulu, Hawaii, U.S.A. 96822 [phone] (808) 956 8615 <halliday@hawaii.edu>

1 Introduction

This paper explores the notion that an individual's health follows a persistent stochastic process. Specifically, we concern ourselves with two tasks. The first is to gain a better understanding of the appropriate way of modeling the evolution of health over the life-cycle. This is important because, while many empirical studies in Economics have investigated the dynamics of both the level of earnings (Lillard and Willis 1978; Abowd and Card 1989) and, more recently, the variance of earnings (Meghir and Pistaferri 2004), few have investigated the dynamics of health.[1] It has been noted by many researchers such as Deaton (1992) and Caballero (1990) that different assumptions about the stochastic process governing income can imply very different types of life-cycle consumption behavior. Accordingly, as health status becomes a more common state variable in structural models, it is becoming increasingly more important that researchers arrive at a better understanding of health dynamics.[2] The second task of this paper is to quantify the relative contributions of unobserved heterogeneity (fixed effects), and state dependence in the determination of health. Doing so is important as each will have very different implications for how health policy should be conducted.

To analyze health dynamics, we utilize data on Self-Reported Health Status (SRHS) from

[1] Contoyannis, Jones and Rice (2004) and Contoyannis, Jones and Leon-Gonzalez (2004) are notable exceptions.

[2] For examples of structural model using health as a state variable, see Rust and Phelan (1997) and Arcidiacono, Heig and Sloan (2004).

the Panel Study of Income Dynamics (PSID). We observe that SRHS is highly persistent in the PSID. Indeed, in our data, the first order auto-correlation of a dummy variable indicating that an individual reports fair or poor health is 0.5661 for men and 0.5643 for women. Simple linear AR(1) regressions using this dummy variable yield t-statistics of 56.33 for men and 62.87 for women. However, while these correlations do indicate a high degree of persistence, they are not informative of the underlying stochastic properties of the health process.

To gain additional insight, we model the evolution of health over the life-cycle as a first order Markov process. The model allows for two sources of persistence: unobserved heterogeneity and state dependence. Unobserved heterogeneity models an individual's (unobserved) ability to cope with idiosyncratic health shocks such as accidents or exposure to disease-causing agents. In our model, not only is the constant term heterogeneous, but the coefficient on lagged health and the coefficients on all functions of age are also heterogeneous. Accordingly, we allow for a great deal of flexibility in heterogeneity. The second source of persistence in the model is state dependence which models the degree to which an individual's ability to cope with a given health shock depends on her health status. State dependence in health captures the idea that people who are ill are less able to cope with health shocks than people who are well. We model state dependence in a flexible manner by allowing for heterogeneity in the coefficient on lagged health as well as heterogeneous interactions between age and lagged health status. This approach contrasts with much of the applied literature on dynamic panel data models which, typically, only allows for heterogeneity in the constant term and, usually, models state dependence as a homogeneous function of the lagged state.[3]

[3]For example, see Magnac (2000), Chay, Hoynes and Hyslop (2001), Contoyannis, Jones and Rice (2004) and Hyslop (1999).

Estimation yields many interesting findings. First, we see that the data favor models with simple homogeneous quadratic or linear functions of age over the more complicated (and more computationally intensive) models with richer forms of heterogeneity and state dependence. This should be consoling to structural modelers who use health status as a state variable and are concerned about the computational tractability of allowing for richer forms of serial correlation in the process. Next, we find large variation in health dynamics within both men and women suggesting that unobserved heterogeneity is an important determinant of health status. The contribution of heterogeneity suggests that a person's health status later on in life has important antecedents earlier in life. Finally, we find evidence of large degrees of state dependence. However, its magnitude depends critically on unobserved heterogeneity and age. Specifically, we find that individuals who are "health-deprived" (*i.e.* people who are innately less able to cope with health shocks) exhibit relatively more state dependence earlier in life than their healthier counterparts. In contrast, in old age, the pattern is reversed so that the degree of state dependence is significantly higher among individuals who are "health-endowed" (*i.e.* people who are innately better able to cope with health shocks).

The relative contributions of heterogeneity and state dependence have important implications for policy. The fact that heterogeneity is an important contributor to health status strengthens the case for policies that focus on babies and children - whose fixed effects are still in formation. Examples of such policies include improving the health of pregnant mothers as is suggested by Barker (1997) and Barker, *et al* (1989) and improving the socioeconomic conditions of households in which children are reared as is suggested by Case, Lubotsky and Paxson (2002). The fact that we also find evidence of state dependence suggest that there is also a role for medical

interventions which target people in adulthood in improving health outcomes. However, our estimates have important implications for exactly who will benefit from such interventions. If interventions target people later on in life then their benefits will disproportionately accrue to people who are health-endowed. If interventions target people in middle age or younger then their benefits will tend to go to people who are health-deprived. Accordingly, if it is the aim of medical interventions to affect the health of those who are the most disadvantaged in their ability to cope with health shocks then it is probably better to act sooner rather than later and target people who are younger than age 50.

The balance of this paper is organized as follows. Section 2 describes the data. Section 3 sets up our model of health dynamics. Section 4 describes our estimation procedure. Section 5 investigates which models are favored by the data. Section 6 discusses the role that heterogeneity plays in the determining health. Section 7 quantifies the degree of state dependence in health. Section 8 concludes.

2 Data

We use data from the PSID from 1984 to 1997. The variables that we work with are SRHS, age and gender. During these years, the SRHS question was only asked of heads of household and their spouses and, thus, our sample is restricted to these individuals. In addition, because the SRHS was not asked prior to 1984, we do not have any data before 1984. The PSID contains an over-sample of low-income families called the Survey of Economic Opportunity (SEO). Because the sample was chosen based on income, we follow Lillard and Willis (1978) and drop it due to

endogenous selection.[4]

We use SRHS as our measure of health. SRHS is a categorical variable that takes on integer values between 1 and 5. 1 means that the individual perceives that their health is excellent; 2 is very good; 3 is good; 4 is fair; 5 is poor. While these data are subjective measures, there is an extensive literature that has shown a strong link between SRHS and health outcomes such as mortality and the prevalence of disease (Mossey and Shapiro 1982; Kaplan and Camacho 1983; Idler and Kasl 1995; Smith 2004). To lower the number of parameters that we estimate, we map the 5-point variable into a 2-point variable. Accordingly, we map all self-reports of fair or poor into unity and all other reports into zero. This is the conventional way of partitioning the SRHS variable in the health economics and epidemiology literatures.

We restrict our sample to individuals between ages 22 and 60. We do not include people younger than age 22 because there are not that many household heads younger than this age. We do not include people older than age 60 to mitigate any possible bias resulting from attrition due to mortality. We drop individuals whose age declines across successive survey years. We also drop individuals whose age increases by more than two years across successive survey years. Finally, we restrict our sample to white men and women. Table 1 reports the descriptive statistics from the resulting sample.

[4] Meghir and Pistaferri (2004) include the SEO in their work. Their reason for its inclusion is that by estimating linear models in differences, the unobserved heterogeneity is purged from the regression equation and, consequently the initial condition problem posed by the endogenous selection of the SEO is solved. However, this will not work in our case as we work with a non-linear model in which it is impossible to purge the model of any fixed effects.

3 Model

We let $h_{i,t} \in \{0,1\}$ denote the health of individual i at age t. Throughout this paper, we refer to individuals for whom $h_{i,t} = 1$ as ill and individuals for whom $h_{i,t} = 0$ as well. We assume that health evolves according to the following process:

$$h_{i,t} = 1(\alpha_i + \gamma_i h_{i,t-1} + \boldsymbol{\rho}_i' \mathbf{T} + \phi_i (t * h_{i,t-1}) + \varepsilon_{i,t} \geq 0) \tag{1}$$

where $\mathbf{T} = [t, t^2]'$. Equation (1) allows for four determinants of health: idiosyncratic risk, aging, state dependence and heterogeneity. We now describe each determinant of health.

Idiosyncratic Risk Idiosyncratic risk is represented by $\varepsilon_{i,t}$. It models illness-causing agents or events that affect individual i when he is of age t. Throughout this paper, we will refer to $\varepsilon_{i,t}$ as a "health shock." $\varepsilon_{i,t}$ can include accident occurrence, disease onset and exposure to bacteria and viruses. Because the effects of these agents and/or events can vary considerably in their intensity, $\varepsilon_{i,t}$ can assume a continuum of values. We assume that $\varepsilon_{i,t}$ is independent of $(\alpha_i, \gamma_i, \boldsymbol{\rho}_i', \phi_i, h_{i,0})$ and that it is distributed *i.i.d.* across time with a logistic distribution.[5] These assumptions imply that

$$P(h_{i,t} = 1 | h_{i,t-1}, ..., h_{i,0}, \theta_i) = \frac{\exp(\boldsymbol{\theta}_i' \mathbf{Z}_{i,t-1})}{1 + \exp(\boldsymbol{\theta}_i' \mathbf{Z}_{i,t-1})} \tag{2}$$

where $\boldsymbol{\theta}_i \equiv (\alpha_i, \gamma_i, \boldsymbol{\rho}_i', \phi_i)'$ and $\mathbf{Z}_{i,t-1} = (1, h_{i,t-1}, \mathbf{T}', t * h_{i,t-1})'$.

[5] Attempts have been made by Hyslop (1999) and Contoyannis, Jones and Rice (2004) to relax the *i.i.d.* assumption by allowing for serial correlation in $\varepsilon_{i,t}$. However, both attempts resulted in negative estimates of the serial correlation in $\varepsilon_{i,t}$. Both papers conclude that this finding is odd and probably reflects an identification issue in these more complicated models.

