DISCUSSION PAPER SERIES

IZA DP No. 3463

Heterogeneity, State Dependence and Health

Timothy J. Halliday

April 2008

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Heterogeneity, State Dependence and Health

Timothy J. Halliday
University of Hawaii at Manoa
and IZA

Discussion Paper No. 3463
April 2008

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3463
April 2008

ABSTRACT

Heterogeneity, State Dependence and Health*

We investigate the evolution of health over the life-cycle. We allow for two sources of persistence: unobserved heterogeneity and state dependence. Estimation indicates that there is a large degree of heterogeneity. For half the population, there are modest degrees of state dependence. For the other half of the population, the degree of state dependence is near unity. However, this may be the result of a high frequency of people in our data who never exit healthy states, potentially resulting in a failure to pin down the state dependence parameter for this segment of the population. We conclude that individual characteristics that trace back to early adulthood and before can have far reaching effects on health.

JEL Classification: I1, C5

Keywords: health, dynamic panel data models, gradient

Corresponding author:

Timothy J. Halliday
Department of Economics
University of Hawai'i at Manoa
2424 Maile Way
Saunders Hall 533
Honolulu, HI 96822
USA
E-mail: halliday@hawaii.edu

* This paper was the first chapter from my dissertation at Princeton University. I would like to thank the editor of this journal and several anonymous referees for excellent comments. In addition, I would like to extend my gratitude to my advisors Chris Paxson and Bo Honoré for their encouragement and guidance.

1 Introduction

We explore the dynamics of health and, in doing so, concern ourselves with two tasks. First, we aim to gain a better understanding of how to model the evolution of health. While many empirical studies 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] 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 its dynamics.[2] Second, we quantify the relative contributions of unobserved heterogeneity and state dependence in the determination of health. Doing so is important as this will have implications for health policy.

Utilizing data on Self-Reported Health Status (SRHS) from the Panel Study of Income Dynamics (PSID), we observe that health is highly persistent. The first order auto-correlation of a dummy variable indicating bad health is 0.5661 and 0.5643 for men and women, respectively. 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 which allows for two sources of persistence. The first is unobserved heterogeneity or the (unobserved) ability to cope with health shocks. The second is state dependence or the

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

[2] For examples of structural models using health as a state variable, see Rust and Phelan (1997), French (2005) and Arcidiacono, Heig and Sloan (2007).

degree to which the ability to cope with a shock depends on health status. Estimation will shed light on the relative contributions of both of these sources of persistence.

The balance of this paper is organized as follows. Section 2 describes the data. In Section 3, we set up our model. In Section 4, we describe our estimation procedure. Section 5 discusses our findings. Finally, in Section 6, we conclude and discuss the relevance of our findings for health policy.

2 Data

We use data from the PSID spanning the years 1984 to 1997. The variables that we employ are SRHS, age and gender. The SRHS question was only asked of heads of household and their spouses and, thus, our sample is restricted to these individuals. We do not employ data prior to 1984 since the SRHS question was not asked in these years. 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.

SRHS is a five-point categorical variable that measures the respondent's assessment of their own health. One is excellent and five is poor. While these data are subjective measures, there is an extensive literature that has shown a strong link between SRHS and more objective health outcomes such as mortality and the prevalence of disease (Mossey and Shapiro 1982; Kaplan and Camacho 1983; Idler and Kasl 1995; Smith 2003).[3] To lower the number of parameters that we

[3] Many objective health measures are not without their limitations. For example, self-reports of specific morbidities such as diabetes or cancer are often inaccurate since many people are unaware that they even have these conditions due to low consumption of medical services. In addition, these measures typically do not account

estimate, we map reports of fair or poor health into unity and all others into zero.

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 or 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.

3 The Empirical Model

We let $h_{i,t} \in \{0, 1\}$ denote the health of individual i at age t. When $h_{i,t} = 1$ then the individual is "ill" and when $h_{i,t} = 0$ she is "well." Health evolves according to the following process:

$$h_{i,t} = 1(\alpha_i + \gamma_i h_{i,t-1} + \boldsymbol{\rho}_i \mathbf{T} + \varepsilon_{i,t} \geq 0), \tag{1}$$

where $\mathbf{T} = [t, t^2]'$.[4] The residual in the model represents idiosyncratic risk or "health shocks" such as accident occurrence, disease onset or exposure to bacteria. We assume that $\varepsilon_{i,t}$ is independent of $(\alpha_i, \gamma_i, \boldsymbol{\rho}_i', h_{i,0})$ and that it is distributed *i.i.d.* across time with a logistic distribution.

for the severity of the condition.

[4] While we acknowledge that a thorough understanding of the linkages between income and health is of vital importance to policy makers, we do not incorporate income into the analysis as doing so would involve much more than simply including income as a strictly exogenous explanatory variable in equation (1). To include income in the analysis, we would have to model income as a predetermined or endogenous variable. This would have made the exercise substantially more complicated.

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

The model has three other key aspects. First, $\boldsymbol{\rho}_i'$ models aging and, thus, allows the effects of health shocks to increase with age. Within the context of the Grossman model of health investment (Grossman 1972), these coefficients can be interpreted as the rate at which the health capital stock depreciates. Second, γ_i models state dependence or the notion that the ability to cope with a given shock will depend on health status. 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.[5] Third, the model allows for a large degree of heterogeneity by allowing all of the elements of $\boldsymbol{\theta}_i$ to vary across individuals. Unobserved heterogeneity models an individual's ability to resist health shocks. Finally, it is important to point out that, while this discussion provides a motivation for our model that is rooted in epidemiology, there are economic motivations which we describe below.

