NBER WORKING PAPER SERIES

THE DYNAMIC EFFECTS OF HEALTH ON THE LABOR FORCE TRANSITIONS OF OLDER WORKERS

John Bound
Michael Schoenbaum
Todd R. Stinebrickner
Timothy Waidmann

Working Paper 6777
http://www.nber.org/papers/w6777

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 1998

We are grateful to David Blau and two anonymous referees for helpful comments, and to Minsong Liang for able research assistance. This work has been supported in part by pilot grants from the National Institute on Aging, 1 RO3 AG14547 01 and 1 RO3 AG14863 01. Michael Schoenbaum has been supported by the Robert Wood Johnson Foundation's Scholars in Health Policy program and by RAND. Todd Stinebrickner has been supported by a post-doctoral training grant from the National Institute on Aging, 1 T32 AG00221 05. An earlier version of this paper was presented at the 1997 International Health and Retirement Surveys Conference in Amsterdam. The views expressed here are those of the author and do not reflect those of the National Bureau of Economic Research.

The Dynamic Effects of Health on the Labor Force Transitions of Older Workers
John Bound, Michael Schoenbaum,
Todd R. Stinebrickner, and Timothy Waidmann
NBER Working Paper No. 6777
November 1998
JEL No. J14, J22, J26

ABSTRACT

This paper addresses the interplay between health and labor market behavior in the later part of the working life. We use the longitudinal Health and Retirement Survey to analyze the dynamic relationship between health and alternative labor force transitions, including labor force exit, job change and application for disability insurance. Specifically, we examine how the timing of health shocks affects labor force behavior. Controlling for lagged values of health, poor contemporaneous health is strongly associated with labor force exit in general and with application for disability insurance in particular. At the same time, our evidence suggests that controlling for contemporaneous health, poor lagged health is associated with continued participation. Thus, it appears that not just poor health, but declines in health help explain retirement behavior. We conclude that modeling health in a dynamic, longitudinal framework offers important new insights into the effects of poor health on the labor force behavior of older workers.

John Bound
Population Studies Center
University of Michigan
1225 South University Avenue
Ann Arbor, MI 48104
and NBER
jbound@umich.edu

Michael Schoenbaum
RAND
1700 Main Street
Santa Monica, CA 90407-2138
mikels@rand.org

Todd R. Stinebrickner
Department of Economics
The University of Western Ontario
Social Sciences Centre
London, Ontario -- CANADA
N6A 5C2
tristenb@jullian.uwo.ca

Timothy Waidmann
The Urban Institute
2100 M Street, NW
Washington, DC 20037
twaidman@ui.urban.org

1. Introduction

The relationship between health and retirement is a dynamic process that can best be examined longitudinally. Health declines with age and may require adaptation or cessation of work activities. Although variation in health status exists at all ages and affects early educational and occupational attainment, it is the decline in health starting in late middle age that is likely to create a mismatch between an individual's capabilities and the requirements of his or her job. Whether and how workers respond to declines in health depends on various factors, including the nature of the declines; their expected persistence; the age at which they occur; and the worker's human capital, economic situation, and preferences for leisure and consumption. Research on the effect of health on retirement has virtually ignored these dynamic issues (Anderson et al., 1986, is a notable exception).

In this paper, we examine the dynamic relationship between health status and labor force behavior among older working-age adults. Consistent with previous studies, we focus on the effects of contemporaneous health status on labor market transitions, but we also include lagged values of health in our models. This permits us to distinguish the labor market effects of persistently poor health from those of health declines. The fact that we include both contemporaneous and lagged health in our models tends to exacerbate problems associated with measuring health status, so we pay particular attention to addressing problems of measurement error and endogeneity.

In addition, previous studies on the effects of health on labor force behavior have focused almost exclusively on explaining labor force exit. However, researchers increasingly view retirement as a process rather than a single event (Honig and Hanoch, 1985; Honig, 1985; Quinn et al., 1990; Ruhm, 1990; Quinn, 1997, 1998) . With the notable exceptions of Honig and Reimers (1987) and Blau et al. (1997), there have been virtually no attempts to model the effect of health on labor force transitions other than retirement.[1]

[1]Using the Retirement History Survey, Honig and Reimers (1987) find little association between poor health and the move from full-time work to partial retirement. Using the HRS, Blau et al. (1997) find that

A more general literature (Baltes and Baltes, 1990; Brim, 1988) documents the adaptations that older adults make in response to deteriorating health. Ceasing to perform an activity is only one response and often the response of last resort. Before this occurs, older adults will expend increased effort, allow more time, and reduce performance standards in order to perform the activity. In the context of labor force behavior, poor health may induce some people to change jobs or even employers, or find ways to accommodate their limitation on the current job, in addition to inducing many people to leave the work force altogether. In this paper, we model these alternatives explicitly.

2. Conceptual Framework

As we have mentioned, the gerontological literature suggests that it is not so much poor but deteriorating health that explains behavior as individuals age. It is possible interpret this view within the context of standard intertemporal labor supply models (Blundell and MacCurdy, forthcoming). We assume that individuals maximize the expected value of future utility:

$$\text{Max}\,\text{E}_t \sum_{j=t}^{T} b^{\,j-t}\,\text{U}(C_j, L_j, Z_j) \tag{1}$$

subject to an intertemporal budget constraint:

$$A_{j+1} = (W_j H_j - C_j) + (1 + r_{j+1}) A_j \tag{2}$$

where C_j and L_j represent consumption and leisure as of time period *j*, respectively; Z_j, a vector of age-specific taste shifters (we imagine that health is an important component of Z_j); W_j, the wage (which we also imagine is a function of health); H_j, hours of work; A_j, assets; and r_j, the rate of interest, all as of time *j*. b is a time rate of preference discount factor; in contrast to the rate of interest, we assume that the preference discount factor is time-invariant. Raising b to the power (*j*-*t*) reduces the weights on utility in future periods in the intertemporal utility function. Consumption and hours of work represent choice variables. Current period taste

the effect of health on the probability that a person changes jobs is much smaller than its effect on labor force exit.

shifters, wages, and assets are known, while future period (*j*>*t*) taste shifters, wages, and rates of interest are exogenous but unknown.[2]

Given the separability built into the utility function, the first order conditions for this problem have a particularly simple form:

$$\mathrm{U}_C(C_t, L_t, Z_t) = l_t \tag{3}$$

$$\mathrm{U}_L(C_t, L_t, Z_t) = l_t W_t \tag{4}$$

where l_t represents the marginal value of wealth as of time *t*, which itself will be a function of the distribution of possible future wages, interest rates and taste shifters.

Equation 4 can be used to understand the labor supply impact of health declines. Health declines can be expected to lower wages and raise the relative valuation of leisure. Both of these effects will work to lower work hours. Health declines will also affect the marginal utility of wealth, though the magnitude of this effect depends on the extent to which health declines are unanticipated and the extent to which they are expected to persist.

To help fix ideas, we first suppose that an individual's health follows a first-order autoregressive process (AR(1)):

$$h_t = a h_{t-1} + v_t \tag{5}$$

In Equation 5, h_t represents health status in time *t*, v_t a random shock occurring in time *t*, and a a numerical constant. Assuming that health shocks are persistent (i.e., $a \approx 1$), negative health shocks can be expected to lower future wages, and to raise future valuations of leisure. As a result, the shock raises the marginal value of wealth, which will tend to increase labor supply. In this framework, the later in life a health shock occurs (i.e., *t* closer to *T* in this model), the smaller its effect on wealth. As a result, the later a health shock occurs during a person's working life, the greater the chance it will induce him or her to leave the work force.[3,4]

---

[2]While there is a large literature that treats health as a form of human capital and thus as endogenous, the literature on the effect of health on labor force behavior has, in general, done as we do and treated health as exogenous.

[3]The magnitude of these effects will depend on a, but the direction will be as described as long as a is big enough that health shocks last through the remainder of a person's working life.

Although our simple model has assumed that wages are exogenous, in general it may be possible for individuals to adapt to health limitations by learning new skills. In this case, future wages will depend on the extent of such investment. The magnitude of such investments will depend on a person's time horizon: the closer he or she is to retirement age, the less he or she will invest in new skills. Thus, the endogeneity of wages will serve to accentuate the effects outlined above: the older workers are when suffering a decline in health, the larger the resulting decline in wages and the more likely they are to leave the work force.[5]

We are particularly interested whether lagged values of health affect labor force behavior even after controlling for contemporaneous health status. We have been assuming that health follows an AR(1) process: while current health helps to predict future health, lagged health contains no predictive information when we condition on current health. In this case, if we compare two individuals who are both currently in poor health, but one of whom has been in persistently poor health while the other was in good health until recently, the two should have comparable expectations about their future health. The only difference between the two is that one suffered a health decline earlier than the other.

In contrast, we might imagine that health does not follow an AR(1), but rather that people who have been in persistently poor health are less likely to recover than are people who have recently suffered a decline in their health. In this case, the person who has been in persistently poor health will have worse expectations about the future than will the person who recently experienced a health shock.[6] If labor market transitions tend to be irreversible (e.g., once one has

---

[4]Previous research using retrospective (Daly and Bound, 1996) and prospective data (Charles, 1997) has found evidence consistent with the notion that the earlier in life a health shock occurs, the less likely it is to lead to (immediate) labor force withdrawal.

[5]Consistent with this possibility, Charles (1997) has found evidence that people who suffer health shocks early in their working lives experience more wage recovery than those who experience such shocks at older ages.

[6]For concreteness, we have been imagining the situation that someone in persistently poor health is more likely to continue in poor health than someone who has recently begun to experience poor health. However, the reverse is clearly theoretically possible. In either case, health will not follow a simple Markov process, with the implication that not only current, but also lagged health will affect expectations regarding future

left a job, one cannot return to that same job), then expectations that one's poor health is likely to persist will increase the odds that a person will leave work. This suggests lagged effects in the same direction as the effects of contemporaneous health.

It also seems possible that lagged health might affect current behavior simply because transitions take time. It may take some time before an individual learns whether or not his or her employer can or will accommodate a health limitation. It may also take time before one can line up a new job. To the extent that these effects are important, they will also generate lagged health effects that work in the same direction as does current health.[7]

In sum, our conceptual framework suggests that we will observe a relationship between lagged values of health and current labor force behavior. When controlling for contemporaneous health, the direction of the effects of lagged health on labor supply is theoretically ambiguous.[8] The direction will depend on the relative importance of workers' adaptations to enable continued work, the extent to which lagged health affects expectations of future health, and the extent to which workers need to prepare before being able to exit the labor force.

3. Data

Data for this research come from the first three waves of interviews of the Health and Retirement Survey, a multi-purpose social science survey conducted by the Survey Research Center (SRC) at the University of Michigan and funded by the National Institute on Aging. The first wave of the survey was conducted in 1992/93; respondents were re-interviewed in 1994 and 1996 and will be re-interviewed at two-year intervals in the future. We use the baseline public

---

health.

