NBER WORKING PAPER SERIES

HEALTH, ECONOMIC RESOURCES AND THE WORK DECISIONS OF OLDER MEN

John Bound
Todd Stinebrickner
Timothy Waidmann

Working Paper 13657
http://www.nber.org/papers/w13657

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2007

Financial support was provided by the National Institute on Aging (R01 AG17579-01), the Robert Wood Johnson through the Economic Research Initiative on the Uninsured at the University of Michigan, and the Social Security Administration through the Michigan Retirement Research Center. Bound would also like to thank the Center For Advanced Study in the Behavioral Sciences, Stanford University for support. Michael Stevens provided able research assistance. The project was made computationally feasible by our access to the supercomputer resources of SHARCNET at the University of Western Ontario which were developed with funding from the Canadian Foundation for Innovation and the Ontario Research and Development Challenge Fund. We would like to thank system administrator Gary Molenkamp and Baolai Ge for their help on the computing front. Stinebrickner has also benefitted from a time-release supported under the SHARCNET fellowship program. We have benefitted from the comments on an earlier version of this paper from seminar participants at the University of Michigan and the Canadian Econometric Study Group and especially from Peter Arcidiacono, Paul Contoyannis, Eric French, Donna Gilleskie, Steve Haider, Marteen Lindeboom, Chris Swann, an anonymous reviewer, and the editors of this Journal. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2007 by John Bound, Todd Stinebrickner, and Timothy Waidmann. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Health, Economic Resources and the Work Decisions of Older Men
John Bound, Todd Stinebrickner, and Timothy Waidmann
NBER Working Paper No. 13657
November 2007, Revised January 2008
JEL No. J14,J22,J26

ABSTRACT

In this paper, we specify a dynamic programming model that addresses the interplay among health, financial resources, and the labor market behavior of men in the later part of their working lives. Unlike previous work which has typically used self reported health or disability status as a proxy for health status, we model health as a latent variable, using self reported disability status as an indicator of this latent construct. Our model is explicitly designed to account for the possibility that the reporting of disability may be endogenous to the labor market behavior we are studying. The model is estimated using data from the Health and Retirement Study. We compare results based on our model to results based on models that treat health in the typical way, and find large differences in the estimated effect of health on behavior. While estimates based on our model suggest that health has a large impact on behavior, the estimates suggest a substantially smaller role for health than we find when using standard techniques. We use our model to simulate the impact on behavior of raising the normal retirement age, eliminating early retirement altogether and eliminating the Social Security Disability Insurance program.

John Bound
Department of Economics
University of Michigan
Ann Arbor, MI 48109-1220
and NBER
jbound@umich.edu

Todd Stinebrickner
Department of Economics
University of Western Ontario
London, Ontario, N6A 5C2
CANADA
and NBER
trstineb@uwo.ca

Timothy Waidmann
The Urban Institute
twaidman@ui.urban.org

Section 1. Introduction

Population aging and other challenges to public and private pension financing highlight the importance of understanding the determinants of retirement behavior. Much of the recent research on the labor force behavior of older workers has focused on the effects of financial incentives such as Social Security and private pensions, generally showing that these incentives have powerful behavioral effects (e.g., Blinder, et al., 1980; Burkhauser and Quinn, 1983; Diamond and Hausman, 1984; Stock and Wise, 1990; Rust and Phelan, 1997; Gruber and Wise, 1999, 2004; Gustman and Steinmeier, 1986, 2000, 2005). Additionally, a substantial amount of research has focused on the effects of the availability of both privately and publicly provided health insurance on retirement behavior (Rust and Phelan, 1997; Gustman and Steinmeier, 1994; Blau and Gilleskie, 2001a, 2006, Forthcoming). At the same time, econometric studies of retirement behavior have provided strong evidence for the importance of health factors (e.g., Quinn, 1977; Gordon and Blinder, 1980; Burkhauser and Quinn, 1983; Diamond and Hausman, 1984; Gustman and Steinmeier, 1986; Quinn et al., 1990; Rust and Phelan, 1997; Bound et al., 1999; Blau and Gilleskie, 2001b). Indeed, in analyses using Census data we found that more than half of men and one third of women who leave the labor force before reaching the Social Security early retirement age of 62 report that health limits their capacity to work (Bound et al., 1997).