3.1 A Reduced Form Model of Health Investment

Our model can be viewed as a reduced form model of health investment. Suppose that agents live until age T with certainty and derive utility from a consumption good denoted by $c_{i,t}$. Utility

[5] It is important to contrast our model with an obvious alternative formulation in which health is determined by a continuous index given by $H_{i,t}$ which follows an AR(1) process and agents report ill health when $H_{i,t}$ is beyond some threshold i.e. $h_{i,t} = 1(H_{i,t} \geq 0)$. While this alternative model does allow health shocks to have persistent effects, it does not allow for state dependence. In other words, in this model, the effects of a shock on future health outcomes are not conditioned by the agent's current health status.

in a given period depends on the health state *a la* Viscusi and Evans (1990) and is denoted by $u(c_{i,t}, h_{i,t})$. The agent's expected lifetime utility is then $E_0\left(\sum_{t=0}^{T}\beta^t u(c_{i,t}, h_{i,t})\right)$. The health state is (partly) the consequence of an endogenous investment decision, $i_{i,t} \in \{0, 1\}$:

$$h_{i,t} = 1\left(g_{i,t}(i_{i,t-1}) + \varepsilon_{i,t} \geq 0\right), \tag{3}$$

where $g_{i,t}(i_{i,t})$ is an individual-specific return to health investment with the property that $g_{i,t}(0) > g_{i,t}(1)$. Income is given by $y_{i,t}$ and investment imposes pecuniary costs of the form $\lambda_{i,1} * h_{i,t} + \lambda_{i,0} * (1 - h_{i,t})$ with $\lambda_{i,1} > \lambda_{i,0}$. Assuming no storage, the individual's budget constraint will be given by

$$c_{i,t} + i_{i,t} * [\lambda_{i,1} * h_{i,t} + \lambda_{i,0} * (1 - h_{i,t})] \leq y_{i,t}. \tag{4}$$

In this simple set-up, health will be a dynamic process similar to equation (1) because investment in equation (3) will depend on health status in the previous period due to state-dependent utility and investment costs. Consequently, a positive degree of state dependence might indicate that health investment is less likely when people are ill.

3.2 An Exogenous State Variable

Our model can be viewed as an exogenous state variable in a life-cycle consumption model. Many recent investigations into life-cycle consumer behavior such as Arcidiacono, Sieg and Sloan (2007), French (2005) and Rust and Phelan (1997) have incorporated exogenous uncertainty over health states. Our investigation will provide additional insights into how this uncertainty should be modeled. Proper modeling is crucial for the conclusions of these models to be valid. Indeed,

Deaton (1992) provides a discussion of how different income processes can lead to radically different consumption behaviors and, thus, demonstrates the sensitivity of the outcomes of economic models to their underlying assumptions.

3.3 An Analogy to State Dependence in Labor Market Outcomes

It is important to point out the relationship between state dependence in health and labor market outcomes. As discussed by Hyslop (1999), many sources of state dependence in labor force participation have been cited including intertemporally nonseparable preferences for leisure (Hotz, Kydland and Sedlacek 1988) and search costs which depend on participation states (Eckstein and Wolpin 1990). However, regardless of the underlying source, understanding the magnitude of state dependence in labor force participation will have policy implications since this tells us about the effectiveness of policies that alleviate short-term unemployment. Similarly, the magnitude of state dependence in health will be informative of the relative importance of unobserved individual characteristics *vis-a-vis* idiosyncratic health shocks. To the extent that the effects of these shocks can be mitigated by improvements in health care and its delivery, understanding the magnitude of state dependence in health will have implications for many health policy debates. In both the cases of labor and health economics, the statistical properties of the data will contain information that is pertinent to the conduct of policy.

4 Maximum Likelihood Estimation

We estimate the model in equation (1) using a Maximum Likelihood Estimation (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 $\tau_i \geq 0$. This causes 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

$$P(h_{i,T_i}, ..., h_{i,\tau_i+1} | h_{i,\tau_i}, \boldsymbol{\theta}_i') = \prod_{t=\tau_i+1}^{T_i} \Lambda\left(\boldsymbol{\theta}_i' \mathbf{Z}_{i,t-1}(2h_{i,t} - 1)\right). \tag{5}$$

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. Our approach is the same as Deb and Trivedi (1997) in that we assume that the population is drawn from a finite number of distinct classes corresponding to varying degrees of latent health.[6] Let $P_{\tau_i}(h_{i,\tau_i} | \boldsymbol{\theta}_a')$ denote the probability of the first observation conditional on

[6] This approach is also similar to Heckman and Singer (1984) who use a discrete distribution to approximate the distribution of unobserved heterogeneity.