[7]The nature of the HRS data tends to exacerbate this issue. The health measures we use were obtained at two-year intervals. If health changed between Wave 1 and Wave 2 or between Wave 2 and Wave 3, we have no idea when in the interval the change occurred. If, the change occurred immediately prior to the Wave 3 survey it may be quite unlikely that the respondent would have already left their job.

[8]Without controlling for contemporaneous health status, we would expect to see that poor health in previous periods raises the likelihood of leaving the workforce. As we will see, declines in health tend to persist. Thus those whose health was poor in the past are likely to experience poor health contemporanously.

release, the beta release of Wave 2, and the alpha release of Wave 3 here. The HRS is described in additional detail in Juster and Suzman (1995).

The HRS covers a representative national sample of non-institutionalized men and women born between 1931 and 1941 (inclusive), so that respondents in the sample frame were aged 50-62 at the time of the first wave. The HRS is the first dataset of its kind to sample women at the same rate as men (*cf.* the Retirement History Survey), thus permitting comprehensive analysis of women's retirement patterns that has not been possible before. In addition, the HRS oversamples Blacks, individuals of Mexican descent, and residents of the state of Florida to permit reliable analysis of these groups. The first wave of HRS was conducted in person in respondents' homes; the response rate was 82%. The total sample size of the first wave is 12,654 respondents. The second wave of the HRS was conducted by telephone; the second wave re-interviewed 11,642 respondents, representing 92% of the original sample.

3.1 Sample

The HRS includes the spouses/partners of the survey population even if they are not themselves in the age range of the sample frame; since respondents out of the sample frame do not constitute a representative sample, they are excluded here. The age-eligible first wave sample consists of 9,824 respondents. From this group, we exclude 274 respondents who were lost to follow-up or had died by Wave 3 (3%), and 1586 respondents who were ever interviewed by proxy (16%). Proxy respondents were not asked all the detailed health questions we use here and we were reluctant to impute these health characteristics for proxy respondents, especially given our focus on measurement error and endogeneity in self-reported health data. We were also concerned that responses for proxies and non-proxies might be sufficiently different to confound our analyses. For similar reasons, we excluded 1263 respondents with missing data for any of the health or demographic measures used here (16% of age-eligible non-proxy respondents). The sample used for analysis thus includes 6701 respondents, 2875 men and 3826 women.

3.2 Measures

Outcomes. Because we are interested in modeling the effects of both contemporaneous and lagged health, we choose to model labor force transitions between Waves 2 and 3 of the HRS to maximize the number of lagged periods for which health data are available.[9] Specifically, we examine the sample of respondents who were employed at Wave 2, defined as respondents who report that they are employed and that they are not currently receiving or have a pending application for Social Security Disability Insurance (DI), Supplemental Security Income (SSI), or disability insurance benefits from any other program except Worker's Compensation or Veteran's benefits.[10] Among these people, we examine four alternative transitions between Waves 2 and 3 (with the following hierarchy): whether they applied for any disability insurance since Wave 2, are employed at the same job, are employed at a different job, or are both not employed and have not applied for disability insurance. These categories are defined to be mutually exclusive.

Demographic. All analyses are conducted separately by sex. The multivariate models outlined below include controls for several demographic and other individual characteristics. These include age in years; educational attainment, given by four categories (less than a high school diploma, high school graduate, some college, college graduate); whether the respondent is non-Hispanic White, non-Hispanic Black, Hispanic, or other; whether the respondent is currently married with spouse present or living with a partner (a single dummy variable); whether the respondent was born in the US; and whether the respondent is a veteran of the US armed forces (omitted for women due to very low prevalence). We also include dummy variables to control for

[9]I.e., in this case, we examine contemporaneous (Wave 3), once-lagged (Wave 2) and twice-lagged (Wave 1) health.

[10]We do include the small number of respondents who report being employed but on sick leave, or who are employed but also identify themselves as "retired." We exclude disability insurance participants even if they report that they are employed because DI, SSI and many federal and state programs prohibit participants from substantial work, and we consequently expect that their work is quite limited. On the other hand, Worker's Compensation participation tends be of short duration, at least relative to DI/SSI participation, and vets' disability does not preclude substantial employment.

whether the respondent has reached age 62 – the Social Security early retirement age – by Wave 2 or Wave 3.

Health. The HRS includes respondents' self-assessed general health status and respondents' reports of the extent to which health limits their ability to work. The HRS also includes a wide range of more detailed measures of self-reported health. We focus on measures of limitation in physical function here. These functional limitation measures assess respondents' difficulty performing 17 activities of daily living (ADLs) and instrumental activities of daily living (IADLs), ranging from the ability to walk, climb stairs, and lift objects such as a bag of groceries, to the ability to use a telephone and pick a dime off a table. We focus on these measures because they are powerful predictors of self-assessed general health status. In exploratory analyses we also included measures of the prevalence of various chronic diseases, but our substantive findings did not change.

Because we are interested in the dynamic response of labor market status to health, we include contemporaneous (Wave 3), lagged (Wave 2) and twice-lagged (Wave 1) health status in our models (to enable analysis of changes over time, all relevant survey questions were repeated in all three waves).

4. Analysis Strategy

4.1 Modeling Health

We seek appropriate ways to measure respondents' health in modeling labor force transitions. The most common health measures used in retirement research have been global questions such as, "Does health limit the amount or kind of work you can perform?" or "How would you rate your health? Is it excellent, very good, good, fair or poor?" However, there are a number of potential problems with such survey measures (Parsons, 1982; Myers, 1982; Anderson and Burkhauser, 1984, 1985; Bound, 1991; Waidmann et al., 1995). First, respondents are asked for subjective judgments, and these judgments may not be entirely comparable across respondents. Second, responses may not be independent of the very labor market outcomes an investigator hopes to explain. Third, since health may represent one of the few legitimate reasons for a working-age adult to be out of work, respondents out of the labor force may mention health limitations to rationalize their behavior. Fourth, since early retirement benefits are often available only for those deemed incapable of work, respondents will have a financial incentive to identify themselves as disabled, an incentive that will be particularly high for those for whom the relative rewards from continuing to work are low. In short, there has been much concern that reporting differences across individuals implies that global measures of health are endogenous to labor force status.

It is important to note that each of these problems will lead to different kinds of biases (Bound, 1991). The lack of comparability across individuals represents measurement error that is likely to lead one to underestimate the impact of health on labor force participation, while the endogeneity of self-reported health is likely to exaggerate its impact. Biased estimates of health's impact on outcomes will also lead to biased coefficients on any variable correlated with health. Finally, the dependence of self-reported health on economic characteristics will lead to biased estimates of the impact of economic variables on participation, regardless of whether one correctly measures the impact of health itself. Longitudinal analysis of the impact of health on retirement will tend to exacerbate these problems. Over a short period of time, one is unlikely to experience many dramatic health status changes. As a result, many observed changes may be spurious (Mathiowetz and Laird, 1994).

Compared with the global measures, the more detailed health indicators in the HRS may be less susceptible to measurement and endogeneity problems, since the questions are narrower and more concrete. Thus, another way to model health is to include each of the detailed health measures as explanatory variables. This makes maximal use of the available information on health status. However, doing so presents difficulties in interpretation. First, there is no obvious way to quantify the marginal effect of changes in health *per se* on the outcomes of interest. Second, the various detailed measures are presumably collinear to some degree (e.g., due to co-morbidity), and such collinearity would also complicate interpreting the estimated coefficients on particular health measures. Finally, even the numerous measures available in the HRS only partly describe individual health; they are subject to measurement error (Edwards et al., 1994; Mathiowetz and Laird, 1994); and, with respect to specific conditions, they cover prevalence but provide little information regarding severity.

The strategy we choose here is to use a latent variable model to construct an index of health. Specifically, we imagine that health as of time t is a linear function of some exogenous factors, such as age and education, X_i, a vector of detailed health measures (i.e., functional limitations in this paper), Z_{t_i}, and other unobserved factors n_{t_i}.

$$h_{t_i} = X_i'\Pi_t + Z_{t_i}'g_t + n_{t_i} \qquad (t = 1,2,3) \tag{6}$$

We assume that n_{t_i} is uncorrelated with both X_i and Z_{t_i} (this assumption is essentially definitional: n_{t_i} is the part of h_{t_i} that is uncorrelated with either X_i or Z_{t_i}). While we do not directly observe h_{t_i}, we do observe self-reported general health status, h_{t_i}, a categorical variable with five levels (excellent, very good, good, fair, and poor). Letting $h_{t_i}^*$ represent self-reported health, the latent counterpart to h_{t_i}, we assume that $h_{t_i}^*$ is a simple function of h_{t_i} and a term reflecting reporting error:

$$h_{t_i}^* = h_{t_i} + m_{t_i} \tag{7}$$

We assume that m_{t_i} is uncorrelated with h_{t_i}. Substituting (6) into (7), we obtain an equation we can estimate:

$$h_{t_i}^* = X_i'\Pi_t + Z_{t_i}'g_t + [n_{t_i} + m_{t_i}] \tag{8}$$

$$h_{t_i}^* = X_i'\Pi_t + Z_{t_i}'g_t + u_{t_i} \tag{8'}$$

Assuming that u_{t_i} is distributed normally, Equation (8) defines an ordered probit model.

The error term in Equation (8), $u_{t_i} = [n_{t_i} + m_{t_i}]$, reflects a number of different factors. The n_{t_i} component reflects aspects of health not captured by X_i and Z_{t_i}, while the m_{t_i} component reflects reporting errors. These errors reflect differences in reporting behavior across individuals and across time for the same individual. The presence of the m_{t_i} terms introduces a number biases in our estimates if we were to use the $h^*_{t_i}$ terms directly when estimating the impact of health on labor market outcomes. If the m_{t_i} terms were completely random, they would represent classical measurement error which will attenuate the estimated effect of health on labor market outcomes. On the other hand, if people use health as a way to rationalize labor market behavior, then one would expect the m_{t_i} terms to be correlated with labor market status. In this context, the use of global self-reported health measures might well exaggerate the effect of health.

To create a summary measure of health, we estimate Equation (8) to obtain estimates of Π_t and g_t. We then use the estimated coefficients to construct

$$\hat{h}^*_{t_i} = X_i'\hat{\Pi}_t + Z_{t_i}'\hat{g}_t \tag{9}$$

for each individual as a proxy for health status in our labor force models. The model is ordered so that larger (more positive) values represent worse health.

Essentially, our latent variable model uses the detailed health information available in the HRS to instrument an endogenous and error-ridden self-reported health measure, $h^*_{t_i}$. Using one error-ridden proxy to instrument another represents one of the more common ways for dealing with errors in variables (Fuller, 1987; Griliches, 1974). In our context, there are three error-ridden variables which we wish to instrument, $h^*_{1_i}$, $h^*_{2_i}$ and $h^*_{3_i}$. In the spirit of IV estimation, we use all of the available exogenous information available to construct instruments for $h^*_{1_i}$, $h^*_{2_i}$ and $h^*_{3_i}$. In particular, this implies using detailed information from all three waves when constructing each of the three proxies $h^*_{t_i}$. Similar IV approaches have been used in the cross sectional analysis of labor market behavior (Stern, 1989; Bound et al., 1996).