Aging The coefficients $\boldsymbol{\rho}_i'$ and ϕ_i are our aging coefficients. They allow the effects of accidents and exposure to illness-causing agents on $h_{i,t}$ to increase with age. In addition, they account for the fact that many diseases like Alzheimer's, cancer, hypertension and heart disease are more likely to manifest later in life. We allow for a quadratic in age to allow for flexibility in aging. In addition, we allow for additional flexibility in aging by allowing it to depend on the individual's health state and to be individual-specific. State dependent aging is modeled by the term $\phi_i(t * h_{i,t-1})$ in equation (1).

State Dependence The coefficients γ_i and $\boldsymbol{\phi}_i$ are our state dependence coefficients. They model the notion that a person who are ill may be less able to cope with a given health shock than when that same person is well. To give a concrete (albeit extreme) example, exposure to a flu virus is more likely to affect a person's health if she is HIV positive than if she is HIV negative. We allow for a richer form of state dependence by allowing it to vary with age. If $\boldsymbol{\phi}_i \neq 0$, then the degree of state dependence will vary with age. Health will exhibit positive state dependence at age t if $\gamma_i + \phi_i t > \mathbf{0}$.

Heterogeneity Our model allows for a large degree of flexibility in heterogeneity. We do this by letting the vector, $\boldsymbol{\theta}_i$, vary across individuals. So, not only is there heterogeneity in the "constant" term, α_i, but there is also heterogeneity in all of the model's parameters. This contrasts with the majority of the dynamic panel data literature in which, typically, only α_i is individual-specific. $\boldsymbol{\theta}_i$ model an individual's resistance to health shocks. Borrowing some jargon from Epidemiology, sometimes, we refer to these parameters as "host resistance."

We argue that allowing for more flexibility in the heterogeneity is important when modeling

the evolution of health for two reasons. First, we can write part of the index in equation (1) as

$$\alpha_i + \gamma_i h_{i,t-1} = \kappa_{i,0}(1 - h_{i,t-1}) + \kappa_{i,1} h_{i,t-1} \tag{3}$$

where $\gamma_i \equiv \kappa_{i,1} - \kappa_{i,0}$ and $\alpha_i \equiv \kappa_{i,0}$. What this calculation tells us is that, as long as we expect to see heterogeneity in both the persistence of illness, or the transition from $h_{i,t-1} = 1$ to $h_{i,t} = 1$, and in the onset of illness, or the transition from $h_{i,t-1} = 0$ to $h_{i,t} = 1$, then we should expect heterogeneity in both the constant coefficient (α_i) and the primary state dependence coefficient (γ_i). Second, we allow for heterogeneous aging coefficients ($\boldsymbol{\rho}_i'$ and $\boldsymbol{\phi}_i$) to allow for the possibility that decreases in host resistance with age will vary across individuals.

4 MLE

We estimate the model in equation (1) using an MLE procedure which has been discussed in Heckman (1981a and 1981b). Individual i ($i = 1, ..., N$) experiences $h_{i,t}$ at time $t \in \{0, ..., T_i\}$. However, the econometrician only observes $h_{i,t}$ for $t \in \{\tau_i, ..., T_i\}$ where $\tau_i \geq 0$, and, thus, we have an initial conditions problem. The procedure that we use accounts for this.

We now construct the likelihood function. The likelihood of a sequence of health outcomes conditional on $(\boldsymbol{\theta}_i', h_{i,\tau_i})$ for individual i for $t = \tau_i, ..., T_i$ is given by

$$P(h_{i,T_i}, ..., h_{i,\tau_i+1} | h_{i,\tau_i}, \boldsymbol{\theta}_i') = \prod_{t=\tau_i+1}^{T_i} \Lambda\left(\boldsymbol{\theta}_i' \mathbf{Z}_{i,t-1}(2h_{i,t} - 1)\right). \tag{4}$$

We assume that the heterogeneity vector has a discrete support where it can take on one of A values so that $\boldsymbol{\theta}_i \in \{\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_A\}$. The probability weight that is associated with each point of support is π_a. This approach is similar to Heckman and Singer (1984) who use a discrete distribution to approximate the distribution of unobserved heterogeneity when estimating duration models in the presence of heterogeneity via nonparametric maximum likelihood.[6] With some abuse of notation, let $P_{\tau_i}(h_{i,\tau_i}|\boldsymbol{\theta}'_a)$ denote the probability of the first observation conditional on $\boldsymbol{\theta}_i = \boldsymbol{\theta}_a$. We can now obtain the unconditional likelihood of observing $(h_{i,\tau_i}, ..., h_{i,T_i})$ via

$$\begin{aligned} &P(h_{i,T_i}, ..., h_{i,\tau_i}) = \\ &\sum_{a=1}^{A} P(h_{i,T_i}, ..., h_{i,\tau_i}|\boldsymbol{\theta}'_a)\pi_a = \\ &\sum_{a=1}^{A} \prod_{t=\tau_i+1}^{T_i} \Lambda(\boldsymbol{\theta}'_a \mathbf{Z}_{i,t-1}(2h_{i,t}-1)) P_{\tau_i}(h_{i,\tau_i}|\boldsymbol{\theta}'_a)\pi_a. \end{aligned} \tag{5}$$

Summing over the heterogeneity addresses the incidental parameters problem (Neyman and Scott 1948).

Our model implies a recursive definition for $P_{\tau_i}(h_{i,\tau_i}|\boldsymbol{\theta}'_a)$. To compute this probability, first, we let the probability of being well in $t = 0$ conditional on $\boldsymbol{\theta}_a$ be given by $p_a \equiv P_0(h_{i,0} = 0|\boldsymbol{\theta}'_a)$.

[6] Heckman and Singer (1984) verify that Kiefer-Wolfowitz conditions are satisfied for a general class of duration models with unobserved heterogeneity. These conditions ensure consistent estimation of the distribution of unobserved heterogeneity and structural parameters in this class of duration models. In addition, they cite a theorem from Lindsay (1983) that says that the nonparametric maximum likelihood estimate of the structural parameters and distribution of the unobserved heterogeneity will be such that the estimate of the heterogeneity distribution is discrete. Our approach contrasts with Heckman and Singer since we assume that the population distribution of the heterogeneity is discrete and, hence, we do not need to verify the Kiefer-Wolfowitz conditions.

The probability of observing $h_{i,t}$ conditional on $\boldsymbol{\theta}_a$ in any subsequent period is then given by

$$\begin{aligned} P_t(h_{i,t}|\boldsymbol{\theta}'_a) &= \sum_{d=0}^{1} P_t(h_{i,t}|h_{i,t-1}=d,\boldsymbol{\theta}'_a)P_{t-1}(h_{i,t-1}=d|\boldsymbol{\theta}'_a) \\ &= \sum_{d=0}^{1} \Lambda((\alpha_a+\gamma_a d+\boldsymbol{\rho}'_a\mathbf{T}+\phi_i t*d)(2h_{i,t}-1))P_{t-1}(h_{i,t-1}=d|\boldsymbol{\theta}'_a). \end{aligned} \tag{6}$$

Substituting, we get

$$\begin{aligned} P_t(h_{i,t}|\boldsymbol{\theta}'_a) = &\left(\sum_{d=0}^{1}\Lambda((\alpha_a+\gamma_a d+\boldsymbol{\rho}'_a\mathbf{T}+\phi_i t*d)(2h_{i,t}-1))\right) \\ &\left(\sum_{d=0}^{1}\Lambda((\alpha_a+\gamma_a d+\boldsymbol{\rho}'_a(\mathbf{T}-\mathbf{1})+\phi_i(t-1)*d)(2h_{i,t-1}-1))\right)... \\ &\left(\sum_{d=0}^{1}\Lambda((\alpha_i+\gamma_i d+\boldsymbol{\rho}'_a\mathbf{1}+\phi_i*d)(2h_{i,1}-1))*(p_a)^{1-d}(1-p_a)^{d}\right). \end{aligned} \tag{7}$$

Using the above recursive formulation, we can calculate $P_{\tau_i}(h_{i,\tau_i}|\boldsymbol{\theta}'_a)$.[7] Of course, this is a burdensome task if τ_i is large since computation will involve calculating the sum of the probabilities of all possible sequences of health outcomes that could have led to h_{i,τ_i}. Fortunately, the above recursive formulation simplifies matters greatly.

Our treatment of the initial condition in (7) imposes no additional parametric assumptions

[7] Heckman (1981a) proposes using this method which involves using the underlying statistical model to calculate $P_{\tau_i}(h_{i,\tau_i}|\boldsymbol{\theta}'_a)$ which can in turn be used to calculate $P(h_{i,T_i},...,h_{i,\tau_i})$. This procedure addresses the initial condition problem that occurs when the stochastic process has been running prior to τ_i. Since our underlying statistical model does not have any time varying regressors, we do not need to concern ourselves with the distribution of the time varying regressors for $t<\tau_i$. However, in the presence of time varying regressors, auxiliary distributional assumptions must be made. In addition, the computations become rather involved. An alternative to this is provided by Wooldridge (2005) who proposes modeling the distribution of the heterogeneity conditional on h_{i,τ_i} and any time varying regressors that may be present. Doing this does not require internal consistency with the underlying statistical model nor does it require computations that are as involved as the previous method, but it does require additional distributional assumptions. A third solution to the initial conditions problem assumes that the process has been running sufficiently long prior to the sampling period and that the process is in equilibrium. It then uses the stationary distribution for the process as the probability of the first observation. However, this will not work in our case as health is non-stationary process.

on the model beyond the assumptions that health evolves according to equation (1) and that the heterogeneity distribution is discrete. The reason is that because equation (1) contains no time varying regressors, we can roll the model back to $t = 0$. In addition, because the heterogeneity has a discrete distribution, we can treat the probability of wellness for heterogeneity type a at $t = 0$ (p_a) as an additional parameter in the model. In other words, because of these two assumptions, we do not need to make an additional parametric assumption about $P_0(h_{i,0} = 0|\boldsymbol{\theta}'_a)$.