$\boldsymbol{\theta}_i = \boldsymbol{\theta}_a$. We can now obtain the unconditional likelihood via

$$P(h_{i,T_i}, ..., h_{i,\tau_i}) =$$
$$\sum_{a=1}^{A} P(h_{i,T_i}, ..., h_{i,\tau_i}|\boldsymbol{\theta}'_a)\pi_a = \tag{6}$$
$$\sum_{a=1}^{A} \prod_{t=\tau_i+1}^{T_i} \Lambda(\boldsymbol{\theta}'_a \mathbf{Z}_{i,t-1}(2h_{i,t}-1)) P_{\tau_i}(h_{i,\tau_i}|\boldsymbol{\theta}'_a)\pi_a.$$

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, we let the probability of being well in $t = 0$ conditional on $\boldsymbol{\theta}_a$ be given by $p_a \equiv P_0(h_{i,0} = 0|\boldsymbol{\theta}'_a)$. The probability of observing $h_{i,t}$ conditional on $\boldsymbol{\theta}_a$ in any subsequent period is then given by

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

Substituting, we get

$$P_t(h_{i,t}|\boldsymbol{\theta}'_a) = \sum_{d_{t-1}=0}^{1} \Lambda((\alpha_a + \gamma_a d_{t-1} + \boldsymbol{\rho}_a \mathbf{T})(2h_{i,t}-1)) * \tag{8}$$
$$\sum_{d_{t-2}=0}^{1} (\Lambda((\alpha_a + \gamma_a d_{t-2} + \boldsymbol{\rho}_a(\mathbf{T}-\mathbf{1}))(2d_{t-1}-1)) P_{t-2}\left(h_{i,t-2} = d_{t-2}|\boldsymbol{\theta}'_a\right)).$$

Using the above formulation, we can calculate $P_{\tau_i}(h_{i,\tau_i}|\boldsymbol{\theta}'_a)$.[7] Of course, this is a burdensome task

[7] Heckman (1981a) proposes using this method which involves using the underlying statistical model to calcu-

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 definition simplifies matters greatly.

We now obtain the likelihood function:

$$L(\boldsymbol{\beta}) = \tag{9}$$
$$\sum_{i=1}^{N} \log \left(\sum_{a=1}^{A} \prod_{t=\tau_i+1}^{T_i} \Lambda(\boldsymbol{\theta}_a' \mathbf{Z}_{i,t-1}(2h_{i,t}-1)) P_{\tau_i}(h_{i,\tau_i}|\boldsymbol{\theta}_a')\pi_a \right),$$

where $\boldsymbol{\beta} \equiv (\boldsymbol{\theta}_1', ..., \boldsymbol{\theta}_A', \pi_1, ..., \pi_{A-1}, p_1, ..., p_A)$ and has dimension $7A-1$. The likelihood function was maximized using the Fletcher-Powell algorithm, a variant of Newton's Method, which only requires the computation of the the gradient vector. To save on computation time, we calculated analytical gradients.[8]

When the number of support points for the mixing distribution exceeds two, estimating π_a directly will often result in some trivial probabilities so that the number of support points effectively collapses to two or (sometimes) three. To avoid this, we follow Arcidiacono and Jones

late $P_{\tau_i}(h_{i,\tau_i}|\boldsymbol{\theta}_a')$ 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.

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

(2003) and note that the MLE of π_a is given by

$$\widehat{\pi}_a = \frac{1}{N}\sum_{i=1}^{N}\frac{f_{i,a}}{q_i}, \tag{10}$$

where $f_{i,a} \equiv \prod_{t=\tau_i+1}^{T_i} \Lambda(\boldsymbol{\theta}_a' \mathbf{Z}_{i,t-1}(2h_{i,t}-1)) P_{\tau_i}(h_{i,\tau_i}|\boldsymbol{\theta}_a')\pi_a$ and $q_i \equiv \sum_{a=1}^{A} f_{i,a}$. This insight suggests the following iterative strategy. First, choose a set of values for the mixing distribution probabilities and call these values π_a^1. Similarly, choose initial values for $\varpi \equiv (\boldsymbol{\theta}_1', ..., \boldsymbol{\theta}_A', p_1, ..., p_A)$ and call these values ϖ^1. Next, calculate the gradient with respect to ϖ using the probabilities π_a^1 and ϖ^1 and iterate to get ϖ^2. Then, evaluate equation (10) using π_a^1 and ϖ^1 to obtain π_a^2. Repeat the process.[9]

5 Estimation Results

5.1 Model Selection

We investigate model selection along two dimensions. The first is the specification of the index inside equation (1) and the second is the number of support points.[10] Our model selection criterion is the Akaike Selection Criterion (AIC) which is proportional to the absolute value of the likelihood function plus the number of estimated parameters (Amemiya 1985). The preferred

[9] To verify that this procedure does, in fact, work, using two support points, we calculated the MLE using this method and using the alternative method in which the probabilities π_a were estimated directly (*i.e.* we differentiated the likelihood function with respect to π_a as well). Both procedures yielded the same estimates.

[10] When testing for the number of support points, likelihood-based test statistics are inappropriate because, under the null hypothesis, one of the probabilities π_a must be set to zero. This places the parameter vector at the edge of a compact set and, thus, violates the regularity conditions of MLE. Consequently, the resulting test statistic will not be χ^2. However, as pointed out by Leroux (1992), model selection criteria do not require that the true parameter lie in the interior of a compact set and, thus, they are an appropriate means of testing for the number of support points.

model has the lowest AIC.