However, as several reviews have noted, important questions remain regarding the magnitude of the effect of health on labor market behavior (Chirikos, 1993; Lumsdaine and Mitchell, 1999; Currie and Madrian, 1999). Moreover, except for early work by Quinn (1977) and research focusing specifically on the effects of changes in Social Security Disability Insurance (DI) on the work force attachment of older men (e.g., Parsons, 1980; Halpern and Hausman, 1986; Bound, 1989; Haveman et al., 1991; Bound and Waidmann, 1992, 2002; Kreider 1999a; Kreider and Riphahn, 2000), no one has studied the effect of the availability of financial resources on the relationship between health and retirement. This despite the likelihood that health and financial factors interact in affecting retirement decisions -- that is, deteriorating health will tend to make continued work less attractive, but individuals will tend to retire only if they have sufficient financial resources.

More fundamentally, previous longitudinal retirement research has suffered from limited measures of health, relying heavily on global measures such as self-rated work limitations and self-rated health. There are a number of potential problems with such survey measures: (1) they are discrete, whereas the construct researchers are interested in measuring is presumably continuous; (2) they are presumably error ridden, since not everyone will use the same scale when responding to survey questions; and (3) they are likely to be endogenous to retirement decisions, since it seems plausible that responses to these global questions will be related to labor market status.[1]

In this paper we use methods designed to address all three of these problems. We use both the global and the more detailed health measures available on the Health and Retirement Study (HRS) within the context of latent variable models. This approach can be thought of as using the detailed health measures available on the HRS to instrument the global measures. We find that the choice of how we model health is substantively important. To preview our results, our estimates confirm the central role health plays in the early retirement behavior of men. Indeed, we estimate the rate of labor force exit before the age of 62 to be 5 times as great for those in poor health than it is for those in average health. Importantly, however, we found that using the standard (binary) model of health would overstate the magnitude of this effect. At the same time, our simulations suggest that the availability of financial resources also plays an important role in determining behavior. In particular, we find that a large fraction of those who leave the workforce in poor health before the age of 62 apply for DI. Consistent with other estimates in the literature, our estimates imply that the DI application

[1]The potential endogeneity of self-rated work limitations or health has received a good deal of attention in the literature (e.g. Parsons, 1982; Myers, 1982; Anderson and Burkhauser, 1984, 1985; Bound, 1991; Waidmann et al.,1995). Indeed, compelling evidence indicates that responses to such questions depend not just on health, but also on features of the individual's social and economic environment. Bound and Waidmann (1992, 2002) demonstrate that the fraction of working aged men in the U.S. who are out of work and identify themselves as limited in their capacity for work tracks the fraction receiving Social Security Disability benefits quite closely, rising in the 1970s, falling in the 1980s, and rising again in the 1990s. Bound and Waidmann argue that the most plausible interpretation of these findings is that exogenous changes in the availability of disability benefits induced a change in reporting behavior. Waidmann et al. (1995) report similar trends for a range of other health measures, including self-reports of overall health and specific chronic conditions.

decision is quite sensitive to benefit levels. Even so, our estimates suggest that seemingly dramatic changes to Social Security rules-increasing the normal retirement age or eliminating the early retirement benefit--would have small spill over effects on the DI program. The reason is simple: most men in their 60s are too healthy to qualify for DI.

Similar approaches to modeling health have been used by other researchers in both cross-sectional (Kreider, 1999a) and longitudinal studies (Lindeboom and Kerkhofs, 2004; Disney et al., 2006). However, to our knowledge no one has embedded such modeling into the kind of dynamic programming models that are currently state of the art for modeling retirement behavior. This is the approach that we take in this paper.