Using these probabilities, taking logs and summing over individuals, we obtain the likelihood function:

$$L(\boldsymbol{\beta}) = \tag{8}$$
$$\sum_{i=1}^{N} \log \left(\sum_{a=1}^{A} \prod_{t=\tau_i+1}^{T_i} \Lambda(\boldsymbol{\theta}'_a \mathbf{Z}_{i,t-1}(2h_{i,t} - 1)) P_{\tau_i}(h_{i,\tau_i}|\boldsymbol{\theta}'_a)\pi_a \right).$$

where $\boldsymbol{\beta} \equiv (\boldsymbol{\theta}'_1, ..., \boldsymbol{\theta}'_A, \pi_1, ..., \pi_{A-1}, p_1, ..., p_A)$. $\boldsymbol{\beta}$ is of dimension $7A - 1$. $L(\boldsymbol{\beta})$ was maximized using the Fletcher-Powell algorithm, a variant of Newton's Method, which only requires the computation of the the gradient vector $\nabla L(\boldsymbol{\beta})$.[8] Since evaluating the likelihood in (8) can be time-consuming, we calculated analytical gradients, as opposed to numerical gradients.[9,10]

[8] Typically, when the distribution of unobserved heterogeneity is treated in the spirit of Heckman-Singer, the LM algorithm is used for optimization. However, Newton's method has also been used (see Baker and Melino 2000, for example).

[9] If we would have worked with the 5-point SRHS variable, the model would have looked something like

$$1(h_{i,t} = k) =$$
$$1(\varphi_{k-1} < \alpha_i + \sum_{j=2}^{5} \gamma_i^j 1(h_{i,t-1} = j) + \boldsymbol{\rho}'_i \mathbf{T} + \sum_{j=2}^{5} \phi_i^j t * 1(h_{i,t-1} = j) + \varepsilon_{i,t} \leq \varphi_k).$$

The number of parameters associated with this model is $11A + 4A + (A - 1) + 4 = 16A + 3$. $11A$ parameters are associated with the index. $4A$ parameters are associated with the initial condition probabilities. $(A - 1)$ parameters are associated with the probability weights associated with each point of support of the heterogeneity. Finally, the are 4 ancillary parameters in the model (the $\varphi'_k s$). For two points of support, this model has 35 parameters.

[10] All computer programs and data used are available upon request from the author.

5 Model Selection

In this section of the paper, we estimate our model while imposing various restrictions on $\boldsymbol{\theta}_i$ in order to investigate which versions of the model are favored by the data. We estimate the model using two points of support for the heterogeneity with the data described in Section 2. Each point of support of the heterogeneity corresponds to a different level of host resistance. Individuals with high levels of host resistance are called "healthy" and individuals with low levels are called "unhealthy." We subscript all parameters that correspond to healthy individuals with H and all parameters that correspond to unhealthy individuals with U. Thus, $\boldsymbol{\theta}_i \in \{\boldsymbol{\theta}_H, \boldsymbol{\theta}_U\}$. Note that the terms "healthy" and "unhealthy" refer to the individual's level of host resistance or overall robustness whereas the terms "ill" and "well" refer to the health state that the individual occupies. To control for gender in a non-parametric fashion, we estimate the model separately for men and women. The parameter estimates for men and women are displayed in Tables 1 and 2, respectively.

TABLES 1 AND 2 HERE

In column 1 of both tables, we impose the following restrictions:

$$\boldsymbol{\rho}_i' = [\rho^1, 0]' \forall i, \phi_i = 0 \forall i. \tag{L}$$

Assumption L imposes a simple homogeneous linear trend on the model, which corresponds to the log-odds of morbidity being linear in age. Testing restriction L is interesting because we know that the log-odds of mortality is linear. Looking at the estimates of γ_H and γ_U in column 1 of both tables, we see substantial evidence of state dependence for both healthy and unhealthy

people. Not surprisingly, in both tables, we see that the estimates of ρ^1 are positive and highly significant, indicating a high degree of non-stationarity. Finally, we see that our estimate of the probability of being a healthy individual (π_H) is 0.8227 for men and 0.7692 for women. This indicates that there is a higher proportion of persistently unhealthy women in our data than persistently unhealthy men, which is consistent with the body of research showing that women having higher morbidity (but lower mortality) than men.[11]

In column 2 of both tables, we allow for a homogeneous quadratic function of age and assume:

$$\boldsymbol{\rho}_i' = [\rho^1, \rho^2]' \forall i, \phi_i = 0 \forall i. \tag{Q}$$

Looking at Tables 1 and 2, three interesting findings emerge. First, the primary state dependence coefficients (γ_H and γ_U) do not change. Second, the probabilities of being healthy (π_H) in both tables do not change either. Third, in Table 1, we see that the log likelihood increases from -6155.8 in column 1 to -6154.1 which yields a likelihood ratio statistic of 3.4 with a corresponding p-value of 0.065. In Table 2, the likelihood ratio statistic is 4.6 with a corresponding p-value of 0.032. So, the data do appear to favor the quadratic model over the linear model, but only at the 90% level for men and the 95% level for women. Thus, relative to the quadratic model, the linear model still performs surprisingly well.

In column 3, we weaken our assumptions even further and allow for heterogeneous aging parameters and assume:

$$\phi_i = 0 \forall i. \tag{HA}$$

[11] For an excellent investigation into this issue, see Case and Paxson (2005).

Assumption HA allows for heterogeneous aging trends, but does not allow for an interaction between lagged health and age. As before, the estimates of the primary state dependence parameters (γ_H and γ_U) and the probability of being healthy (π_H) remain unchanged. Testing Assumption L against Assumption HA, we obtain p-values of 0.122 for men and 0.133 for women. When compared to the simple linear model, the model with heterogeneous aging performs worse than the quadratic model.

In column 4, we estimate the unrestricted model with heterogeneous aging parameters ($\boldsymbol{\rho}_i'$) and a heterogeneous interaction between lagged health and age ($\boldsymbol{\phi}_i$). We now see that the estimates of the primary state dependence parameters (γ_H and γ_U), are affected greatly. However, this does not necessarily imply that the resulting transition dynamics have changed. We investigate this issue further on in the paper and show that, in fact, the transition dynamics are relatively unaffected. Next, we see that the estimate of the probability of being healthy (π_H) remains unchanged for both men and women. Overall, we see that allowing for richer heterogeneity and state dependence does not affect the estimates of π_H. Finally, when we test the unrestricted model against the linear model, we obtain p-values of 0.136 for men and 0.168 for women. Relative to the linear model, the unrestricted model performs worse than both the quadratic model and the heterogeneous aging model.

Tables 1 and 2 also report the Akaike Selection Criterion (AIC) for each model. The AIC is

$$AIC = -\frac{2}{N}\log L(\widehat{\beta}) + \frac{2p}{N} \propto -\log L(\widehat{\beta}) + p$$

where $p = \dim(\widehat{\beta})$ (Amemiya 1985). The selection criterion is to choose the model with the smallest AIC. The AIC is reported for each model in the last row of Tables 1 and 2. As we would

expect from our perusal of the likelihood ratios, we find that the AIC criterion favors the quadratic model above the linear model and the linear model above the other models. Interestingly, the data favor the simpler linear and quadratic models above the more complicated (and more computationally intensive) models that allow for richer heterogeneity and state dependence. This bodes well for researchers who are concerned about computing time.

6 Heterogeneity

This section of the paper investigates the role that heterogeneity plays in determining health. To do this, we map our parameter estimates into transition probabilities and investigate the difference in these transitions across healthy and unhealthy people. Figures 1 and 2 display $P_t(h_{i,t} = 1|h_{i,t-1} = 1, \boldsymbol{\theta}'_a)$ for $a \in \{H, U\}$ and $t \in \{23, ..., 60\}$. Figure 1 corresponds to unhealthy men and Figure 2 corresponds to healthy men. We graph these profiles for the linear, quadratic, heterogeneous aging and unrestricted specifications. These figures elucidate the persistence of illness in the PSID for both heterogeneity types.

FIGURES 1 AND 2 HERE

Figure 1 shows that illness is highly persistent for unhealthy men. Even at age 25, illness has roughly a 50% chance of remaining at age 26. At age 60, the persistence of illness is greater than 85%. Interestingly, the estimated profiles do not vary across model specifications.

Figure 2 displays the persistence of illness for healthy men. This figure shows much lower persistence than Figure 1. At age 25, illness has a 5% chance of persisting until the next year.

By age 60, this probability is somewhere between 0.3 and 0.4. These transition profiles differ from one-another substantially more than those in Figure 1. However, this is exactly what we would expect to see since illness tends to be less prevalent among healthier men and, thus, the data should not contain an abundance of information on $P_t(h_{i,t} = 1|h_{i,t-1} = 1, \boldsymbol{\theta}'_H)$. Accordingly, we would expect greater variation across models in the estimates of this transition probability.

FIGURES 3 AND 4 HERE

Figures 3 and 4 display the probability of moving from $h_{i,t-1} = 0$ to $h_{i,t} = 1$ for healthy and unhealthy men. These figures depict the probability of the onset of illness or equivalently one minus the persistence of wellness. Figure 3 shows a very high probability of the onset of illness among unhealthy men. In contrast, Figure 4 shows a very low probability of illness manifesting itself among healthy men. Even at age 60, this probability is less than 0.05. So, while we see that illness is a highly persistent state for unhealthy men, we see that wellness is a highly persistent state for healthy men. Finally, in both figures, all four models yield similar transition profiles.