5.1.1 Index

In Table 2, we report the AIC for four indices with two and three support points. The indices are defined in the table. When we only have two points of support, we see that the AIC slightly favors the homogeneous quadratic model for both men and women. When we move to three points, the AIC still favors the homogeneous quadratic model for men, but now favors the heterogeneous quadratic model for women. What is important to note, however, is that our choice of index does not alter the AIC by a large margin.

5.1.2 Support Points

Table 3 reports the AIC results for the number of support points. We consider up to four points of support. We did not venture beyond four points due to computational limitations. For each value of A that we considered, we estimated the model with a homogeneous quadratic function of age.[11] We see that the AIC increases with the number of support points, but at a decreasing rate. In contrast to altering the index, adding support points has a dramatic effect on the AIC. The preferred model has four points of support which suggests that there is a tremendous amount of heterogeneity in health.[12] The results of this table stand in contrast to results in Deb and Trivedi (1997) who find that only two points of support were necessary when estimating a

[11] When choosing the number of support points, we did not concern ourselves with the index selection for two reasons. The first is that, as indicated by Table 2, changing the index did not alter the AIC tremendously. The second is that the computations in this exercise were quite intensive. Utilizing more complicated indices, such as the heterogeneous quadratic model, would only have made it worse.

[12] Presumably, if we had continued to add support points, we would have found evidence of even more heterogeneity. However, because the relationship between the AIC and A appears to be concave, we conjecture that eventually the selection criterion would have started to decline.

model for the demand of medical care.

5.2 Health Dynamics

Tables 4 and 5 report the parameter estimates and their standard errors for the homogeneous quadratic model with four support points for men and women.[13] This model had the lowest AIC of all the models that we considered.[14] Each column of the tables corresponds to a separate support point which we call a "type." We have defined each according to the magnitude of α_a. The lowest value (*i.e.* most negative) of α_a is defined to be "Type 1" and the highest is "Type 4."[15]

In Figures 1 through 4, we take the parameter estimates for men and map them into health transition probabilities.[16] Each figure corresponds to a separate type and plots two profiles. The first is the probability of being ill today conditional on having been ill yesterday. We call this profile the persistence of illness. The second is the probability of being ill today conditional on having been well yesterday. We call this profile the onset of illness. We plot 95% confidence bands around each profile.[17] It is important to realize that because many of the standard errors in Tables 4 and 5 are quite large, some of these confidence bands include zero or unity.

The figures show a large degree of heterogeneity in health. Figure 1, which corresponds to

[13] Standard errors were calculated using the "sandwich" standard errors. The gradient vector from the likelihood function was used to calculated the average of its outer product. To calculate the Hessian, we numerically differentiated the gradient vector.

[14] We also calculated the AIC for a homogeneous quadratic model with a homogeneous state dependence parameter for $A = 4$. The model with the heterogeneous state dependence parameter was preferred.

[15] It is important to emphasize that the probability of being a certain type is independent of age in our analysis. However, if we were to have modeled mortality as well, then the probability of being a particular type would depend on age since the unhealthy types would have higher probabilities of dying.

[16] The figures for women, which we do not report, were similar.

[17] The δ-method was used to calculate the standard errors.

Type 1 men, shows that the persistence of illness is close to unity and that the onset of illness is close to zero at all ages. Taken at face value, this suggests that Type 1 men exhibit a tremendous degree of state dependence. Figures 2 through 4 correspond to Types 2 through 4. Type 2 men are the healthiest and Type 4 men are the unhealthiest. These figures show far more muted degrees of state dependence than Figure 1.

Figures 5 through 8 display the degree of state dependence, which is defined to be the difference between the persistence and onset profiles for men. Each figure corresponds to a separate type and includes a 95% confidence band. The degree of state dependence is close to unity for Type 1 men and women. The degree of state dependence for Type 2 men and women is very low - below 10% for most ages. For Types 3 and 4, we see a more intermediate degree of state dependence that is somewhere between 10% and 20%.

At this point, we subject the reader to a caveat concerning the high degree of state dependence that we uncovered for Type 1 people. We conjecture that this is a consequence of the fact that these individuals have an initial probability of being ill that is below 1% and an extremely low probability of falling ill from that point onward. Consequently, the data do not contain a wealth of information on the persistence of illness for these types. This makes it difficult to pin down γ_1. Thus, we believe that our estimations are telling us that these types have very low propensities of falling ill, but are not terribly informative of their degree of state dependence. Also, it is worth mentioning that Halliday (2007a) used the same data, but alternative semi-parametric tests, and did not find strong evidence of state dependence.

To better see this, in Tables 6 and 7, we report the frequencies of 4 year sequences of health, for men and women, starting from ages 30, 40 and 50. Both tables show that the sequence

$(0,0,0,0)$ is, by far, the most frequent and so, the healthy state is highly persistent for the vast majority of the individuals in our data. In contrast, if it were the case that the degree of state dependence actually was unity for half of the population, we would also expect to see a large frequency for the sequence $(1,1,1,1)$, but we do not.

6 Conclusion

This paper investigated the evolution of health over the life-course by estimating several specifications of a flexible model of health dynamics which allowed for two sources of persistence: unobserved heterogeneity and state dependence. Our analysis suggested that altering the linear index of our model did little to improve its fit. In contrast, adding support points to the mixing distribution led to large improvements in fit. We found that at least four support points were necessary, indicating a large degree of heterogeneity in our data. This suggests that much of what determines health in adulthood can be traced back to childhood and is consistent with recent work by Case, Paxson and Lubotsky (2002). We found modest degrees of state dependence for approximately half of the population. For the other half, we found that it was near unity. However, because the likelihood of falling ill was so low for this part of the population, we do not believe that we can say anything conclusive about their degree of state dependence.