In many ways, the method used here is similar to methods used in the health science literature to summarize detailed health measures. In contrast, we derive weights for combining particular detailed health measures within the context of our econometric framework. The health science literature includes numerous alternative methods for summarizing health data, in particular for summarizing information on functional ability into a single measure (e.g., Suurmeijer et al., 1994). Other researchers have summarized measures of morbidity (e.g., House et al., 1990). The

CES-D depression scale (Radloff, 1977) is included in the HRS. In general, each summary measure represents a weighted average of a number of components, where the weighting scheme is derived theoretically but outside the context of an econometric model.

It is worth mentioning another respect in which $\hat{h}^*$ differs from other indices in the literature. The fact that $\hat{h}^*$ is a function not only of *Z* (the detailed health measures), but also of *X* (other individual characteristics), means that $\hat{h}^*$ has been adjusted for the extent to which, for all the reasons mentioned above, the detailed health measures available in the HRS only partially measure health differences across the population. As a result, even if the *Z* vector does not adequately account for health differences in the population, $\hat{h}^*$ will. $\hat{h}^*$ will act as a valid proxy for h_{t_i} even when both h_{t_i} and Z_{t_i} suffer from measurement error and even when there is a strong correlation between $h^*_{t_i}$ and the error in the behavioral equations we are trying to estimate.

In this framework, $\hat{h}^*$ will act as a valid proxy for h_{t_i} under two conditions. The first is that Z_{t_i} be exogenous, i.e., that there be no correlation between Z_{t_i} and the errors in the behavioral equations we are interested in estimating. This seems plausible for the functional limitation variables we use here, since they do not require physician diagnosis, and since the nature of the questions -- which gauge the ability to perform specific, narrow and routine tasks -- may reduce the scope for rationalization.[11] The second condition is that X_i must not enter the reporting equation (Equation 7), for instance because those more likely to be out of work, minorities, and those with less education might be more likely to report themselves in poor health. However, violation of this second condition will only bias the estimated effects of demographic variables in our behavioral models, which is not of central concern in this study; parameter estimates for health will be consistent.

[11]It is possible to test this assumption by comparing estimates based on one set of Z variables with estimates based on a different set. In ongoing research (Bound et al., 1998) we are doing just that; our preliminary results suggest that the parameter estimates in our behavioral equations are insensitive to our choice of Z variables. In particular, estimates that use such things as reports of heart attacks to form Z give results similar to estimates that use the functional limitation measures. Since we suspect that the reporting of heart attacks is unlikely to be endogenous to labor force status, we conclude that the functional variables used in this paper would also appear to be relatively immune from endogenous reporting errors.

4.2 Modeling Labor Force Outcomes

The remainder of this section describes our econometric framework for examining the effect of health on labor force behavior. As discussed above, deteriorating health may lead people to change their labor market behavior in a variety of ways. In particular, we imagine that deteriorating health may induce individuals to leave the workforce, change jobs or apply for disability insurance. To examine these possibilities, we estimate labor force models that distinguish between those who leave the workforce without applying for disability insurance, those who leave the workforce to apply for disability insurance, and those who continue to work but change jobs. We assume that the value to an individual of each of these respective outcomes can be approximated as a linear function of the same exogenous factors that affect health, X_i; health as of time 1, h_{1_i}, time 2, h_{2_i}, and time 3, h_{3_i}; and other unobserved factors, e_i^k:

$$V_i^k = X_{i'} b^k + l_3^k h_{3_i} + l_2^k h_{2_i} + l_1^k h_{1_i} + e_i^k \qquad (k = r, d, c, b) \qquad (10)$$

where k=r,d,c,b indexes the possible choices the person could make at time 3 (r for retirement, d for application for disability benefits, c for changing jobs, b for remaining in the same job as before). As above, we do not directly observe true health, h_{1_i}, h_{2_i} and h_{3_i}. However, we do observe the discrete counterparts of the self-reported latent health variables, $h_{1_i}^*$, $h_{2_i}^*$ and $h_{3_i}^*$.

We therefore substitute the values of true health from Equation 6 above into Equation 10 to get:

$$\begin{aligned} V_i^k = {} & X_{i'}[\,b^k + (l_3^k)\Pi_3 + (l_2^k)\Pi_2 + (l_1^k)\Pi_1] + Z_{3_i}'\,[(l_3^k)g_3] \\ & + Z_{2_i}'\,[(l_2^k)g_2] + Z_{1_i}'\,[(l_1^k)g_1] + [(l_3^k)n_{3_i} + (l_2^k)n_{2_i} + (l_1^k)n_{1_i} + e_i^k] \end{aligned} \qquad (11)$$

In the context of this multinomial model, we treat continuing to work for the same employer as the base case. To identify the coefficients in this multinomial model, we specify the following differenced equation:

$$\begin{aligned} V_i^k - V_i^b = {} & X_i'\left[\left(b^k - b^b\right) + \left(l_3^k - l_3^b\right)\Pi_3 + \left(l_2^k - l_2^b\right)\Pi_2 + \left(l_1^k - l_1^b\right)\Pi_1\right] \\ & + Z_{3_i}' g_3\left(l_3^k - l_3^b\right) + Z_{2_i}' g_2\left(l_2^k - l_2^b\right) + Z_{1_i}' g_1\left(l_1^k - l_1^b\right) \\ & + \left(l_3^k - l_3^b\right)n_{3_i} + \left(l_2^k - l_2^b\right)n_{2_i} + \left(l_1^k - l_1^b\right)n_{1_i} + e_i^k - e_i^b \end{aligned} \qquad (12)$$

If the stochastic variables in the previous equations are jointly normally distributed, then Equations (8) and (12) specify a series of six correlated equations which are used to estimate the model coefficients. Equation (8) specifies three ordered probit models for self-reported health which can be used to identify the coefficients Π_t and g_t once the variance of the stochastic processes in Equation 8 is normalized to unity.

We use the multinomial probit models in Equation (12) to identify the other coefficients. However, as is well-known in discrete choice models, coefficients can only be identified relative to the base case. The vectors of differenced coefficients are identified, given a normalization of the variance of the composite errors in the differenced equations to be 1. The fact that only differences can be estimated implies that without loss of generality we can set the coefficients of the base case equal to zero and estimate other coefficients relative to these zero values.

We freely estimate the correlations between the health equations and each of the equations in (12). While in theory it should also be possible to allow the errors across the various equations in (12) to be freely correlated, in practice these correlations are only tenuously identified (Keane, 1992). To deal with this we assume that e_i^b and the various e_i^k are all independent across equations, which implies that the correlation between equations in (12) is being driven by the common unobserved components, n_{3_i}, n_{2_i} and n_{1_i}. In this context, $Var(n)$ is not identified without making strong assumptions about independence of the stochastic components of the model; as a result, we choose to parametrically vary $Var(n)$.

We note one aspect of using our method for constructing health proxies in the context of these longitudinal models: when we use Wave 1, Wave 2, and Wave 3 functional limitation variables to predict overall health status in each wave ($\hat{h}^*$), the three proxies are very highly positively correlated. As a result, when these variables are included in behavioral models, the parameter estimates on the various health variables will be negatively correlated, and anything that affects the coefficients on one health variable will tend to affect the coefficients on the others in the opposite direction.

As specified to this point, the labor force model (i.e., Equation 12) is defined for a self-selected sample of the population: those working as of time 2. In particular, Waves 1 and 2 health are extremely strong predictors of Wave 2 employment status. Those in poor health at Wave 2 who continue to work presumably have unobserved characteristics that encourage work. For instance, they may be in better health than our health proxy suggests, they may have a strong commitment to the workforce, or perhaps they lack the financial capacity to leave the workforce. In any of these cases, we would expect that those who were in poor health but also working as of Wave 2 would be more likely to continue to work as of Wave 3, than would a random sample of those in poor health as of Wave 2.[12] This self-selection will tend to lead us to underestimate the effect of health on labor force behavior. To correct for this, we add to our model an additional equation which represents baseline work status. This equation is analogous to the labor force outcome equation (Equation 12) but models the binary outcome of whether the respondent is working as of Wave 2, and it is jointly estimated with the other equations in the model.

While taking account of the self-selected nature of our sample is important, doing so in a credible way is not easy. Given the non-linear nature of our model, the correlation between the selection equation and the retirement equations – which we designate as r_k, (k=r,d,c) – is formally identified. However, such identification is coming solely from the functional-form assumptions we have made – assumptions that we have made more for convenience than because they are theoretically justified.

One alternative we tried was to find variables that might affect the selection but not the retirement equations.[13] However, even here, identification is coming partly from the functional-

[12]Put differently, we would expect that individuals adapt in one way or another to deteriorating health. For most of those in poor health as of Wave 2, this adaptation may have already occurred – many will have left the work force. For those who have not already left, workforce attrition may not be that much higher than for those in good health.

[13]In our case, factors that have a strong influence in the timing of retirement, such as defined benefit pensions, can be used to construct variables that should influence behavior as of Wave 2 but would not influence behavior in Wave 3 (e.g. controlling for pension eligibility at Wave 3, we imagine that eligibility for pensions as of Wave 2 should not affect behavior in Wave 3). The pension and age variables we constructed in this way did have powerful effects on the timing of retirement for men, and estimated

form assumptions and partly from the specified exclusion restrictions. As a result, the results presented here are based on a strategy of parametrically varying the components of r_k through what we believe is a reasonably large range.[14] In our case, it seems plausible that the components of r_k are negative. The residual in the selection equation represents preferences for work together with unmeasured pecuniary and non-pecuniary rewards for working, some of which may be job-specific. Presumably these factors will all be correlated strongly and positively through time. Thus, the observed factors that lead a person to work as of Wave 2 will be positively correlated with working, and with working in the same job, as of Wave 3.

The models are estimated by full maximum likelihood. For the multinomial retirement decision, likelihood contributions are derived from as many as seven correlated equations, and this large number of equations implies that the integral necessary to compute the likelihood contributions (the joint probabilities) cannot be evaluated analytically or precisely approximated using standard numerical integral approximation techniques. Instead, we estimate the models using simulated maximum likelihood, taking advantage of simulation techniques recently derived by Geweke (1991), Hajivassiliou (1990), and Keane (1994) for the multinomial probit problem (referred to as the GHK simulator). The GHK method provides a way to recursively simulate the values of the stochastic components of the model and to compute the desired joint probability for a particular person conditional on this set of values. The simulated likelihood contribution for the person is obtained by repeating this process D times and averaging the resulting D joint probabilities (D=20 here). The simulations are done using antithetic acceleration (Geweke, 1988).

---

correlations based on such exclusion restrictions are consistent with the range of r we report in the text. However, for women, the pension variables had only a weak effect on the timing of retirement.)

[14]The strategy we follow is consistent with the strategy advocated by Leamer (1978).