FIGURES 5 AND 6 HERE

Figures 5 displays $P_t(h_{i,t} = 1|h_{i,t-1} = 1, \boldsymbol{\theta}'_a)$ for $a \in \{H, U\}$ and Figure 6 displays $P_t(h_{i,t} = 1|h_{i,t-1} = 0, \boldsymbol{\theta}'_a)$ for $a \in \{H, U\}$. These figures gives us a notion of the degree of heterogeneity that is present in the data. Both figures show large variation across healthy and unhealthy

people suggesting that there is substantial heterogeneity present in the data.[12]

FIGURES 7 AND 8 HERE

Finally, in Figures 7 and 8, we compare the transition dynamics of men to those of women. Figure 7 displays $P_t(h_{i,t} = 1|h_{i,t-1} = 1, \boldsymbol{\theta}'_a)$ for $a \in \{H, U\}$ for both men and women and Figure 8 displays $P_t(h_{i,t} = 1|h_{i,t-1} = 0, \boldsymbol{\theta}'_a)$ for $a \in \{H, U\}$ for men and women. Each figure displays a total of 16 profiles. Interestingly, these figures show very similar profiles for men and women suggesting that the primary differences in morbidity rates across sexes are being driven by a relatively higher proportion of of unhealthy women (*i.e.* a higher π_H) rather than substantial differences in transition profiles.

7 State Dependence

This section of the paper is concerned with quantifying the degree of state dependence in the health process. This exercise will allow us to assess the dynamic consequences of hypothetical medical interventions which improve a person's health at a point in time. The reason for this is that if the degree of state dependence is large, then medical interventions which move people from illness to wellnes will have large dynamic benefits or multiplier effects.

[12] In this paper, we do not provide a formal test of the null that the data contain no heterogeneity. A formal test for the presence of any unobserved heterogeneity in the data would test the null hypothesis that $\pi_H = 0$. However, deriving the distribution theory of this statistic is a non-trivial task since, under the null, there are unidentified nuisance parameters.

To do this, we calculate

$$SD(t;\boldsymbol{\theta}_i) \equiv P_t(h_{i,t}=1|h_{i,t-1}=1,\boldsymbol{\theta}_i') - P_t(h_{i,t}=1|h_{i,t-1}=0,\boldsymbol{\theta}_i'). \tag{9}$$

$SD(t;\boldsymbol{\theta}_i)$ gives us the reduction in the probability of illness at t resulting from a medical intervention that changes the health state of individual i from illness to wellness at $t-1$. Knowledge of these functions allows us to better understand at what ages and for whom medical interventions will be most potent.

FIGURE 9 HERE

Figure 9 displays $SD(t;\boldsymbol{\theta}_a)$ for $t \in \{23,...,60\}$ and $a \in \{H,U\}$ for men. Prior to age 60, we see that there is more state dependence among the unhealthy than among the healthy. The implication then is that interventions which target people below 60 years of age will have greater dynamic effects on the less robust individuals in society than on society's more robust individuals. We see that state dependence among unhealthy men is the greatest between ages 40 and 45. In contrast, state dependence for healthy men increases until age 60 suggesting that it does not peak until sometime thereafter. So, if policy aims to improve the health of society's most disadvantaged in terms of health (i.e. those for whom $\boldsymbol{\theta}_i = \boldsymbol{\theta}_U$) then the case for medical interventions targeting people in middle age or younger is strengthened as this is where the degree of state dependence among the unhealthy is greatest.[13] Figure 9 shows a general agreement among the four models in the estimates of $SD(t;\boldsymbol{\theta}_U)$. There is also a general

[13] In other words, an intervention which targets all middle-aged people will have a greater impact on unhealthy individuals than on healthy individuals.

consensus among models in the estimates of $SD(t;\boldsymbol{\theta}_H)$; although, the unrestricted model with time-variant state dependence parameters is somewhat of an outlier.

Figure 10 displays $SD(t;\boldsymbol{\theta}_a)$ for $t \in \{23, ..., 85\}$ and $a \in \{H, U\}$ for men. Because we estimated our models using a sample of people between the ages of 23 and 60, Figure 10 shows out-of-sample predictions. The figure shows that the degree of state dependence among the unhealthy declines in old age. In contrast, we see a precipitous rise in state dependence among the healthy beyond age 60. Among the healthy, we see greater dispersion among the four models past age 60 than before it. We also see that, for the unrestricted model, the heterogeneous aging model (Assumption HA) and the quadratic model (Assumption Q), that state dependence is greatest around 80 years of age. For the linear model, state dependence is increasing until 85 years of age. These out-of-sample predictions suggest that, among the elderly, there is considerably more state dependence among the healthy than the unhealthy so that medical interventions which target the elderly will disproportionately affect the most robust people in society.

Figures 11 and 12 display $SD(t;\boldsymbol{\theta}_a)$ for $a \in \{H, U\}$ for women. Figure 11 displays until age 60 and Figure 12 displays until age 85. As we see for unhealthy men, we also see that state dependence for unhealthy women is the greatest between ages 40 and 45. However, unhealthy women exhibit slightly less state dependence than unhealthy men, overall. In addition, we see that the highest degree of state dependence for healthy women does not occur until past age 60 just as it does for healthy men. In contrast to healthy men, however, state dependence for healthy women peaks between ages 70 and 75 and is lower overall than it is for men.

8 Conclusions

The paper investigated the dynamics of health status in the PSID. To do this, we estimated several specifications of a flexible model describing the evolution of health over the life-course which allowed for two sources of persistence: unobserved heterogeneity and state dependence. The data favored the models with the simpler forms of heterogeneity and state dependence above the more complicated models. In addition, we found that both unobserved heterogeneity and state dependence play important roles in the determination of health. However, the magnitude of state dependence depended critically on the individual's age and unobserved characteristics.

The results of this paper shed light on how health policy should be conducted and, thus, have implications for the gradient: the much-studied but little-understood statistical correlation between health and socioeconomic status (Adler, *et al* 1994). If it is the case that the gradient is largely determined by the causal impact of health status on earnings and wealth - as suggested by Smith (1999) - then the relevant policy prescription is to directly target health *via* medical interventions (Deaton 2002). The argument for medical interventions is further strengthened if health exhibits a high degree of state dependence as this implies that the intervention will have large dynamic effects which operate through the causal effect of health on itself.

Our results indicate that, among the "health-deprived" or those who are innately less-able to cope with health shocks, there is a large degree of state dependence through middle age, but its magnitude dissipates greatly in old age. However, among the "health-endowed" or those who are innately better-able to cope with health shocks, we find the opposite so that there is very little state dependence early in life, but a large degree of it in old age. The implication is

then that interventions which target people earlier on in the life-course will have benefits that disproportionately accrue to the disadvantaged. In contrast, interventions that target the elderly will disproportionately benefit those who are advantaged in terms of their heterogeneity. Now, if it is the case that being health-endowed is correlated with socioeconomic status as is suggested by Halliday (2004), then interventions which target the elderly would likely improve the health of the rich while leaving the health of the poor relatively unaffected, thereby, exacerbating the gradient. This is not to argue that health policies which target the elderly are undesirable Indeed, any policy that improves health - whether it is the health of the advantaged or disadvantaged - is good. We merely want to point out that health policies that target people at different points in the life-cycle are likely to have different effects on the gradient.

This paper suggests several further research topics. First, additional work should be done to incorporate mortality into the existing framework, while addressing the selection bias that it induces. As discussed by Wooldridge (2000), correcting selection bias due to non-random attrition - which encompasses mortality - is difficult, particularly, when the econometrician is unwilling to make stringent assumptions. The reason is that while the sample may be random in the initial time period, the sample is, generally, non-random for every subsequent period. Consequently, we expect any estimates of a transition probability in these subsequent periods to suffer from selection bias even if mortality were to be added as a third (absorbing) state.

Second, further work should estimate models with higher orders of state dependence. The main challenge of working with models with higher orders of state dependence concerns treating the initial condition. In the case of first order state dependence calculation of the probability of the first observation involves the summation of 2^{τ_i} probabilities. In the case of second order

state dependence, we would have to sum 4^{τ_i-1} probabilities. Generally, for Kth order state dependence, we have to sum over $\left(2^K\right)^{\tau_i-K}$ sequences. Thus, allowing for higher orders of state dependence is likely to increase the computational burden by a fairly large margin.

Finally, additional work should estimate structural models that incorporate health as a state variable. While there have been some studies that have done so such as Rust and Phelan (1997) and Arcidiacono, Heig and Sloan (2004), this is still a relatively new field. Particular attention should be paid to how assumptions about the health process affect estimation results. Typically, in structural models, assumptions are made for the sake of computational tractability. Such an exercise would allow us to see how innocuous these assumptions are and, thus, would shed light on the issue of identification in structural estimation.

References

[1] Abowd, J. and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57, 411-445.

[2] Adler, N.E., T. Boyce, M.A. Chesney, S. Cohen, S. Folkman, R.L. Kahn and S.L. Syme (1994): "Socioeconomic Status and Health: The Challenge of the Gradient," *American Psychologist*, 49, 15-24.

[3] Amemiya, T. (1985): *Advanced Econometrics*. Cambridge, MA: Harvard University Press.

[4] Arcidiacono, P., H. Sieg and F. Sloan (2004): "Living Rationally Under the Volcano? Heavy Drinking and Smoking Among the Elderly," unpublished manuscript, Duke University.

[5] Baker, M. and A. Melino (2000): "Duration Dependence and Non-parametric Heterogeneity: A Monte Carlo Study," *Journal of Econometrics*, 96, 357-390.