Can the estimates in this paper inform us about health policy? While this paper can be criticized as being too "reduced form," we believe that our approach, which is focused on deepening our understanding of the statistical properties of the data while make parametric restrictions that are as weak as possible, can be informative of policy. In fact, because measuring health is so difficult and incorporating it into life-cycle consumption models often results in models

that are very hard to estimate and potentially fragile in the face of mis-specified distributional and modeling assumptions, many authors such as Adams, Hurd, McFadden, Merrill and Ribeiro (2003), Adda, Banks and Van Gaudecker (2006) and Halliday (2007b) have also adopted less structural approaches in health applications.

To this end, we contend that the results of this paper shed light on the gradient: the much-studied but little-understood statistical correlation between health and socioeconomic status (Adams, Hurd, McFadden, Merrill and Ribeiro 2003). 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* improvements in health care and its delivery (Deaton 2002). The argument for health policies is further strengthened if health exhibits a high degree of state dependence since this implies that interventions will have large dynamic effects.

Our reading of the results leads us to conclude that, while improvements in medical care will lead to modest improvements in health, there may be larger potential gains to identifying and then targeting factors that influence individual heterogeneity. Our reasoning for this is that we uncover relatively modest degrees of state dependence for most people. For the rest of the population, we do uncover an enormous degree of state dependence, but we have good reasons, which we outlined above, for thinking that this is a result of the fact that this segment of the population almost never gets sick prior to age 60. On the other hand, we do uncover a large amount of heterogeneity which indicates that much of the persistence that we observe in the aggregate is driven by individual characteristics which can be traced back to early adulthood and before.

References

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

[2] Adams, H.P., M.D. Hurd, D. McFadden, A. Merrill and T. Ribeiro (2003): "Healthy, Wealthy and Wise? Tests for Direct Causal Pathways between Health and Socioeconomic Status," *Journal of Econometrics*, 112, 3-56.

[3] Adda, J., J. Banks and H.M. von Gaudecker (2006): "The Impact of Income Shocks on Health: Evidence from Cohort Data," unpublished manuscript.

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

[5] Arcidiacono, P. and J.B. Jones (2003): "Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm," *Econometrica*, 71, 933-946.

[6] Arcidiacono, P., H. Sieg and F. Sloan (2007): "Living Rationally Under the Volcano? Heavy Drinking and Smoking Among the Elderly," *International Economic Review*, 48, 37-65.

[7] 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.

[8] 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.

[9] 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.

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

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

[12] Deb, P. and P. Trivedi (1997): "Demand for Medical Care by the Elderly: A Finite Mixture Approach," *Journal of Applied Econometrics*, 12, 313-336.

[13] Eckstein, Z. and K. Wolpin (1989): "Dynamic Labor Force Participation of Married Women and Endogenous Work Experience," *Review of Economic Studies*, 56, 375-390.

[14] Grossman, M. (1972): "On the concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80, 223-255.

[15] French, E. (2005): "The Effects of Health, Wealth and Wages on Labour Supply and Retirement Behavior," *Review of Economic Studies*, 72, 395-427.

[16] Halliday, T. (2007a): "Testing for State Dependence with Time Variant Transition Probabilities," *Econometric Reviews*, 26, 1-19.

[17] Halliday, T. (2007b): "Income Risk and Health," unpublished manuscript.

[18] 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.

[19] 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.

[20] 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.

[21] Hotz, J.V., Kydland, F.E. and G.L. Sedlacek (1988): "Intertemporal Preferences and Labor Supply," *Econometrica*, 56, 335-360.

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

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

[24] 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.

[25] Leroux, B.G. (1992): "Consistent Estimation of a Mixing Distribution," *Annals of Statistics*, 20, 1350-1360.

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

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

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

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

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

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

[32] Smith, J. (2003): "Health and SES Over the Life-Course," unpublished manuscript, RAND.

[33] Viscusi, K. and W.N. Evans (1990): "Utility Functions that Depend of Health Status: Estimates and Economic Implications," *American Economic Review*, 80, 353-374.