5. Results

5.1 Descriptive Statistics

Table 1 shows that three-quarters of men in our HRS sample were employed at Wave 2. Employed men were about 18 months younger, on average; they also had significantly higher educational attainment and were more likely to be married than men who were not employed. There were no differences in the fraction who were veterans or born in the US. As has been shown previously, men who are not employed are disproportionately Black. Employed women also have relatively high educational attainment, but they are less likely to be married than women who are not employed. Also in contrast to men, Black and White women are similarly represented among the employed and not employed, while Hispanic women are less likely to be employed.

As expected, the distributions of general health status and self-reported work limitation vary dramatically between employed and not-employed people of both sexes. Respondents who are employed are much more likely to report their health as "excellent" and are very much less likely to report their health as "poor." Results for work limitations are even starker: fewer than 10 percent of employed people identified themselves as having any health problem that limits or prevents work, compared with 33 percent of not-employed women and 43 percent of not-employed men.

Table 2 gives the distribution of labor force status at Wave 3 among respondents who were employed at Wave 2, along with demographic and health characteristics by labor force status. Among men, 85 percent of the individuals employed at Wave 2 were still employed at Wave 3; of these, 78 percent were in the same job as at Wave 2. Fifteen percent of men left employment; of these, 18 percent applied for some type of disability insurance and the rest left the labor force without applying for disability coverage. Distributions are generally similar among women.

Among both men and women, people who changed jobs are somewhat better educated than people who stayed at the same job. Married women are somewhat underrepresented among female job changers, while Blacks are somewhat underrepresented among male job changers.

Compared with respondents who stayed employed, respondents who applied for disability have disproportionately lower education. They are also more likely to be Black and less likely to be married, especially among women. In terms of education, race/ethnicity and other demographic characteristics, men and women who left employment without applying for disability look roughly like respondents who stayed employed, although such men are disproportionately lower educated and are less likely to be married, and such women are more likely to be married.

Comparing health across the columns of Table 2 shows very large differences and clear longitudinal patterns. As with other characteristics, respondents who kept the same job and those who changed jobs have similar distributions of general health and work limitations. Compared with people who stayed employed, people who left Wave 2 employment to apply for disability coverage report worse general health and much higher work limitation at Waves 1 and 2. By Wave 3, these large differences have become huge: among respondents who were employed at Wave 2 but applied for disability coverage by Wave 3, some six times as many identify their overall health as "poor" as among respondents who stayed employed, and virtually all identify themselves as being limited or disabled. Finally, differences between respondents who stayed employed and those who left employment without applying for disability are in the same direction as for those who apply for disability, but much smaller in magnitude: health differences are quite small at Wave 1, but by Wave 2 and especially by Wave 3 respondents who left employment report somewhat worse general health and much higher prevalence of work limitations.

5.2 Health Dynamics

Our conceptual model suggests that the dynamic relationship between health and labor force behavior should be sensitive to the form of the health trajectories people experience. If health shocks can be expected to be persistent, they will have different effects than if the typical health shock is short-lived. To examine the times-series properties of the typical health shock experience by HRS respondents, we estimated ordered probit models using self-rated health at Wave 3 as the dependent variable and demographics and the proxy measures for once-lagged (Wave 2) and twice-lagged (Wave 1) health as explanatory variables.[15]

Table 3 presents results from these autoregressions. Coefficients have been rescaled so they can be interpreted as partial correlation coefficients. We report chi-squared statistics comparing the model with only the demographic variables to models that first include just once-lagged health and then once-lagged and twice-lagged health. While the coefficients on twice-lagged health are statistically significant, the chi-square statistics indicate that the twice-lagged proxy provides relatively little explanatory power beyond what is contained in the once-lagged proxy.

Figures 1 plots the impulse functions for these models. The plots suggest significant persistence of health shocks: five periods (10 years) after a negative health shock, health will typically still be much worse that one could have expected before the shock. It is worth noting that since we include age as a covariate in all our models, the best way to think about our health variables is as deviations around age- (and sex-) specific means. Thus, the fact that the estimates show reversion to the mean does not imply that individuals' health typically improves after a negative shock, but that it improves relative to the health of other people the same age. (From the point of view of our conceptual model, it would seem as if such deviations around age-specific means is what we want, since people should be able to anticipate the mean age effect.)

Given the left-censored nature of the HRS, we note an important matter for interpretation. Conceptually, we would like to include not just lagged and twice-lagged health measures in our models, but also health measures in the more distant past. Short of such measures, the coefficients on the health variables in the model will, to some extent, reflect forces of past history. It seems plausible that this will be more true of twice-lagged health (the earliest measure we have

[15]The proxies ($\hat{h}_{t_i}{}^*$) were constructed using the methods described above (Equation 8). However, in contrast to Equation 8, in which we use detailed health measures (Z_{t_i}) from all three waves to construct each proxy, for the auto-regressions we do not use Z_{3_i} to construct the Wave 1 and Wave 2 proxies.

available) than of once-lagged or contemporaneous health. Intuitively, health at baseline will serve as a proxy for health in all previous periods.[16]

5.3 Multivariate Models

To examine the patterns in Tables 1 and 2 more formally, and in particular to address the issues of endogeneity and measurement error outlined above, we turn to multivariate models. Parameter estimates for the health proxy equations (Equation 8) are presented in Appendix Table A1.

Tables 4a (men) and 4b (women) present coefficient estimates for the multinomial models (each equation also controls for all the demographic variables described above). We also present Wald (chi-square) test statistics, corresponding to the test that the lagged health coefficients are zero in the respective behavioral equations; that all three health coefficients are zero in each of the respective behavioral equations; that the coefficients on the lagged health measures are zero across all three behavioral equations; that all the health coefficients are zero across all three behavioral equations; and that all health coefficients are zero across the three behavioral equations and the selection equation.

The health proxies are scaled in such a way that higher (more positive) values of the proxies represent worse health. As described above, we try to take account of initial conditions by jointly estimating the retirement equations with a selection equation modeling employment at baseline. The tables present two sets of results: one for which we set the correlation between the error term in the Wave 2 selection equation and the error terms of the respective behavioral equations (r_k) equal to 0, and another where we set each r_k equal to -0.75.[17] Setting r_k equal to 0 is

---

[16]It is easy to see that in a linear model, if the explanatory variables follow an AR(1), excluding high ordered lagged values of these variables from ones estimating equations will effect only the highest ordered lagged values included in the model (Pakes and Griliches, 1984).

[17]Recall that we have imagined that r_k should be negative. Thus r_k =0 represents an upper bound on plausible values for these paremeters. r_k =-0.75 is meant to represent something like a lower bound. Coefficients on the health variables vary monotonically as we vary r_k .

essentially equivalent to ignoring the self-selected nature of the employed population, while r_k =-0.75 is intended to capture fairly strong selection effects. (In addition, we need to specify values of the components of $Var(n)$; when r_k =0, we set $var(n_t)$ and $corr(n_t, n_s), t \neq s$ equal to 0, while when r_k =-0.75, we set $var(n_t)$ and $corr(n_t,n_s)$ equal to 0.75.) Our substantive findings were not generally sensitive to the values of $Var(n)$.)

For men when r_k =0, "poor" contemporaneous (Wave 3) health is a powerful predictor of applying for disability insurance, and also of leaving the labor force without applying for disability benefits. Effects on changing jobs are also positive, but smaller and not statistically significant. Our conceptual framework, and the fact that lagged health follows a pattern that is close to AR(1), suggests that the effects of lagged health should be non-positive in predicting application for disability insurance, and other labor force exit, while the effects on changing jobs are more ambiguous. Results in the table are consistent with these expectations, although the coefficients on once-lagged and twice-lagged health are not significant (at the 0.05 level) for any of the outcomes. With respect to both applying for disability and other labor force exit, the effects of lagged health are negative: controlling for contemporaneous health status, those in good health as of Wave 1 and/or 2 are more likely to choose these outcomes than those in poor health.

It seems plausible that the coefficients on health in these specifications reflect initial conditions, i.e., those who continue to work despite being in poor health as of Wave 2 must have unobserved characteristics that make work relatively attractive to them. As discussed above, we try to take account of such selection by jointly estimating the retirement and Wave 2 employment equations, parametrically varying the error correlations ($Var(n)$). Consistent with our expectations, raising r_k increases the overall magnitude of the effect of health on labor market behavior (based on the chi-square statistic of the joint significance of the health variables across all three behavioral equations). Adjusting for initial conditions by setting r_k =-0.75 attenuates the effect of contemporaneous health on each of the outcomes. Poor contemporaneous health does not appear to have much of an impact on whether or not individuals change jobs but remains a significant predictor of both types of labor force exit. When r_k =-0.75, lagged health becomes a significant predictor of changing jobs: worse health in Wave 2 raises the probability of changing jobs by Wave 3. For both types of labor force exit, the coefficients on once-lagged health are closer to zero (and statistically insignificant).[18] The effects of lagged health are not significant for the other two outcomes, but the point estimates change sign and are thus the same sign as the effects of contemporaneous health. Setting r_k=-0.75 attenuates but does not eliminate the effect of twice-lagged health, but the point estimates are still negative.

It is possible to translate the estimated effects of current and lagged health on behavior into the effects of baseline health and health changes. Thus, our results indicate that, when we set r_k=0, changes in health have a larger effect on behavior than does baseline health. In contrast, setting r_k=-0.75 tends to increase the relative importance of baseline health. Given that we interpret the two values of r_k as bounds, it is impossible to know where between these extremes the truth lies.

Results for women are again similar to those for men. When r_k=0, contemporaneous health is positively associated with each of the labor force outcomes. Once-lagged health is positively associated with changing jobs, negatively associated with applying for DI, and has no effect on non-DI labor force exit. Twice-lagged health is negatively associated with each outcome. Setting r_k=-0.75 attenuates the effects of contemporaneous and twice-lagged health but does not change the respective direction of these effects. Once-lagged health, however, becomes positively associated with each outcome.[19]

Interpreting the magnitude of the coefficients on the health variables is difficult. To aid interpretation, Table 5 presents the results of a number of simulations based on the estimates in Table 4. As a baseline, we simulate the probability of working as of Wave 2, and of the respective labor force outcomes as of Wave 3 conditional on having been working as of Wave 2, for each

---

[18]For men, the values of r at which the coefficient on lagged health hit zero was around -0.10 for changing jobs, -0.55 for applying for DI, and –0.65 for exiting without applying for DI.

[19]For women, the values of r_k at which the coefficient on lagged health hit zero was positive for changing jobs, around –0.55 for applying for DI, and around zero for exiting without applying for DI

individual in our data set. We then calculate the simple average of the probability of Wave 2 employment and a weighted average of the Wave 3 outcomes, using the probability of Wave 2 employment as the weight (this weighting is conceptually and substantively very similar to simulating the Wave 3 labor force outcomes only for respondents who were employed at Wave 2 and taking the simple average across this restricted sample). These predicted probabilities resemble but are not identical to the actual sample frequencies (this is a characteristic of probit models).