[6] Barker, D.J.P. (1997): "Maternal Nutrition, Fetal Nutrition and Diseases in Later Life," *Nutrition*, 13, 807-813.

[7] Barker, D.J.P., C. Osmond, J. Golding, D. Kuh and M.E.J. Wadsworth (1989): "Growth in Utero, Blood Pressure in Childhood and Adul Life, and Mortality from Cardiovascular Disease," *British Medical Journal*, 298, 564-567.

[8] Caballero, R.J. (1990): "Consumption Puzzles and Precautionary Savings," *Journal of Monetary Economics*, 25, 113-136.

[9] Case, A., D. Lubotsky and C. Paxson (2002): "Economic Status and Health in Childhood: The Origins of the Gradient," *American Economic Review*, 92, 1308-1334.

[10] Case A. and C. Paxson (2005): "Sex Differences in Morbidity and Mortality" forthcoming *Demography*.

[11] Chay, K.Y., H. Hoynes and D. Hyslop (2001): "A Non-experimental Analysis of True State Dependence in Monthly Welfare Participation Sequences," unpublished manuscript, UC-Berkeley.

[12] Contoyannis, P., A.M. Jones and R. Leon-Gonzalez (2004): "Using Simulation-based Inference with Panel Data in Health Economics," *Health Economics*, 13, 101-122.

[13] Contoyannis, P., A.M. Jones, and N. Rice (2004): "Simulation-based Inference in Dynamic Panel Probit Models: An Application to Health," *Empirical Economics*, 29, 49-77.

[14] Deaton, A. (1992): *Understanding Consumption*. Oxford: Oxford University Press.

[15] Deaton, A. (2002): "Policy Implications of the Gradient of Health and Wealth," *Health Affairs,* 21, 13-30.

[16] Halliday, T. (2004): *Essays on the Economics of Health and Migration*, unpublished Ph.D. dissertation, Princeton University.

[17] Heckman, J.J. (1981a): "Statistical Models for Discrete Panel Data," in *Structural Analysis of Discrete Data*, ed. by Charles Manski and Daniel McFadden. Cambridge, MA: MIT Press.

[18] Heckman, J.J. (1981b): "Heterogeneity and State Dependence," in *Structural Analysis of Discrete Data*, ed. by Charles Manski and Daniel McFadden. Cambridge, MA: MIT Press.

[19] Heckman, J.J. and B. Singer (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52, 271-320.

[20] Hyslop, D. (1999): "State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women," *Econometrica*, 67, 1255-1294.

[21] Idler, E.L. and S.V. Kasl (1995): "Self-Ratings of Health: Do They Also Predict Changes in Functional Ability?" *Journal of Gerentology*, 50, S344-S353.

[22] Kaplan, G.A. and T. Camacho (1983): "Perceived Health and Mortality: A 9 Year Follow-up of the Human Population Laboratory Cohort," *American Journal of Epidemiology*, 177, 292.

[23] Lillard, E.L., and R. Willis (1978): "Dynamic Aspects of Earnings Mobility," *Econometrica*, 46, 985-1012.

[24] Lindsay, B. (1983): "The Geometry of Mixture Likelihoods, Part I" *Annals of Statistics*, 11, 86-94.

[25] Magnac, T. (2000): "Subsidised Training and Youth Employment: Distinguishing Unobserved Heterogeneity from State Dependence in Labour Market Histories," *Economic Journal*, 110, 805-837.

[26] Meghir, C. and L. Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72, 1-32.

[27] Mossey, J.M. and E. Shapiro (1982): "Self-Rated Health: A Predictor of Mortality Among the Elderly," *American Journal of Public Health*, 71, 100.

[28] Neyman, J. and E. Scott (1948): "Consistent Estimates Based on Partially Consistent Observations," *Econometrica*, 16, 1-32.

[29] Rust, J. and C. Phelan (1997): "How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets," *Econometrica*, 65, 781-831.

[30] Smith, J. (1999): "Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status," *Journal of Economic Perspectives*, 13, 145-166.

[31] Smith, J. (2004): "Health and SES Over the Life-Course," unpublished manuscript, RAND.

[32] Wooldridge, J. (2000): *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

[33] Wooldrdige, J. (2005): "Simple Solutions to the Initial Conditions Problem for Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity," forthcoming *Journal of Applied Econometrics*.