For our method to yield valid results, responses to the more detailed measures available on the HRS must be exogenous to labor market behavior. While many researchers have worked with the presumption that the more specific health measures available on the HRS are less susceptible to the kinds of problems we have discussed, these measures may not be completely immune. We discuss evidence below that suggests to us that these measures are, indeed, exogenous. The latent variable model we use is computationally intensive, which may explain why other researchers have continued to use the global health measures available on the HRS. For this reason, it seems important to know whether our model yields answers that are substantially different from those we obtain if we follow the standard practice of simply including discrete health indicators in our behavioral models, and thus we include estimates from both approaches.

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 and Kozy, 1996; Quinn, 1997, 2000). While poor health induces many individuals to leave the work force altogether, it may induce others to merely change jobs or find ways to accommodate their limitation on their current job. A more general literature on the adaptations that older adults make in response to deteriorating health indicates that ceasing to perform an activity is often the response of last resort (Baltes and Baltes, 1990; Brim, 1988). Before this occurs, older adults will expend increased effort, allow more time, and reduce performance standards in order to perform the activity. However, to date very little research has attempted to model the effect of health on labor force transitions other than retirement.[2]

As far as we know, with the exception of work in progress by Rust and his colleagues (Rust, Benitez-Silva, Buchinsky 2001), we are the only researchers modeling retirement behavior to distinguish the application for DI benefits from other modes of labor force exit. Doing so seems crucial because: (a) our estimates suggest that, of those in poor health, the number that apply for disability is greater than the number that simply leave the workforce, and (b) the financial incentives involved in the two behaviors are quite different.

In our modeling and estimation to this point we have focused on single men nearing retirement age during the 1990s. We focus on single individuals to avoid the very significant complication arising from trying to model the joint labor supply decisions of married men and women. While, as a result, our sample is not representative of the population, we believe that we can learn a considerable amount by examining the behavior of this group.[3]

In section 2 we describe our dynamic programming (DP) model and the methods that we use to solve the value functions that are the key inputs into our estimation procedure. In Section 3 we describe the estimation methods we use. In Section 4 we present parameter estimates and simulations which highlight important aspects and implications of our model. Additionally we compare results based on our model to results that simply use self reported work limitations as the measure of health. In Section 5 we conclude.

[2] Exceptions include Honig and Reimers (1987), who find little association between poor health and the move from full-time work to partial retirement in their analysis of Retirement History Survey data; Blau and Gilleskie (2001a), who find that the effect of health on the probability that a person changes jobs is much smaller than its effect on labor force exit (using HRS data); and Bound et al. (1999), who find that poor health is a significant predictor of labor force exit -- particularly of exit combined with application for DI -- and of job change among people who choose to stay in the labor market.

[3] We focus on men because we do not have reliable information on the financial incentives facing single women. Their Social Security Benefits will often depend on their ex-husbands earnings, something we do not have information on. Further, while 39 members (20%) of our baseline behavioral sample do marry during the observation period, we ignore this possibility in the choice set.

Section 2. Model Specification and Solution

Model and Estimation Overview

We model the behavior of males who are working as of a "baseline" time (t=0) which corresponds to the first wave (1992) of the Health and Retirement Study. The basic behavioral model is a dynamic programming model in which individuals take into account that current period decisions may have substantial effects on their future utility. Central to this model is a set of current period utility equations that allows a person to construct the expected lifetime utility, or value, that he will receive from each option that he considers in each year that he makes a decision.

The solution of the value functions and the estimation of the parameters of these "behavioral" equations is complicated by our desire to address two issues. First, those who are working at our baseline time period are a select group of individuals. For example, from the standpoint of understanding the effects of health on behavior, it is possible that the individuals in poor health who are still working at t=0 have unobserved characteristics and preferences regarding work that are on average different from those of individuals who are in poor health at time t=0 but are no longer working. Second, although our model posits that individuals make decisions based on actual health, as mentioned earlier, it is self-reported health that is observed in our data.