Next, we simulate behavior under the counterfactual assumptions that everyone in the sample was in good health at all three interviews, was in good health as of the first interview and poor health during the second and third, etc. To do this, we first calculated the average value of $\hat{h}^*_{1_i}$ for those who report "good" (excellent, very good or good) health and for those in "poor" (fair or poor health) health as of Wave 1, which we can refer to as $\hat{h}^*_{g_1}$ and $\hat{h}^*_{p_1}$. Note that these quantities do not vary across the population (all models are sex-specific). We did the same calculation for $\hat{h}^*_{2_i}$ and $\hat{h}^*_{3_i}$. For the simulations, we then replaced $\hat{h}^*_{t_i}$ with $\hat{h}^*_{t_g}$ and $\hat{h}^*_{t_p}$, respectively, as designated in the table.

Table 5 indicates that health is a very strong predictor of labor force behavior. When r_k =0, the vast majority of men in consistently good health remained in the labor market, while those who exited did so without applying for disability insurance. Of men whose health was good through Wave 2 and then declined (Column 2), only 30 percent remained in the labor force, and of those who left, the majority applied for disability insurance. Compared with this group, twice as many men whose health was good through Wave 1 and then declined (Column 3) remained in the labor force. Still, about half of the men in Column 3 who exited the labor force between Waves 2 and 3 applied for DI. Finally, the large majority of men whose health declined prior to Wave 1 (Column 4) remained in the labor force (although fewer did so than among men in consistently good health), and only a small fraction applied for DI. These results are consistent with the hypothesis that people adapt to relatively early health shocks in ways that enable continued labor force participation.

However, it is also interesting to note patterns in job change. In absolute terms, the fraction of men who change jobs is lower for men in Column 2 than for Columns 1, 3 or 4; while in percentage terms, the fraction of men who change jobs is higher for men in Columns 2 and 3 than for men in Columns 1 and 4. This suggests that in the short run, health shocks may inhibit job change, while in the medium run, job change may be an important mechanism by which people adapt to health shocks to enable them to remain in the work place.

Comparing Column 1 with Columns 3 and 4 provides evidence on the delayed, as opposed to immediate, effects of health shocks. Compared with Column 1, men whose health declined between Waves 1 and 2 but who remained employed at Wave 2 were much less likely to still be in the same job by Wave 3 and much more likely to apply for DI or otherwise exit the labor force. However, patterns among men whose health declined prior to Wave 1 are very similar to those among men in persistently good health.

The r_k =0 results are descriptively valid regardless of how important selection is – i.e., among those working as of Wave 2, those who experience health declines between Wave 2 and Wave 3 are more likely to leave the workforce than are those who experienced persistently poor health. However, these patterns reflect the causal effect of health on behavior only if selection is unimportant. If selection is important, then some of this difference reflects unobserved differences between these two groups – unobserved differences that were generated by the endogenous selection into the working sample. In the second panel of table 5 we present results for the model with r_k =-0.75 to illustrates the potential importance of selection.

Interpreting simulations when r_k =-0.75 requires more care than does interpretation of the results that ignore sample selection. We are interested in understanding the causal effect of health and changes in health on behavior. To get at this, we use our models to simulate the effect of health and health change on behavior. Conceptually, what we are doing is fixing health at specified values and letting both observed and unobserved factors in the population vary. Thus the simulations reflect the behavior we could expect of someone who had specified health characteristics, observable characteristics that are typical of those working as of Wave 2, and unobserved characteristics that are typical of the entire population. The simulated probabilities will not line up very closely with the actual sample probabilities because the sample probabilities reflect a non-random, rather than random sample of unobserved characteristics.

Differences between men with earlier versus later onset of poor health remain when we set r_k =-0.75. As before, applying for disability insurance is an important outcome among all the groups who suffered a negative health shock. Changing jobs becomes a more important alternative in all columns. Among men whose health was good through Wave 2 and then declined, barely a quarter remained in the work force by Wave 3 – and the majority of these changed jobs. The remainder were approximately evenly split between men who applied for DI and men who exited the labor force without doing so. Among men whose health declined between Waves 1 and 2, somewhat more were still working by Wave 3 relative to men with later onset of poor health (Column 2). However, almost all those who remained in the labor force by Wave 3 changed jobs. Compared with Columns 2 and 3, a higher fraction of men whose health shock occurred prior to Wave 1 remained employed at Wave 3. A large absolute number of these men changed jobs. However, job change as a fraction of all employed men fell; given the patterns in Columns 2 and 3, this suggests that many of these men had already changed jobs prior to Wave 3. These findings reinforce the idea that job change may be an important mechanism by which people adapt to health shocks to enable them to remain in the work place – and those job changes don't occur immediately.

Results for women are broadly similar to those for men. When r_k =0, women with the most recent health shock (Column 2) were less likely to remain employed and to change jobs, and much more likely to apply for DI, than women whose health had declined between Waves 1 and 2 (Column 3). Women in Column 3, in turn, were less likely to remain employed and more likely to apply for DI than women whose health declined prior to Wave 1 (Column 4). However, they were much more likely to change jobs. When r_k =-0.75, disability insurance and job change become more important outcomes for all groups. Among women whose health declined between Waves 2 and 3, 70 percent of those who stayed employed changed jobs. However, among those whose health declined earlier, virtually all who stayed employed at Wave 3 changed jobs, again suggesting that job change – in addition to DI application – are important responses to the onset of poor health.

For comparison with these dynamic models, we reestimated our multivariate models omitting lagged and twice-lagged health as explanatory variables. Overall findings were consistent with those in Tables 4 and 5 (although the simpler models made them easier to interpret). As in our previous models, men and women in poor contemporaneous health were more likely to leave the labor force without applying for DI and much more likely to apply for DI than were respondents in good current health (both effects were statistically significant), while current health had relatively little effect on the likelihood of changing jobs. Also as before, the direction of these effects was generally unaffected by varying r_k from 0 to -0.75, though the magnitude increases as the magnitude of r_k is increased. Coefficients and significance tests are presented in Appendix Tables 2a (men) and 2b (women), which are analogous to Tables 4a and 4b.

6. Discussion and Conclusions

The results we have presented confirm that health is a very important determinant of labor force patterns for older men and women. Poor health leads many older workers to withdraw from the labor force. Among people in poor health, more than half of those who exit the labor force apply for DI. Among those who keep working, many change jobs within several years of the onset of their poor health, suggesting that changing jobs is an important way that older workers adapt to enable continued labor force participation. Together, these results confirm the value of modeling alternative labor force outcomes, beyond the binary outcome of labor force withdrawal.

The high rate of labor force exit among people in poor health raises the question of what people live on when they leave the labor force. Although many people apply for DI, many do not (and of course not all those who apply for DI receive it). In preliminary analysis using the HRS, we found that people with a health-related work limitation have significantly lower household income, on average, than respondents who are not employed but report no such limitation. They also receive a much higher fraction of their income from public transfer programs, and a lower fraction from earnings (presumably representing spouses' earnings), assets and, among men, private pensions. This issue should be the subject of future research.

Our findings also suggest that the relationship between health and labor force behavior is dynamic. Respondents whose health declined relatively recently are more likely to apply for DI or to leave the labor force without applying for DI, and are less likely to change jobs, than are respondents whose health declined earlier. Overall, the earlier a health shock occurs in our models, the less likely it is to lead to labor force exit. In terms of lagged effects, our models suggest that lagged health affects behavior even controlling for contemporaneous health status. These are issues that obviously cannot be examined without controlling for both contemporaneous and lagged values of health.

While our findings suggest that the effect of health on labor force behavior is dynamic, precisely estimating the effect of changes in health status on behavior is difficult for two distinct reasons. First, while the direction of the effects of lagged and contemporaneous health on alternative labor force transitions does not depend on how we account for initial conditions, the magnitude does. In terms of the bounds we use in our models, lagged effects matter more at one extreme ($r_k = 0$) than at the other ($r_k = -0.75$). Establishing how much behavior depends on the level of health, and how much on the trajectory, requires researchers to credibly control for initial conditions. Doing so is difficult, and we leave this to future research. Second, even conditional on a choice of r_k, the coefficients on the lagged health variables were imprecisely estimated and in some specifications were not jointly statistically significant at conventional levels. Presumably this has to do with the high multicollinearity between our measures of contemporaneous, lagged, and twice-lagged health.

BIBLIOGRAPHY

1. Anderson, K, and R. Burkhauser (1984). The Importance of the Measure of Health in Empirical Estimates of the Labor Supply of Older Men. *Economics Letters* 16:375-80.
2. Anderson, K.H. and R.V. Burkhauser, (1985) "The Retirement-Health Nexus: A New Measure of an Old Puzzle." *Journal of Human Resources*, 20:315-330.
3. Anderson, K., R. Burkhauser and J.F. Quinn, (1986) Do Retirement Dreams Come True? The Effect of Unanticipated Events on Retirement Plans. *Industrial Labor Relations Review* 39:518-526.
4. Baltes, P.B., and Baltes, M.M. (1990). Psychological perspectives on successful aging: a model of selective optimization with compensation. In P.B. Baltes and M.M. Baltes (Eds.), *Successful Aging: Perspectives from the Behavioral Sciences* (pp. 1-34) New York: Cambridge University Press.
5. Blau, D.M., D.B. Gilleskie and C. Slusher (1997). The effects of health on employment transitions of older men. Presented at the 1997 annual meeting of the Population Association of America, Washington DC.
6. Blundell, R., MacCurdy, T. "Labor Supply" in Orley Ashenfelter and David Card Ed, *Handbook of Labor Economics*, Vol 3, forthcoming.
7. Bound, J. (1991). Self-reported versus objective measures of health in retirement models. The *Journal of Human Resources*, 26, 106-138.
8. Bound, J., Schoenbaum, M. and Waidmann, T. (1996). Race Differences in Labor Force Attachment and Disability Status. *The Gerontologist* 36:311-321
9. Bound, J., Schoenbaum, M, Waidmann, T., and Stinebrickner, T (1998). Modeling the Effects of Health on Retirement. To be presented at the 1998 NBER Summer Institute Workshop on Aging and Healthcare, Boston.
10. Brim, O.G. (1988). Losing and wining. *Psychology Today* 48-52.
11. Charles, K.K. (1997). *The Longitudinal Structure of Earnings Losses Among Work-Limited Disabled Workers: Age of Onset Effects and Adjustment*. University of Michigan, 1997.
12. Daly, M. and J. Bound (1996). The Adaptation of Workers to the Onset of Health Limitations. *Journal of Gerontology: Social Sciences*, 51B:S53-S60
13. Edwards, W.S., Winn, D.M., Kurlantzick, V., Sheridan, S., Berk, M.L., Retchin, S., & Collins, J.G. (1994). Evaluation of National Health Interview Survey Diagnostic Reporting. *Vital and Health Statistics*, 2(120).
14. Fuller, W.A. (1987). *Measurement Error Models*. New York: Wiley.
15. Geweke, J. (1988). Antithetic Acceleration of Monte Carlo Integration in Bayesian Inference. *Journal of Econometrics*. 38:73-89.
16. Geweke, J. (1991). Efficient Simulation from the Multivariate Normal and Student-t Distributions Subject to Linear Constraints. Computer Science and Statistics: Proceedings of the Twenty-Third Symposium on the Interface. 571-578.
17. Griliches, Z. (1974). Errors in variables and other unobservables. *Econometrica* 42:971-998.
18. Hajivassiliou, V. (1990): Smooth Simulation Estimation of Panel Data LDV Models. Unpublished Manuscript.
19. Honig, M. (1985) Partial Retirement Among Women. *Journal of Human Resources*. 20:613-621.