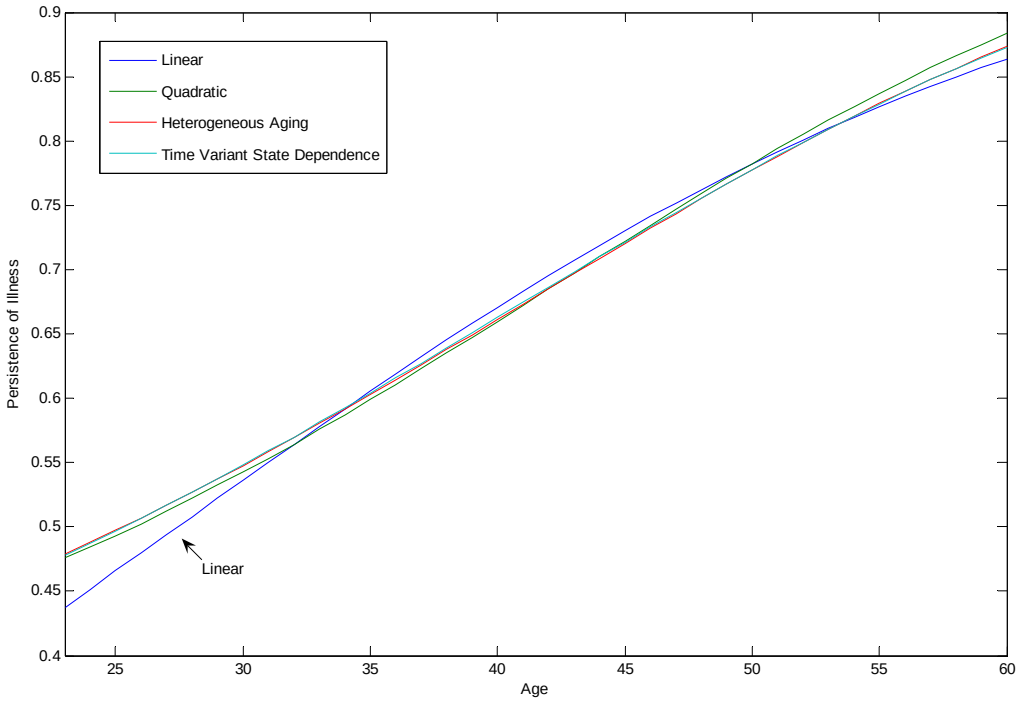

Figure 1: Persistence of Illness, Unhealthy Men

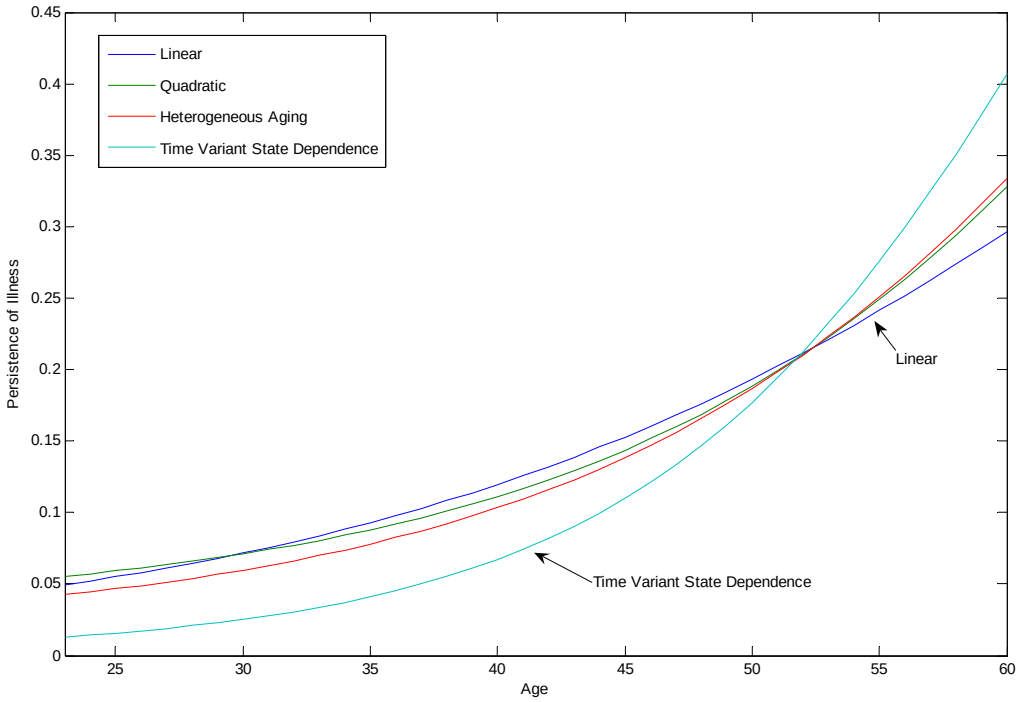

Figure 2: Persistence of Illness, Healthy Men

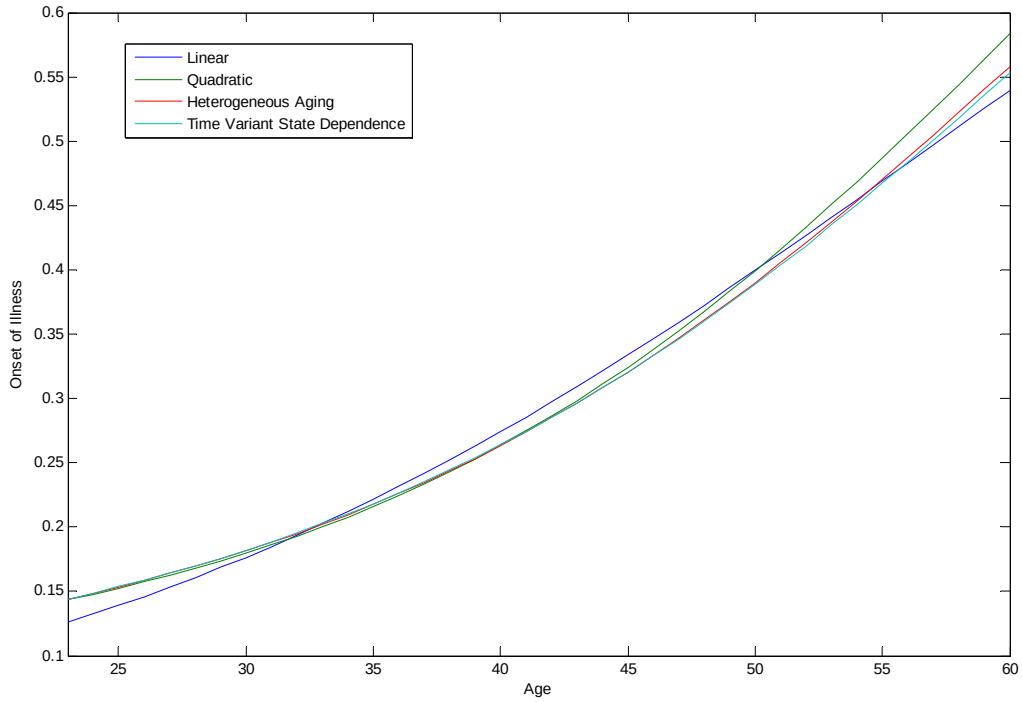

Figure 3: Onset of Illness, Unhealthy Men

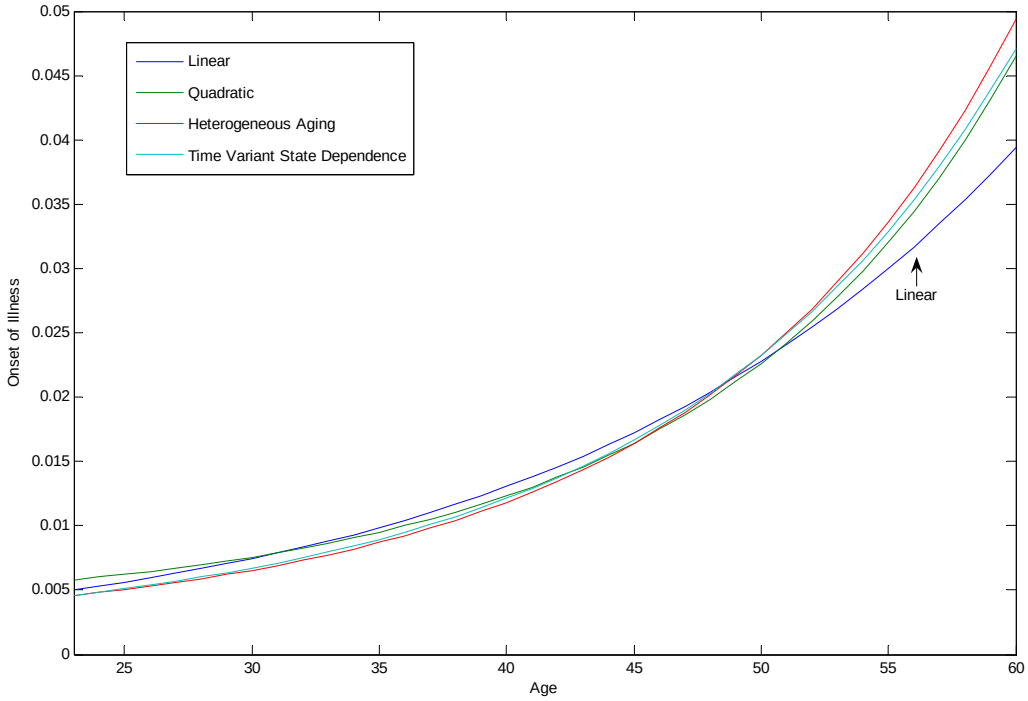

Figure 4: Onset of Illness, Healthy Men

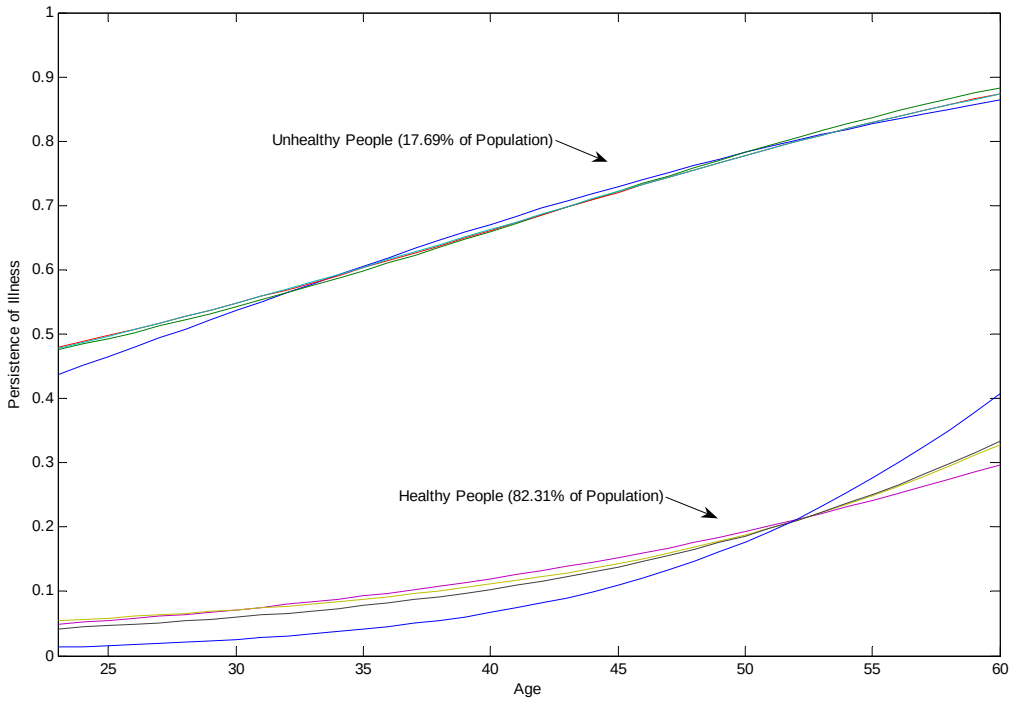

Figure 5: Persistence of Illness, Men

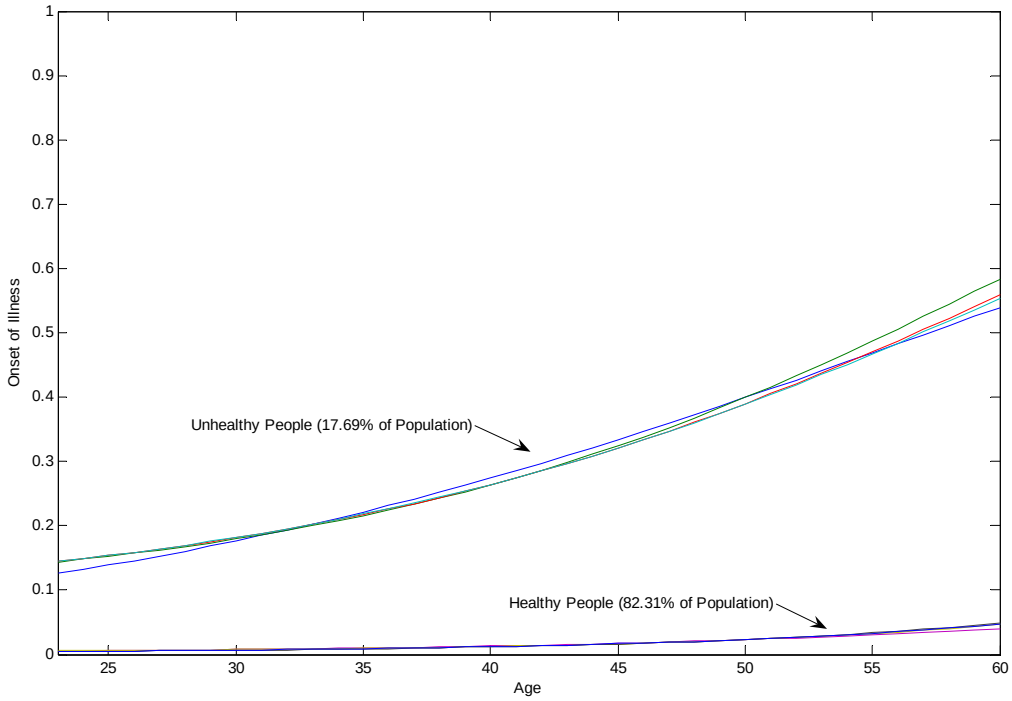

Figure 6: Onset of Illness, Men

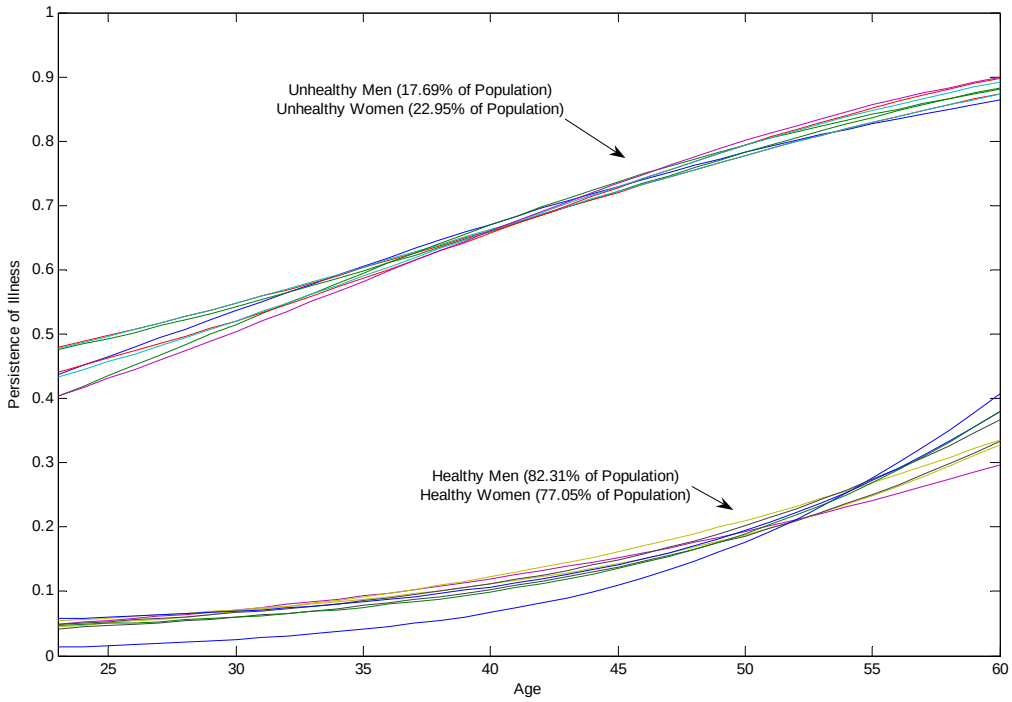

Figure 7: Persistence of Illness, Men and Women

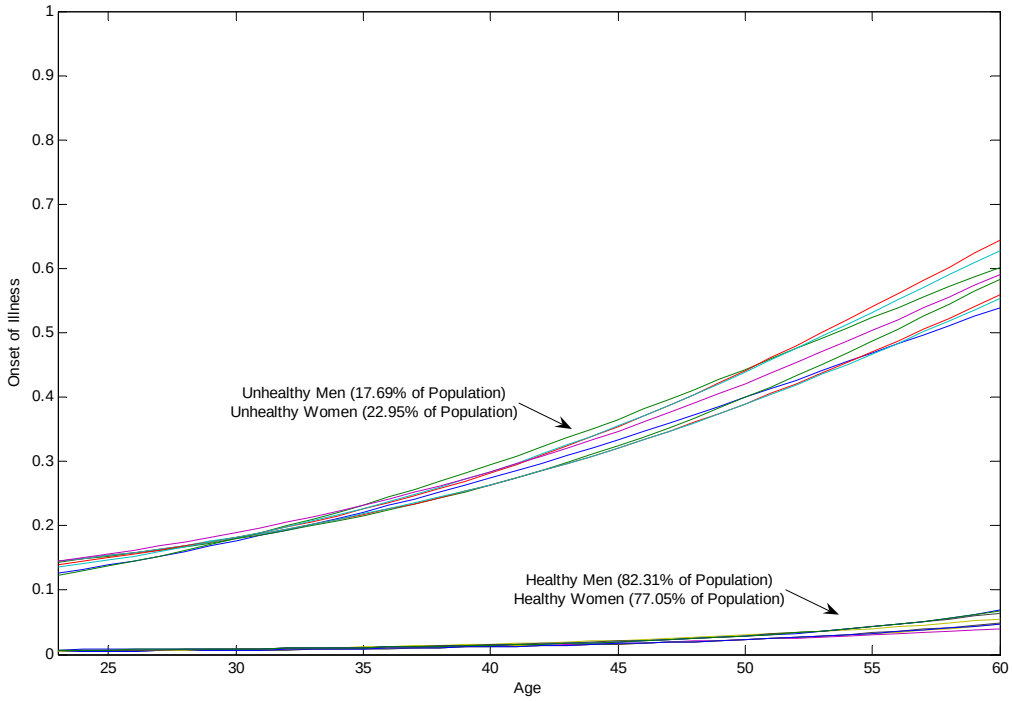

Figure 8: Onset of Illness, Men and Women

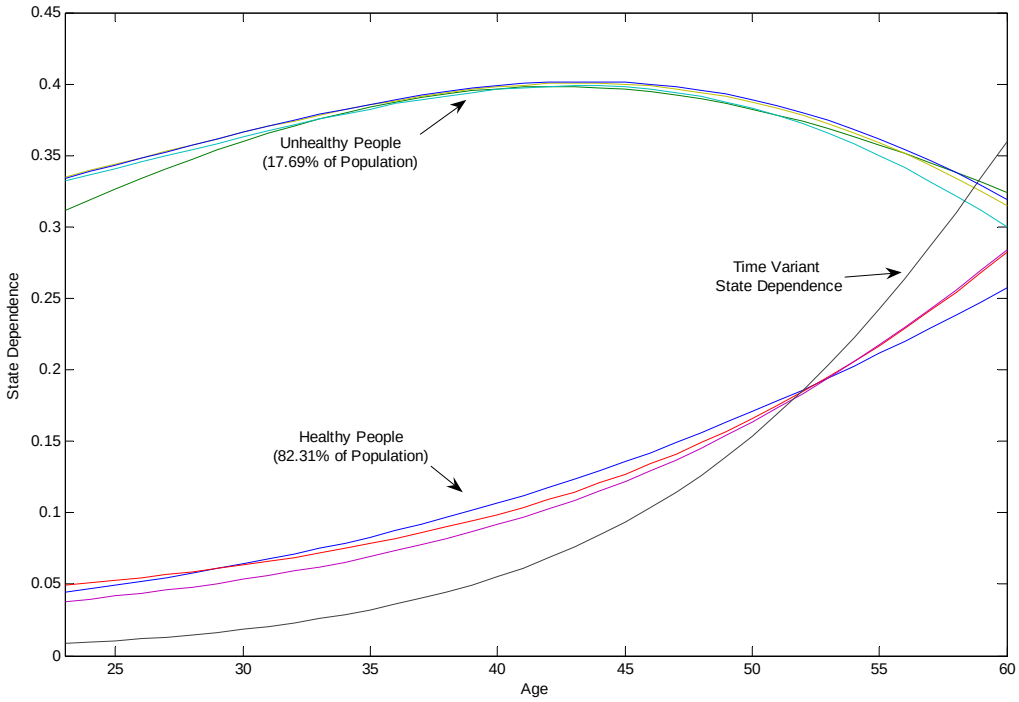

Figure 9: State Dependence, Men

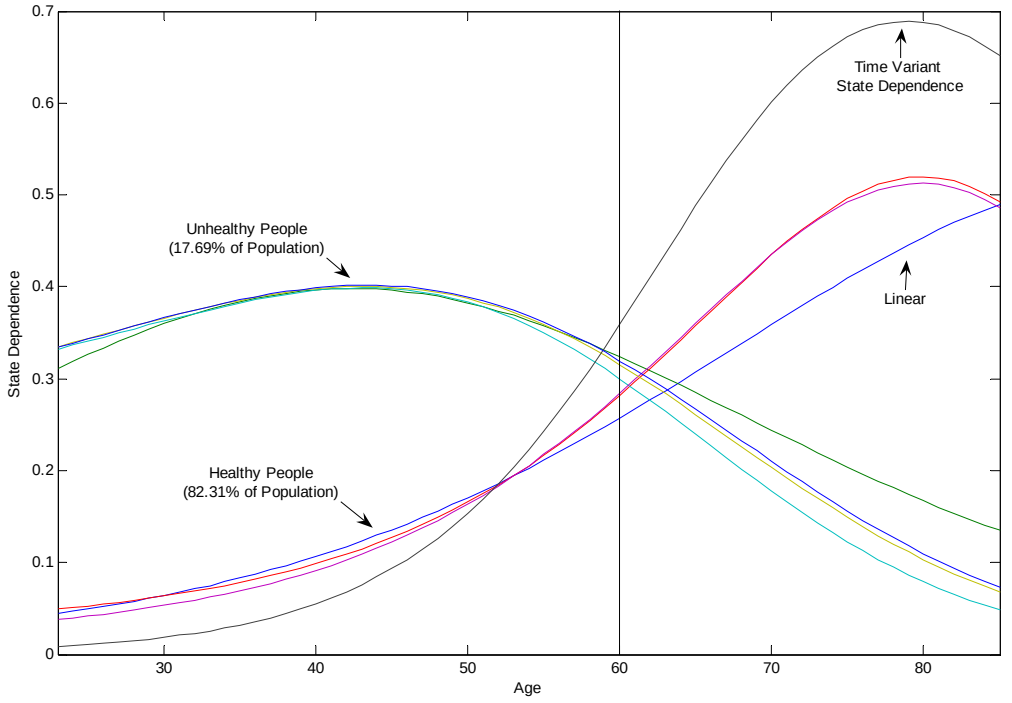

Figure 10, State Dependence, Men, Out-of-Sample

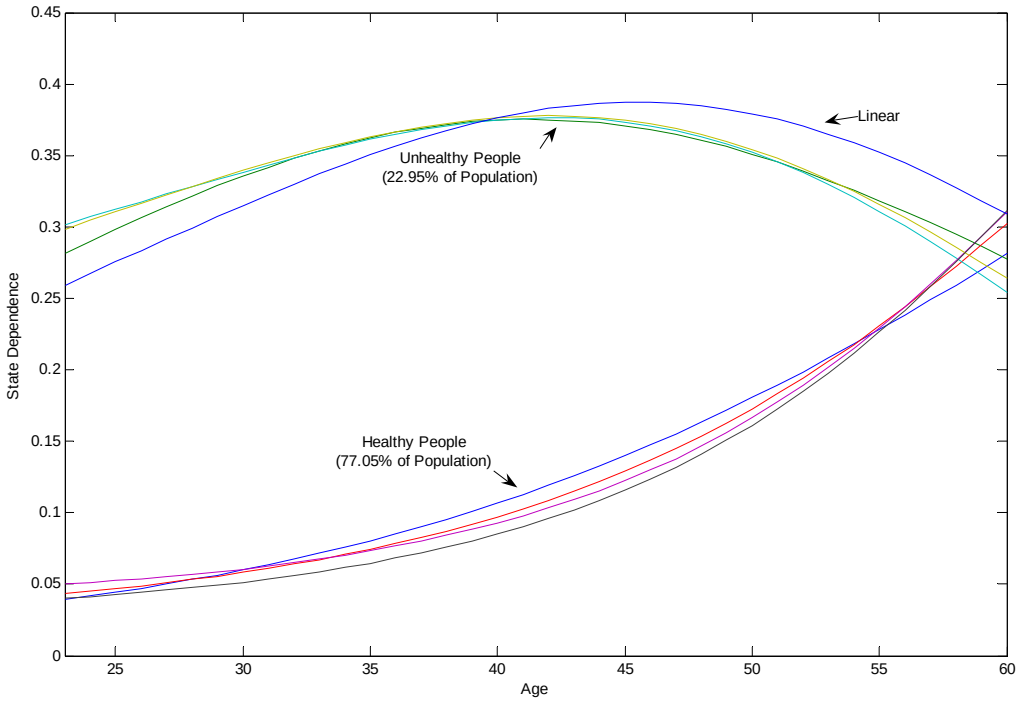

Figure 11: State Dependence, Women

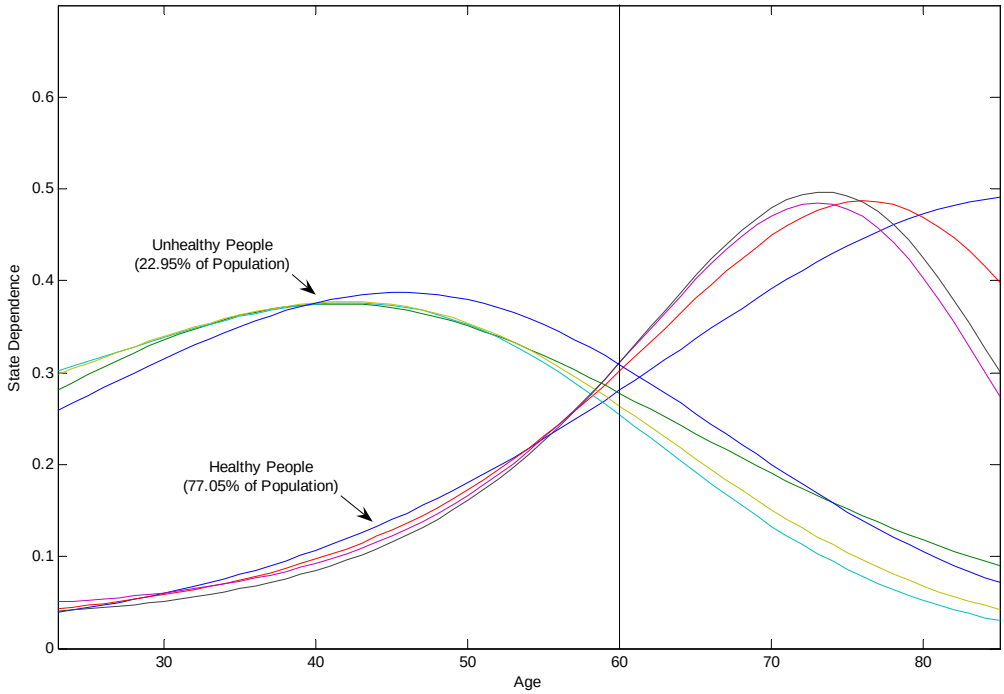

Figure 12, State Dependence, Women, Out-of-Sample

Table 1: Descriptive Statistics

Women

	Mean	**25% Quantile**	**75% Quantile**	**Standard Deviation**
SRHS (5-Point)	2.22	1	3	0.99
SRHS (2-Point)	0.10	0	0	0.30
Age	39.10	31	46	9.82
Panel Duration*	8.21	4	14	4.45
$N = 4186$**				

Men

SRHS (5-Point)	2.10	1	3	0.98
SRHS (2-Point)	0.08	0	0	0.27
Age	39.34	32	46	9.56
Panel Duration*	8.44	4	14	4.46
$N = 3923$**				