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

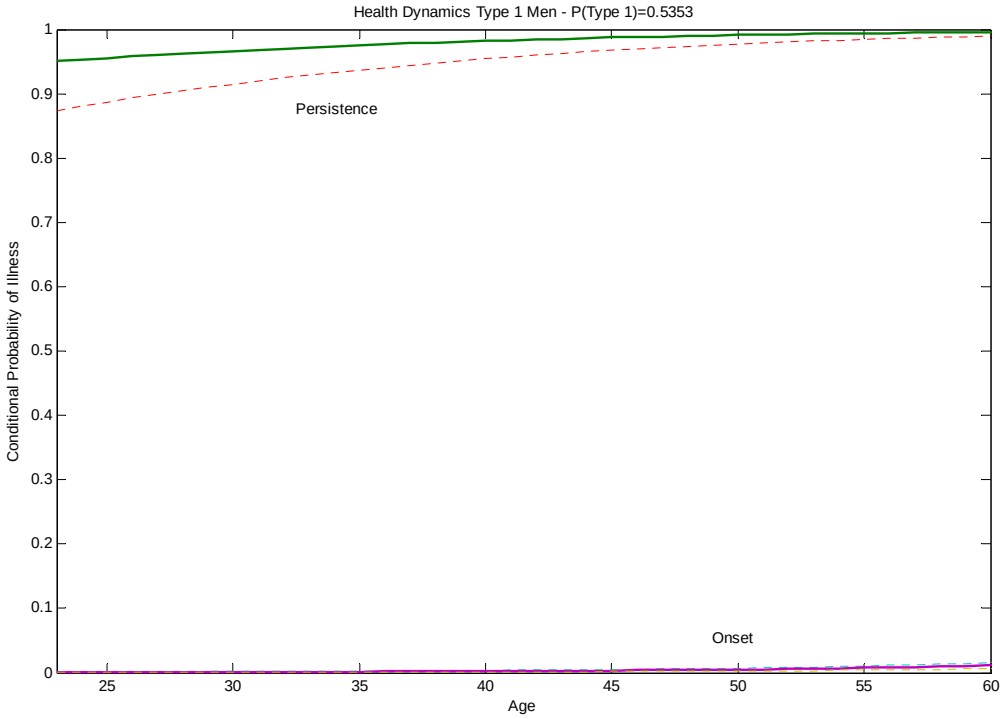

Figure 1

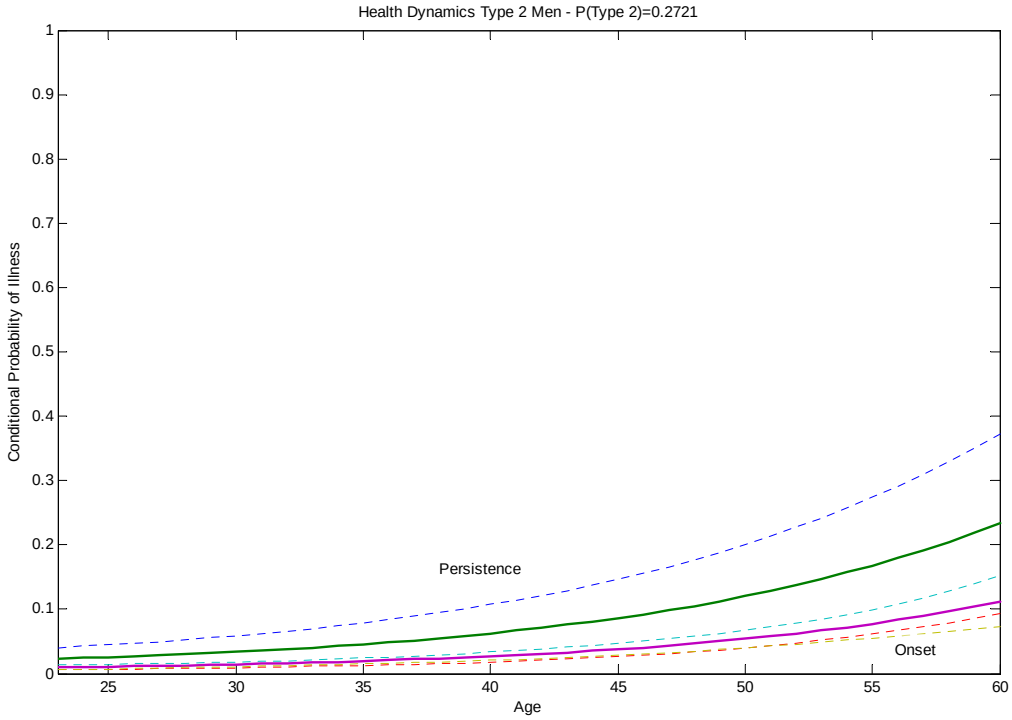

Figure 2

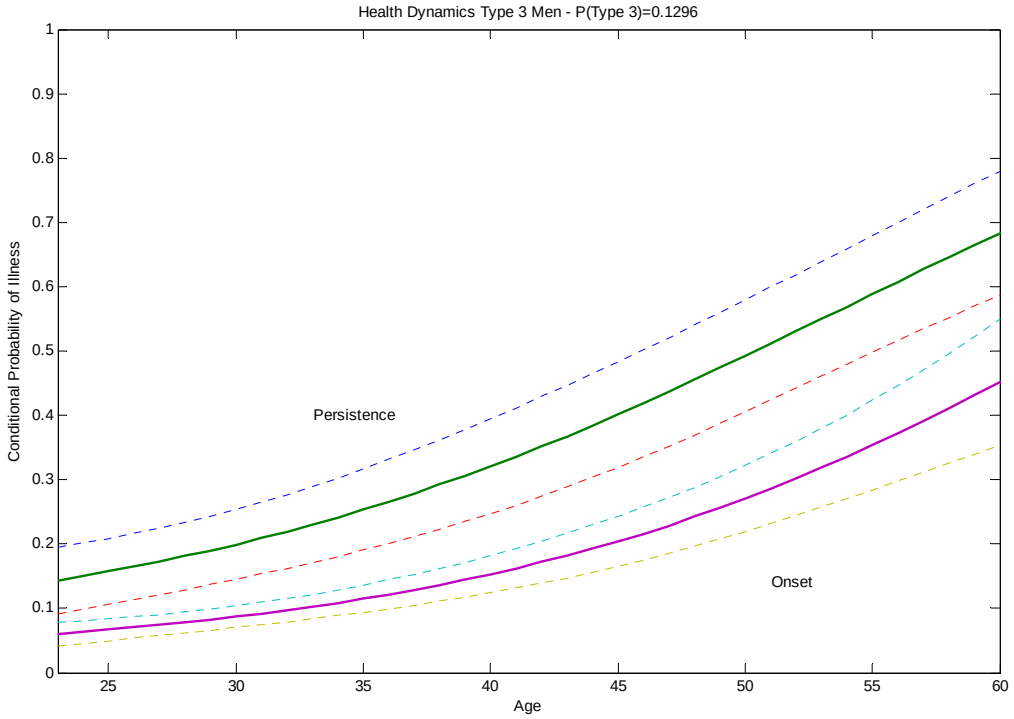

Figure 3

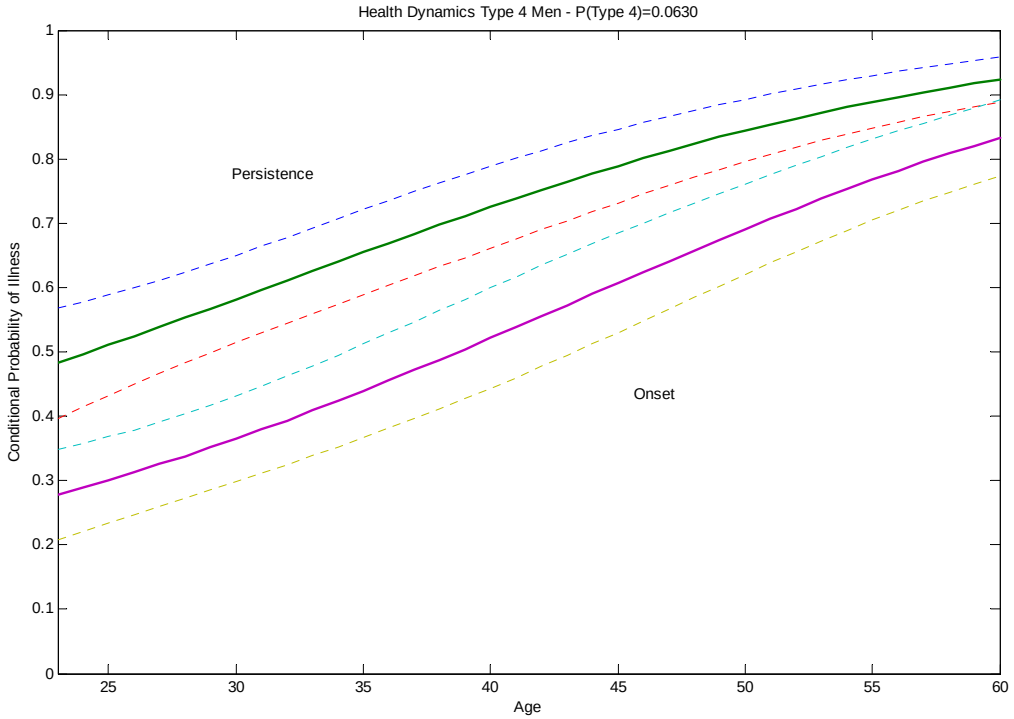

Figure 4

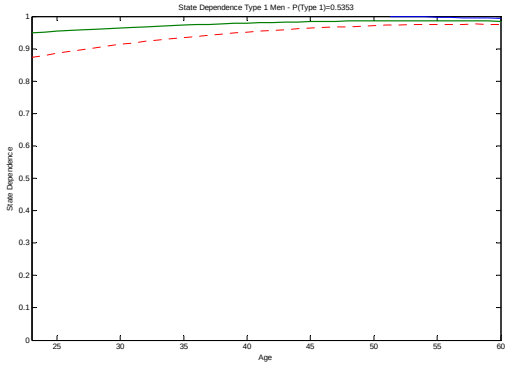

Figure 5

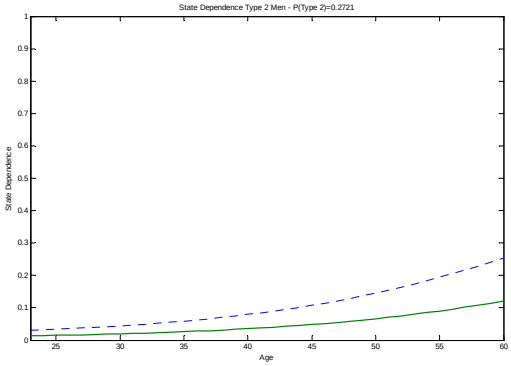

Figure 6

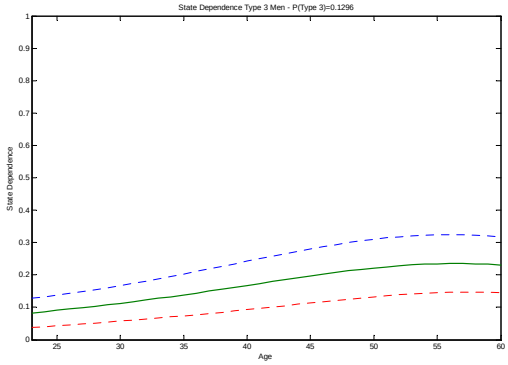

Figure 7

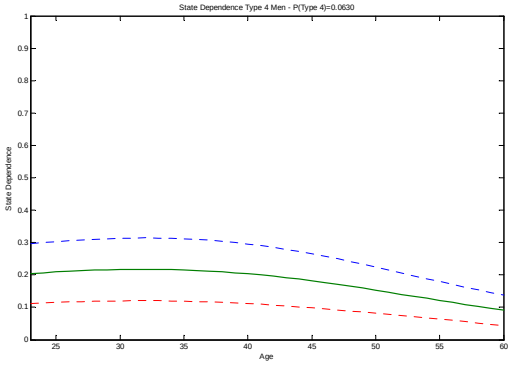

Figure 8

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: AIC for Index Selection

	Aging Function	$\boldsymbol{\rho}$ Hetero?	γ Hetero?	$A = 2$ Men	$A = 2$ Women	$A = 3$ Men	$A = 3$ Women
Linear Model	Linear	No	Yes	6163.8	7564.5	6084.9	7453.9
Homo Quad - Homogeneous γ	Quad	No	No	6164.5	7564.4	6083.8	7451.9
Homo Quad	Quad	No	Yes	6163.1*	7563.2*	6083.5*	7452.9
Hetero Quad	Quad	Yes	Yes	6163.9	7564.7	6085.6	7450.9*

*Denotes the model with the lowest AIC.

Table 3: AIC for Selection of the Number of Support Points

Points of Support	Men	Women