20. Honig, M and Hanoch, G. (1985). Partial Retirement as a Separate Mode of Retirement Behavior. *Journal of Human Resources*. 20:21-46.
21. Honig, M., and Reimers,-C. (1987). Retirement, Re-entry, and Part-time Work. *Eastern-Economic-Journal*; 13(4), October-December 1987, pages 361-71.
22. House, J.S., Kessler, R.C., Herzog, A.R., Mero, R.P., Kinney, A.M, and M.J. Breslow (1990). Age, Socioeconomic Status, and Health. *The Milbank Quarterly* 68(3): 383-411
23. Juster, FT, Suzman, R. (1995). An overview of the Health and Retirement Survey. *J Human Resources* 30(Supplement): S7-S56.
24. Keane, MP. (1992) A Note on Identification in the Multinomial Probit Model. *Journal-of-Business-and-Economic-Statistics*; 10(2): 193-200.
25. Keane, MP. (1994): A computationally Practical Simulation Estimator for Panel Data, *Econometrica* 62(1), 95-116.
26. Leamer, E.E. (1978) *Specification Searches : Ad Hoc Inference with Nonexperimental Data* New York: Wiley, 1978.
27. Mathiowetz, N.A. and Laird, T.J. (1994). Getting Better? Change or error in the measurement of functional limitations. *Journal of Economic and Social Measurement* 20: 237-262.
28. Myers, R.J. (1982). Why do people retire early from work? *Aging and Work*, 5:83-91.
29. Pakes, A, Griliches Z. (1984). Estimating distributed lags in short panels with an application to the specification of depreciation patterns and capital stock constructs. *Review of Economic Studies* 51:243-62.
30. Parsons, D. O. (1982). The male labor force participation decision: *Health*, reported health, and economic incentives. Economica, 49, 81-91.
31. Quinn, J.F. (1998). New Paths to Retirement. *Boston College Working Paper 406*.
32. Quinn, J.F. (1997). The Role of Bridge Jobs in the Retirement Patterns of Older Americans in the 1990s. In *Retirement Prospects in a Defined Contribution World*, D. Salisbury, Ed., Employee Benefit Research Institute, 1997, pp 25-39.
33. Quinn, J.F., R.V. Burkhauser and D. Myers, (1990*) Passing the Torch: The Influence of Economic Incentives on Work and Retirement*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.
34. Radloff, L.S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological. Measurement*, 1:385
35. Ruhm, C.J. (1990). Bridge Jobs and Partial Retirement. *Journal of Labor Economics*. 8:482-501.
36. Stern, S. (1989) Measuring the effect of disability on labor force participation. *Journal of Human Resources*, Summer.
37. Suurmeijer, T.P.B.M., D.M. Doeglas, T. Moum, S. Briangon, B. Krol, R. Sanderman, F. Guillemin, A. Bjelle, and W.J.A. van den Heuvel (1994). The Groningen Activity Restriction Scale for Measuring Disability: Its Utility in International Comparisons. *American Journal of Public Health* 84(8):1270-1273
38. Waidmann, T., Bound, J. and Schoenbaum, M. (1995). The illusion of failure: Trends in the self-reported health of the U.S. elderly. *Milbank Quarterly* 73(2): 253-287.

Table 1: Mean demographic and health characteristics, by Wave 2 labor force status

	Men		Women	
	Employed	**Not employed**	**Employed**	**Not employed**
N	2115	760	2154	1672
Fraction of sample	0.74	0.26	0.56	0.44
Age	55.50	56.97	55.34	56.44
Education				
<HS	0.15	0.25	0.15	0.28
HS grad	0.36	0.40	0.44	0.45
Some college	0.22	0.18	0.22	0.17
College grad	0.27	0.17	0.19	0.10
Married/partner	0.88	0.76	0.68	0.76
Race/ethnicity				
Non-Hispanic White	0.87	0.79	0.82	0.79
Non-Hispanic Black	0.07	0.14	0.11	0.12
Hispanic	0.04	0.06	0.04	0.07
Other	0.02	0.02	0.02	0.01
US-born	0.92	0.95	0.93	0.92
Veteran	0.58	0.59	0.01	0.01
Fair/poor health status	0.09	0.36	0.10	0.32
Limited/disabled	0.08	0.43	0.09	0.33

Source: Authors tabulations using HRS data.
Sample: Age-eligible men and women responding for selves in waves 1, 2 and 3 with no missing health or demographic information.

Table 2: Mean demographic and health characteristics, by Wave 3 labor force status
Sample: respondents employed at Wave 2

	Men				Women			
	Same Job	**Diff. Job**	**Disability**	**Not employed**	**Same Job**	**Diff. Job**	**Disability**	**Not employed**
N	1394	393	58	258	1385	320	65	377
Fraction of sample	0.66	0.19	0.03	0.12	0.65	0.15	0.03	0.18
Age*	55.26	55.13	54.83	57.66	55.11	54.76	55.42	56.63
Education								
<HS	0.15	0.10	0.30	0.20	0.14	0.11	0.45	0.18
HS grad	0.35	0.35	0.34	0.38	0.44	0.43	0.36	0.43
Some college	0.21	0.25	0.25	0.20	0.21	0.31	0.15	0.22
College grad	0.28	0.30	0.10	0.21	0.20	0.16	0.03	0.18
Married/partner*	0.89	0.90	0.86	0.83	0.68	0.61	0.44	0.77
Race/ethnicity								
Non-Hispanic White	0.86	0.90	0.82	0.85	0.82	0.84	0.73	0.84
Non-Hispanic Black	0.08	0.04	0.07	0.08	0.11	0.11	0.20	0.10
Hispanic	0.04	0.04	0.09	0.06	0.04	0.04	0.03	0.04
Other	0.02	0.02	0.01	0.02	0.02	0.00	0.03	0.01
US-born	0.92	0.94	0.87	0.91	0.92	0.95	0.94	0.93
Veteran	0.56	0.61	0.55	0.60	0.01	0.01	0.00	0.01
Fair/poor health status								
Wave 1	0.08	0.07	0.30	0.14	0.09	0.10	0.34	0.13
Wave 2	0.08	0.08	0.28	0.18	0.09	0.11	0.50	0.12
Wave 3	0.10	0.09	0.64	0.19	0.09	0.12	0.71	0.17
Limited/disabled								
Wave 1	0.08	0.05	0.23	0.11	0.08	0.11	0.22	0.10
Wave 2	0.08	0.08	0.33	0.14	0.06	0.12	0.31	0.14
Wave 3	0.10	0.08	0.97	0.23	0.09	0.16	0.94	0.25

***Note:** value at Wave 1
Source: Authors tabulations using HRS data.

Table 3: Autoregression of Wave 3 health on lagged and twice-lagged health
Sample: whole study sample

	Women		Men	
	Estimated β	***(SE)***	**Estimated β**	***(SE)***
Wave 2 (lagged) health	0.711	*(0.070)*	0.765	*(0.084)*
Wave 1 (twice-lagged) health	0.208	*(0.071)*	0.159	*(0.081)*
N	3826		2875	
χ^2 for W2 health (DF=1)	1295.6		876.8	
χ^2 for W2 and W1 health (DF=2)	1304.1		880.6	

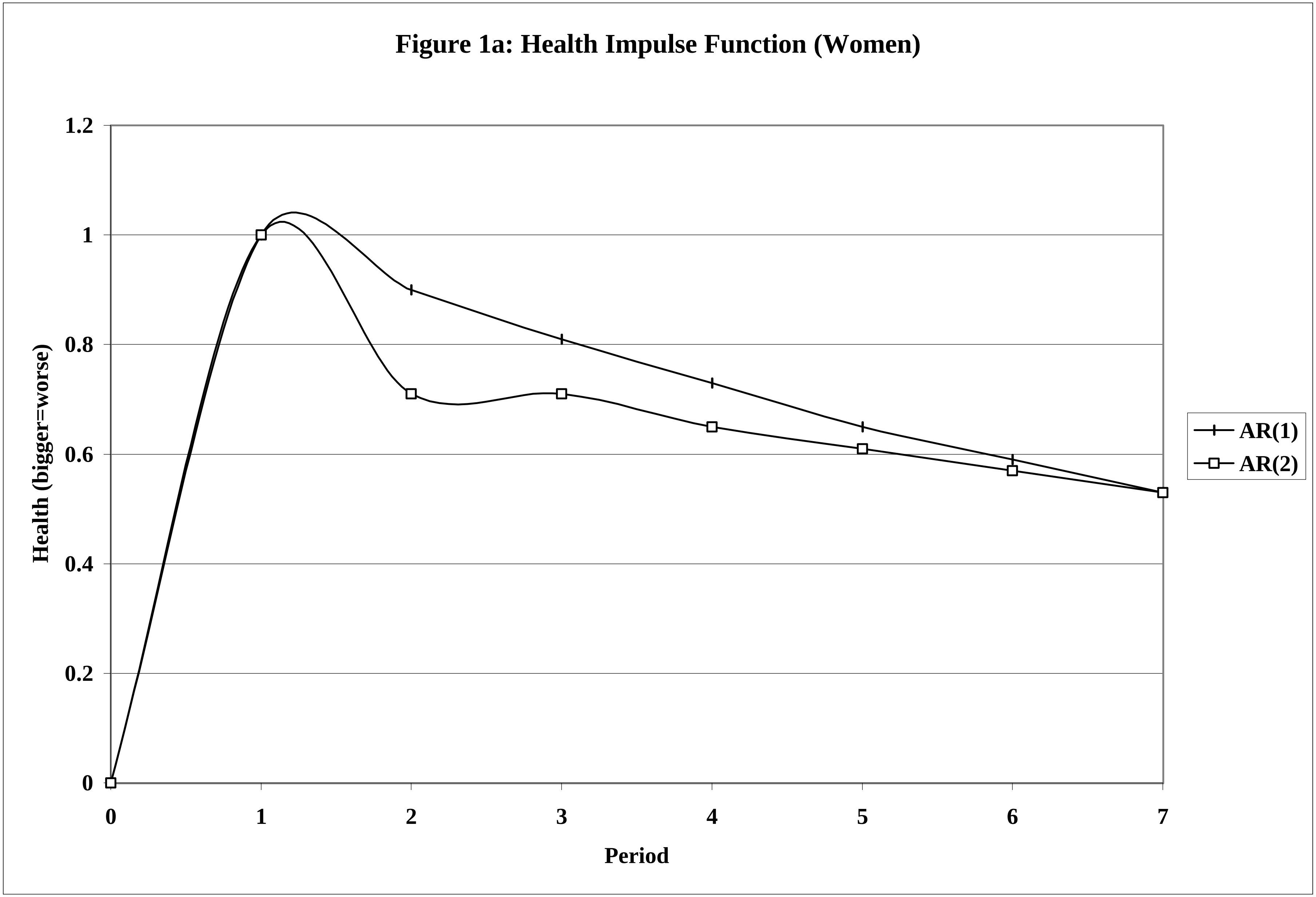
Figure 1a: Health Impulse Function (Women)
1.2
1
0.8
0.6
0.4
0.2
0
Health (bigger=worse)
0
1
2
3
4
5
6
7
Period
AR(1)
AR(2)