*Panel duration refers to the length of time that the individual was in the panel.

**N is the number of individual observations, not individual-time observations.

Table 2: Parameter Estimates - Men

Parameter	**Variable**	**(1)**	**(2)**	**(3)**	**(4)**
α_H	$Cons$	-6.5898	-5.5831	-6.2433	-6.3772
		(0.2107)	(0.6644	(1.2888)	(1.3299)
α_U		-3.3280	-2.2256	-2.3281	-2.3608
		(0.2034)	(0.6654)	(0.7636)	(0.7853)
γ_H	$h_{i,t-1}$	2.3309	2.3056	2.2657	0.0586
		(0.2711)	(0.2602)	(0.2662)	(1.4489)
γ_U		1.6855	1.6885	1.6972	1.6811
		(0.1210)	(0.1194)	(0.1163)	(0.5061)
ρ_H^1	$age/10$	0.5661	0.0449	0.2674	0.3583
		(0.0512)	(0.3309)	(0.6056)	(0.6294)
ρ_U^1				0.1209	0.1426
				(0.3810)	(0.3823)
ρ_H^2	$(age/100)^2$	$-$	6.3697	4.6724	3.3898
			(4.0547)	(6.8829)	(7.1765)
ρ_U^2				5.1034	4.7766
				(4.6566)	(4.6534)
ϕ_H	$(age/10) * h_{i,t-1}$	$-$	$-$	$-$	0.4285
					(0.2716)
ϕ_U					0.0050
					(0.1193)
π_H		0.8227	0.8246	0.8224	0.8226
		(0.0031)	(0.0030)	(0.0042)	(0.0042)
p_H		0.9828	0.9812	0.9807	0.9796
		(0.0002)	(0.0002)	(0.0002)	(0.0042)
p_U		0.8847	0.8919	0.8958	0.9000
		(0.0078)	(0.0070)	(0.0079)	(0.0073)
$L(\widehat{\beta})$		-6155.8	-6154.1	-6152.9	-6151.6
Likelihood Ratio*		$-$	3.4	5.8	8.4
			(0.065)	(0.122)	(0.136)
AIC**		6163.8	6163.1	6163.9	6164.6

*The likelihood ratio statistic comparing columns (2), (3) and (4) to column (1).
**$AIC = -L(\widehat{\beta}) + \dim(\widehat{\beta}) \propto \frac{-2}{N} L(\widehat{\beta}) + \frac{2}{N} \dim(\widehat{\beta})$
+Standard Errors are given in parentheses.