$A = 1$	11798.0	13631.0
$A = 2$	6163.1	7563.2
$A = 3$	6083.5	7452.9
$A = 4$	6062.9	7422.3

The homogeneous quadratic model was employed in the estimation.

Table 4: Parameter Estimates for Preferred Model - Men

	Type 1	Type 2	Type 3	Type 4
α_a	−8.0789	−5.6632	−3.7868	−1.9916
	(0.7920)	(0.7537)	(0.7400)	(0.7238)
γ_a	9.9901	0.8776	0.9597	0.8335
	(0.9093)	(0.3874)	(0.1915)	(0.2075)
ρ^1	0.3598	0.3598	0.3598	0.3598
	(0.3657)	(0.3657)	(0.3657)	(0.3657)
ρ^2	3.9879	3.9879	3.9879	3.9879
	(4.3498)	(4.3498)	(4.3498)	(4.3498)
p_a	0.9907	0.9578	0.8757	0.9999
	(0.0055)	(0.0655)	(0.1387)	(0.0001)
π_a	0.5353	0.2721	0.1296	0.0630

Table 5: Parameter Estimates for Preferred Model - Women

	Type 1	Type 2	Type 3	Type 4
α_a	−6.8666	−5.5826	−3.1537	−1.4090
	(0.7049)	(0.6721)	(0.6584)	(0.6586)
γ_a	9.5925	0.7514	0.8067	0.8779
	(1.0865)	(0.5424)	(0.1340)	(0.2033)
ρ^1	0.2494	0.2494	0.2494	0.2494
	(0.3267)	(0.3267)	(0.3267)	(0.3267)
ρ^2	4.4630	4.4630	4.4630	4.4630
	(3.8702)	(3.8702)	(3.8702)	(3.8702)