Table 4a: Multinomial probit estimates of labor force transitions by Wave 3, among respondents employed at Wave 2
Comparison group: respondents employed in the same job in Waves 2 and 3
Men (N=2115)

Selection Parameter	$\rho=0$[a]			$\rho=-0.75$[a]		
	Diff. Job	**Apply for DI**	**Not emp., not apply for DI**	**Diff. Job**	**Apply for DI**	**Not emp., not apply for DI**
	Coeff. ***(SE)***	**Coeff.** ***(SE)***	**Coeff.** ***(SE)***	**Coeff.** ***(SE)***	**Coeff.** ***(SE)***	**Coeff.** ***(SE)***
twice-lagged health	-0.20	-0.61	-0.32	-0.18	-0.39	-0.24
(Wave 1)	*(0.18)*	*(0.34)*	*(0.19)*	*(0.16)*	*(0.22)*	*(0.17)*
once-lagged health	-0.03	-0.47	-0.42	0.44	0.16	0.07
(Wave 2)	*(0.24)*	*(0.37)*	*(0.27)*	*(0.21)*	*(0.27)*	*(0.22)*
contemporaneous health	0.27	1.83	0.91	0.05	1.05	0.61
(Wave 3)	*(0.20)*	*(0.28)*	*(0.20)*	*(0.18)*	*(0.18)*	*(0.18)*
χ^2**, lagged health vars (DF=2)**	1.87	13.10	10.96	4.56	3.37	2.33
χ^2**, all health vars (DF=3)**	2.75	62.26	28.81	24.17	292.95	44.51
χ^2**, lagged health vars, behavior eqs (DF=6)**		17.29			8.48	
χ^2**, all health vars, behavior eqs (DF=9)**		67.68			331.21	
χ^2**, all health vars, incl. select. eq. (DF=12)**		256.11			498.23	
Log likelihood		-12744.08			-12753.02	

[a]**Note:** Represents correlation between error terms of behavioral & selection equations
Source: Author's calculations using Health and Retirement Survey data
Sample: Age-eligible men and women responding for selves in waves 1-3 with no missing health/demographic data

Table 4b: Multinomial probit estimates of labor force transitions by Wave 3, among respondents employed at Wave 2
Comparison group: respondents employed in the same job in Waves 2 and 3
Women (N=2154)

Selection Parameter	ρ=0[a]			ρ=–0.75[a]		
	Diff. Job	**Apply for DI**	**Not emp., not apply for DI**	**Diff. Job**	**Apply for DI**	**Not emp., not apply for DI**
	Coeff. ***(SE)***	**Coeff.** ***(SE)***	**Coeff.** ***(SE)***	**Coeff.** ***(SE)***	**Coeff.** ***(SE)***	**Coeff.** ***(SE)***
twice-lagged health	-0.50	-0.55	-0.28	-0.39	-0.38	-0.20
(Wave 1)	*(0.19)*	*(0.28)*	*(0.17)*	*(0.16)*	*(0.21)*	*(0.15)*
once-lagged health	0.31	-0.24	-0.01	0.63	0.17	0.37
(Wave 2)	*(0.23)*	*(0.34)*	*(0.20)*	*(0.19)*	*(0.26)*	*(0.17)*
contemporaneous health	0.27	1.58	0.47	0.08	1.04	0.24
(Wave 3)	*(0.14)*	*(0.22)*	*(0.13)*	*(0.13)*	*(0.18)*	*(0.12)*
χ^2**, lagged health vars (DF=2)**	7.57	14.04	5.74	11.48	3.96	4.78
χ^2**, all health vars (DF=3)**	12.95	91.09	27.91	38.65	263.69	67.12
χ^2**, lagged health vars, behavior eqs (DF=6)**		20.07			18.23	
χ^2**, all health vars, behavior eqs (DF=9)**		98.66			307.97	
χ^2**, all health vars, incl. select. eq. (DF=12)**		334.02			502.68	
Log likelihood		-16697.08			-16703.74	

[a]**Note:** Represents correlation between error terms of behavioral & selection equations
Source: Author's calculations using Health and Retirement Survey data
Sample: Age-eligible men and women responding for selves in waves 1-3 with no missing health/demographic data

Table 5: Simulated effects of health status on employment at Wave 3, among respondents employed at Wave 2 (weighted for selection). Wave 1-3 health enters behavioral equation.

Twice-lagged health (Wave 1): Once-lagged health (Wave 2): Contemporaneous health (Wave 3):	Actual Actual Actual (0)	Good Good Good (1)	Good Good Poor (2)	Good Poor Poor (3)	Poor Poor Poor (4)
Men (ρ=0)					
Employed at Wave 2 (selection)	*0.73*	*0.80*	*0.92*	*0.53*	*0.44*
Different job	0.19	0.19	0.09	0.18	0.17
Not employed, not app. for disability	0.12	0.13	0.32	0.20	0.13
Applied for disability	0.02	0.00	0.38	0.21	0.05
Same job	0.66	0.68	0.21	0.42	0.65
Men (ρ=-0.75)					
Employed at Wave 2 (selection)	*0.73*	*0.80*	*0.91*	*0.52*	*0.44*
Different job	0.23	0.24	0.14	0.31	0.33
Not employed, not app. for disability	0.17	0.17	0.39	0.30	0.27
Applied for disability	0.05	0.02	0.35	0.35	0.20
Same job	0.54	0.57	0.12	0.04	0.19
Women (ρ=0)					
Employed at Wave 2 (selection)	*0.56*	*0.63*	*0.78*	*0.35*	*0.32*
Different job	0.15	0.15	0.09	0.29	0.15
Not employed, not app. for disability	0.18	0.18	0.24	0.22	0.22
Applied for disability	0.03	0.00	0.43	0.23	0.10
Same job	0.65	0.67	0.24	0.26	0.53
Women (ρ=-0.75)					
Employed at Wave 2 (selection)	*0.56*	*0.63*	*0.78*	*0.34*	*0.32*
Different job	0.25	0.26	0.15	0.38	0.31
Not employed, not app. for disability	0.30	0.31	0.30	0.30	0.40
Applied for disability	0.06	0.02	0.49	0.31	0.26
Same job	0.39	0.41	0.06	0.00	0.04

Source: Authors' tabulations using estimates from Tables 4a and 4b.
Sample: Age-eligible men and women responding for selves in waves 1-3 with no missing health or demographic data

Appendix Table A1: Construction of health indices (ordered probit models)
Dependent variable: self-rated general health (5 categories)
Sample: whole study sample