Table 3: Parameter Estimates - Women

Parameter	**Variable**	**(1)**	**(2)**	**(3)**	**(4)**
α_H	$Cons$	-6.7052	-5.6806	-4.7666	-4.9900
		(0.2049)	(0.5511)	(1.1776)	(1.2609)
α_U		-3.4484	-2.4152	-2.6885	-2.5774
		(0.1864)	(0.5475)	(0.6343)	(0.6437)
γ_H	$h_{i,t-1}$	2.1615	2.1273	2.1182	1.8592
		(0.2434)	(0.2442)	(0.2500)	(1.2819)
γ_U		1.5781	1.5834	1.5894	1.1173
		(0.0958)	(0.0953)	(0.0956)	(0.4257)
ρ_H^1	$age/10$			-0.3237	-0.2295
		0.6440	0.1090	(0.5418)	(0.5677)
ρ_U^1		(0.0443)	(0.2748)	0.2533	0.2585
				(0.3271)	(0.3245)
ρ_H^2	$(age/100)^2$			11.3870	10.4251
		$-$	6.5370	(6.0463)	(6.2427)
ρ_U^2			(3.3753)	4.7052	3.8708
				(4.0856)	(3.9990)
ϕ_H	$(age/10) * h_{i,t-1}$	$-$	$-$	$-$	0.0440
					(0.2658)
ϕ_U					0.1180
					(0.1051)
π_H		0.7692	0.7709	0.7708	0.7710
		(0.0073)	(0.0072)	(0.0047)	(0.0072)
p_H		0.9781	0.9767	0.9767	0.9776
		(0.0002)	(0.0003)	(0.0002)	(0.0002)
p_U		0.8537	0.8580	0.8575	0.8537
		(0.0103)	(0.0099)	(0.0076)	(0.0106)
$L(\widehat{\beta})$		-7556.5	-7554.2	-7553.7	-7552.6
Likelihood Ratio*		-	4.6	5.6	7.8
			(0.032)	(0.133)	(0.168)
AIC**		7564.5	7563.2	7564.7	7565.6

*The likelihood ratio statistic comparing columns (2), (3) and (4) to column (1).
**$AIC = -L(\widehat{\beta}) + \dim(\widehat{\beta}) \propto \frac{-2}{N} L(\widehat{\beta}) + \frac{2}{N} \dim(\widehat{\beta})$
+Standard Errors are given in parentheses.