p_a	0.9997	0.9587	0.8874	0.7432
	(0.0054)	(0.0285)	(0.0868)	(0.2245)
π_a	0.4093	0.3495	0.1804	0.0608

Table 6: Health Sequence Frequencies - Men

$(h_{i,t-3}, h_{i,t-2}, h_{i,t-1}, h_{i,t})$	$t = 33$	$t = 43$	$t = 53$
$(1,1,1,1)$	14	19	42
$(1,1,1,0)$	20	35	44
$(1,1,0,1)$	3	7	15
$(1,0,1,1)$	6	9	8
$(0,1,1,1)$	20	35	44
$(1,1,0,0)$	30	63	58
$(1,0,1,0)$	8	15	11
$(0,1,1,0)$	13	35	29
$(1,0,0,1)$	4	9	9
$(0,1,0,1)$	11	17	4
$(0,0,1,1)$	27	61	65
$(1,0,0,0)$	123	137	109
$(0,1,0,0)$	97	83	59
$(0,0,1,0)$	100	85	52
$(0,0,0,1)$	123	137	108
$(0,0,0,0)$	3649	3153	1327

Table 7: Health Sequence Frequencies - Women

$(h_{i,t-3}, h_{i,t-2}, h_{i,t-1}, h_{i,t})$	$t = 33$	$t = 43$	$t = 53$
$(1,1,1,1)$	10	29	44
$(1,1,1,0)$	22	49	39
$(1,1,0,1)$	5	13	12
$(1,0,1,1)$	6	19	12
$(0,1,1,1)$	22	49	39
$(1,1,0,0)$	42	84	54
$(1,0,1,0)$	12	18	15
$(0,1,1,0)$	25	48	27
$(1,0,0,1)$	10	12	17
$(0,1,0,1)$	13	24	15
$(0,0,1,1)$	41	78	55
$(1,0,0,0)$	142	168	134
$(0,1,0,0)$	110	96	97
$(0,0,1,0)$	111	102	97
$(0,0,0,1)$	142	168	135
$(0,0,0,0)$	3736	2853	1258