	Women						**Men**					
	Wave 1		**Wave 2**		**Wave 3**		**Wave 1**		**Wave 2**		**Wave 3**	
			N=3826						**N=2875**			
Wave 1 functional limitations												
Jog one mile	0.09	*(0.02)*	0.06	*(0.02)*	0.04	*(0.02)*	0.22	*(0.03)*	0.12	*(0.03)*	0.11	*(0.03)*
Walk several blocks	0.12	*(0.04)*	0.07	*(0.04)*	0.00	*(0.04)*	0.23	*(0.05)*	0.07	*(0.05)*	0.08	*(0.05)*
Walk one block	0.11	*(0.05)*	0.04	*(0.05)*	0.05	*(0.05)*	-0.03	*(0.07)*	0.04	*(0.07)*	-0.04	*(0.07)*
Walk across room	-0.10	*(0.06)*	-0.11	*(0.06)*	0.01	*(0.06)*	-0.05	*(0.09)*	-0.03	*(0.09)*	-0.06	*(0.09)*
Sit two hours	-0.02	*(0.03)*	0.06	*(0.03)*	0.01	*(0.03)*	-0.01	*(0.03)*	-0.04	*(0.03)*	-0.02	*(0.03)*
Stand after sitting	0.03	*(0.03)*	-0.06	*(0.03)*	0.00	*(0.03)*	-0.02	*(0.04)*	-0.05	*(0.04)*	-0.10	*(0.04)*
Get in and out of bed	0.05	*(0.04)*	0.03	*(0.04)*	-0.07	*(0.04)*	-0.05	*(0.06)*	-0.12	*(0.06)*	-0.03	*(0.06)*
Go up several flights of stairs	0.12	*(0.03)*	0.10	*(0.03)*	0.08	*(0.03)*	0.16	*(0.04)*	0.07	*(0.04)*	0.13	*(0.04)*
Go up one flight of stairs	0.07	*(0.04)*	-0.03	*(0.04)*	-0.01	*(0.04)*	0.07	*(0.06)*	0.03	*(0.06)*	-0.04	*(0.06)*
Lift 10 pounds	0.13	*(0.03)*	0.02	*(0.03)*	0.02	*(0.03)*	0.12	*(0.05)*	0.11	*(0.05)*	-0.02	*(0.05)*
Kneel or stoop	0.06	*(0.03)*	0.01	*(0.03)*	-0.05	*(0.03)*	0.04	*(0.04)*	0.02	*(0.04)*	0.00	*(0.04)*
Pick up dime	0.06	*(0.04)*	-0.01	*(0.04)*	0.02	*(0.04)*	0.06	*(0.05)*	0.11	*(0.05)*	-0.03	*(0.05)*
Reach above shoulders	0.04	*(0.04)*	0.03	*(0.04)*	0.04	*(0.04)*	0.12	*(0.05)*	0.08	*(0.05)*	0.08	*(0.05)*
Push large objects	0.08	*(0.03)*	0.05	*(0.03)*	0.03	*(0.03)*	0.16	*(0.05)*	0.07	*(0.05)*	0.09	*(0.05)*
Bathe without help	0.16	*(0.06)*	0.00	*(0.06)*	-0.02	*(0.06)*	-0.08	*(0.09)*	-0.24	*(0.09)*	-0.18	*(0.09)*
Eat without help	-0.08	*(0.09)*	0.00	*(0.09)*	-0.15	*(0.09)*	-0.20	*(0.12)*	0.01	*(0.12)*	0.04	*(0.12)*
Dress without help	-0.01	*(0.07)*	-0.05	*(0.07)*	0.00	*(0.07)*	0.14	*(0.10)*	-0.08	*(0.09)*	0.10	*(0.09)*
Wave 2 functional limitations												
Jog one mile	0.06	*(0.03)*	0.14	*(0.03)*	0.08	*(0.03)*	0.08	*(0.03)*	0.15	*(0.03)*	0.10	*(0.03)*
Walk several blocks	-0.02	*(0.04)*	0.11	*(0.04)*	0.02	*(0.04)*	-0.01	*(0.06)*	0.07	*(0.06)*	-0.04	*(0.05)*
Walk one block	0.07	*(0.06)*	0.03	*(0.06)*	0.02	*(0.06)*	-0.13	*(0.08)*	-0.09	*(0.08)*	-0.05	*(0.08)*
Walk across room	-0.05	*(0.07)*	0.08	*(0.07)*	0.03	*(0.07)*	0.27	*(0.10)*	0.19	*(0.10)*	0.03	*(0.10)*
Sit two hours	0.02	*(0.03)*	0.07	*(0.03)*	0.00	*(0.03)*	-0.02	*(0.04)*	0.03	*(0.04)*	-0.01	*(0.04)*
Stand after sitting	0.00	*(0.03)*	0.06	*(0.03)*	0.02	*(0.03)*	0.08	*(0.04)*	0.09	*(0.04)*	0.02	*(0.04)*
Get in and out of bed	0.08	*(0.05)*	0.07	*(0.05)*	-0.01	*(0.05)*	-0.08	*(0.06)*	0.08	*(0.06)*	0.06	*(0.06)*
Go up several flights of stairs	0.11	*(0.03)*	0.13	*(0.03)*	0.13	*(0.03)*	0.11	*(0.04)*	0.22	*(0.04)*	0.12	*(0.04)*
Go up one flight of stairs	0.00	*(0.04)*	0.13	*(0.04)*	0.01	*(0.04)*	-0.02	*(0.07)*	0.01	*(0.06)*	0.02	*(0.06)*
Lift 10 pounds	0.06	*(0.03)*	0.09	*(0.03)*	0.05	*(0.03)*	0.00	*(0.05)*	0.06	*(0.05)*	0.03	*(0.05)*
Kneel or stoop	-0.01	*(0.03)*	0.01	*(0.03)*	-0.03	*(0.03)*	-0.06	*(0.04)*	0.02	*(0.04)*	0.04	*(0.04)*
Pick up dime	0.07	*(0.04)*	0.07	*(0.04)*	0.05	*(0.04)*	0.09	*(0.06)*	0.05	*(0.06)*	0.06	*(0.06)*
Reach above shoulders	0.08	*(0.04)*	0.15	*(0.04)*	0.03	*(0.04)*	0.12	*(0.05)*	0.08	*(0.05)*	0.05	*(0.05)*
Push large objects	0.05	*(0.03)*	0.12	*(0.03)*	0.07	*(0.03)*	0.05	*(0.05)*	0.18	*(0.05)*	0.08	*(0.05)*
Bathe without help	0.00	*(0.06)*	0.00	*(0.06)*	-0.01	*(0.06)*	0.19	*(0.10)*	0.15	*(0.10)*	0.09	*(0.10)*
Eat without help	-0.23	*(0.18)*	-0.11	*(0.20)*	-0.14	*(0.19)*	0.00	*(0.15)*	-0.18	*(0.14)*	-0.15	*(0.14)*
Dress without help	-0.05	*(0.07)*	0.01	*(0.07)*	-0.10	*(0.07)*	-0.20	*(0.10)*	0.06	*(0.10)*	-0.09	*(0.09)*
Wave 3 functional limitations												
Jog one mile	0.01	*(0.02)*	0.03	*(0.02)*	0.00	*(0.02)*	0.06	*(0.02)*	0.05	*(0.02)*	0.09	*(0.02)*
Walk several blocks	0.06	*(0.03)*	0.05	*(0.03)*	0.16	*(0.03)*	0.10	*(0.05)*	0.07	*(0.05)*	0.22	*(0.05)*
Walk one block	-0.03	*(0.04)*	0.01	*(0.04)*	0.15	*(0.04)*	0.01	*(0.06)*	0.11	*(0.06)*	-0.04	*(0.06)*
Walk across room	-0.12	*(0.05)*	-0.03	*(0.05)*	0.01	*(0.05)*	-0.12	*(0.07)*	-0.27	*(0.07)*	0.04	*(0.07)*
Sit two hours	0.06	*(0.03)*	0.06	*(0.03)*	0.09	*(0.03)*	0.02	*(0.04)*	0.01	*(0.04)*	0.07	*(0.04)*
Stand after sitting	-0.01	*(0.03)*	0.01	*(0.03)*	0.08	*(0.03)*	0.02	*(0.04)*	0.02	*(0.04)*	0.12	*(0.04)*
Get in and out of bed	0.03	*(0.04)*	0.08	*(0.04)*	0.18	*(0.04)*	0.06	*(0.06)*	0.09	*(0.06)*	0.06	*(0.06)*
Go up several flights of stairs	0.09	*(0.03)*	0.09	*(0.03)*	0.16	*(0.03)*	0.06	*(0.03)*	0.08	*(0.03)*	0.15	*(0.03)*
Go up one flight of stairs	0.04	*(0.04)*	0.04	*(0.04)*	0.08	*(0.04)*	-0.03	*(0.05)*	0.00	*(0.05)*	0.07	*(0.05)*
Lift 10 pounds	0.05	*(0.03)*	0.01	*(0.03)*	0.14	*(0.03)*	0.03	*(0.05)*	0.00	*(0.05)*	0.12	*(0.05)*
Kneel or stoop	-0.02	*(0.03)*	0.00	*(0.03)*	0.07	*(0.03)*	-0.01	*(0.04)*	0.01	*(0.03)*	-0.02	*(0.03)*
Pick up dime	-0.02	*(0.04)*	0.04	*(0.04)*	0.10	*(0.04)*	0.12	*(0.05)*	0.06	*(0.05)*	0.08	*(0.05)*
Reach above shoulders	0.04	*(0.03)*	-0.01	*(0.03)*	0.07	*(0.03)*	0.04	*(0.04)*	0.06	*(0.04)*	0.10	*(0.04)*
Push large objects	0.09	*(0.03)*	0.10	*(0.03)*	0.12	*(0.03)*	0.10	*(0.04)*	0.12	*(0.04)*	0.16	*(0.04)*
Bathe without help	-0.04	*(0.05)*	-0.04	*(0.05)*	0.04	*(0.05)*	-0.01	*(0.08)*	-0.02	*(0.08)*	-0.02	*(0.08)*
Eat without help	0.05	*(0.06)*	-0.03	*(0.06)*	0.12	*(0.06)*	0.03	*(0.09)*	-0.03	*(0.09)*	0.05	*(0.09)*
Dress without help	-0.03	*(0.04)*	0.03	*(0.04)*	-0.01	*(0.04)*	0.08	*(0.05)*	0.01	*(0.05)*	0.03	*(0.05)*
Chi-square statistics *(p-value)*												
H^0: $\gamma_t=0$ df=17	1430.0	*(0.000)*	1596.5	*(0.000)*	1658.4	*(0.000)*	1022.6	*(0.000)*	969.3	*(0.000)*	989.0	*(0.000)*
H^0: $\gamma_1=\gamma_2=\gamma_3=0$ df=54	1640.4	*(0.000)*	1767.8	*(0.000)*	1807.9	*(0.000)*	1188.6	*(0.000)*	1149.3	*(0.000)*	1157.7	*(0.000)*

Source: Authors' tabulations using Health and Retirement Survey data.
Sample: Age-eligible men and women responding for selves in waves 1, 2 and 3 with no missing health or demographic information
Notes: Coefficients *(standard errors)* from ordered probit models of self-rated health (1=excellent, 5=poor). Functional limitation variables are standard normal rescalings of categorical responses (1=not at all difficult, 2=a little difficult, 3=somewhat difficult, 4=very difficult/can't do/don't do). If A% respond "1," B% respond "2", etc., then those reponding "1" are recoded to Φ^{-1}(A), those responding "2" are recoded Φ^{-1}(A+B), etc., where Φ^{-1}() is the inverse of the standard normal distribution. Rescaling is done separately by gender. Models also include controls for age, race, hispanic ethnicity, education, marital status, foreign born, and veteran status (for men).

Appendix Table A2a: Multinomial probit estimates of labor force transitions by Wave 3, among respondents employed at Wave 2
Comparison group: respondents employed in the same job in Waves 2 and 3
Men (N=2115)

Selection Parameter	ρ=0[a]			ρ=–0.75[a]		
	Diff. Job	**Apply for DI**	**Not emp., not apply for DI**	**Diff. Job**	**Apply for DI**	**Not emp., not apply for DI**
	Coeff. *(SE)*	**Coeff.** *(SE)*	**Coeff.** *(SE)*	**Coeff.** *(SE)*	**Coeff.** *(SE)*	**Coeff.** *(SE)*
twice-lagged health	--	--	--	--	--	--
(Wave 1)	--	--	--	--	--	--
once-lagged health	--	--	--	--	--	--
(Wave 2)	--	--	--	--	--	--
contemporaneous health	0.06	0.91	0.28	0.28	0.90	0.47
(Wave 3)	*(0.06)*	*(0.12)*	*(0.06)*	*(0.06)*	*(0.06)*	*(0.07)*
χ^2, all health vars, behavior eqs (DF=3)		63.33			273.20	
χ^2, all health vars, incl. select. eq. (DF=6)		245.63			451.02	
Log likelihood		-12755.39			-12757.50	

[a]**Note:** Represents correlation between error terms of behavioral & selection equations
Source: Author's calculations using Health and Retirement Survey data
Sample: Age-eligible men and women responding for selves in waves 1-3 with no missing health/demographic data

Appendix Table A2b: Multinomial probit estimates of labor force transitions by Wave 3, among respondents employed at Wave 2
Comparison group: respondents employed in the same job in Waves 2 and 3
Women (N=2154)

Selection Parameter	$\rho=0$[a]			$\rho=-0.75$[a]		
	Diff. Job	**Apply for DI**	**Not emp., not apply for DI**	**Diff. Job**	**Apply for DI**	**Not emp., not apply for DI**
	Coeff. *(SE)*	**Coeff.** *(SE)*	**Coeff.** *(SE)*	**Coeff.** *(SE)*	**Coeff.** *(SE)*	**Coeff.** *(SE)*
twice-lagged health	--	--	--	--	--	--
(Wave 1)	--	--	--	--	--	--
once-lagged health	--	--	--	--	--	--
(Wave 2)	--	--	--	--	--	--
contemporaneous health	0.13	0.99	0.24	0.30	0.91	0.40
(Wave 3)	*(0.05)*	*(0.10)*	*(0.05)*	*(0.05)*	*(0.05)*	*(0.05)*
χ^2**, all health vars, behavior eqs (DF=3)**		94.75			332.90	
χ^2**, all health vars, incl. select. eq. (DF=6)**		334.90			555.33	
Log likelihood		-16707.82			-16711.63	

[a]**Note:** Represents correlation between error terms of behavioral & selection equations

[b]**Note:** Tests joint signifance of health variables in the retirement equation

[c]**Note:** Tests joint signifance of health variables in the selection and retirement equations