We address the former concern by adding a reduced form initial conditions equation that describes whether a person is working at our baseline time period. Following Bound (1991) we address the latter concern by adding a latent health equation that formally describes the relationship between self-reported health, health reporting error, and true health. The presence of these additional equations has several practical implications that increase the difficulty of the solution and estimation of our model. First, in order for the additional equations to serve their purpose, our estimation procedure must allow correlations between certain unobservables that appear in the initial conditions equation, the health equation, and the behavioral equations. Our use of a multivariate normal distribution, which allows these correlations, implies that closed form solutions do not exist for integrals that are needed to compute value functions or for the likelihood contributions that serve as inputs into the Maximum

Likelihood algorithm that is used for estimation. Second, our health framework produces a continuous measure of true health that is serially correlated over time, a well-known challenge for researchers employing dynamic, discrete choice estimation methods (Stinebrickner 2000). These issues, when combined with our desire to include unobserved heterogeneity, our use of up to six years of observed choices in addition to the initial condition, and our need to include a non-trivial number of state variables other than health, imply that the computational burden of solving and estimating our model is very high.

In this section we describe the behavioral portion of our model and the methods used to solve value functions given the presence of the serially correlated health variable. This discussion implicitly assumes that the true health of each individual is known at each point that a person makes a decision. In reality, true health is not observed. In Section 3, we describe the modifications that we make to our model to address this issue and the sample selection/initial conditions issue, and we describe the estimation method that we implement to deal with the non-standard features of our model.

2.1 Choice Set

Each individual has a finite decision horizon beginning at year t=1 (1993) and ending at year t=T.[4] At each time t, an individual chooses an activity state from a finite set of mutually exclusive alternatives D_t. $D_t \subset \{C, B, N, A\}$ where C is the option of remaining in the person's career job (defined to be the job that the person held at baseline, t=0), B is the option of accepting a bridge job (defined to be a job other than the job held at baseline), A is the option of leaving the workforce and applying for Disability Insurance, and N is the option of leaving the workforce without applying for Disability Insurance (often referred to hereafter as the "non-work" option). Let $d^j(t)=1$ if option j is chosen (j=C, B, N, A) at time t and zero otherwise.

At any time t<T, a person can choose any of the options in the set {C, B, N, and A} unless it is ruled out by one or more of the following two assumptions. First, we assume that a person imagines that he will not return to his career job in the future if he leaves his career job in any year t. With

[4]In practice, we assume that T is the year that the person turns 70 years old. After year T, individuals are assumed to remain out of the workforce for the remainder of their lives. We assume that all individuals die by the age of 100.

respect to this assumption, we allow for both the possibility that a person could leave his career job by choice and the possibility that a person may get exogenously displaced from his career job for a reason such as a plant closing. Notationally, we let L(t) be an indicator of whether a person who is working in a career job at time t-1 gets exogenously displaced before time t. Second, we assume that a person imagines that, if he applies for Disability Insurance and is approved for benefits, he will remain out of the workforce (i.e., he will be in option N) and collect his Disability Insurance payments for the remainder of his life. Notationally, we let DI(t) indicate whether a person has been approved for disability benefits as of time t. A person can apply for Disability Insurance if he is less than the normal retirement age for Social Security Retirement Benefits (65 for most of our sample).[5] These assumptions imply that sufficient to characterize the choice set D_t is the person's age at t, the person's choice at t-1, whether the person becomes displaced from his career job between time t-1 and time t if he was working in his career job at time t-1, and whether the person has been approved for Disability Insurance at any time in the past.

This choice set implies that we do not formally model an individual's optimal consumption/savings decision. Rather, consistent with much previous research in the dynamic, discrete choice literature we assume that a person consumes all of his "income" in year t.[6] In Appendix E we describe the tradeoff between approximation quality and model "size" that influenced our decision not to expand the choice set to model the consumption/savings decision (and other endogenous decisions). Instead we view pension wealth, non-pension wealth, Social Security earnings, Disability Insurance payments, and other

[5]For this and other elements of the choice set, we carefully follow the eligibility and benefit rules of the Social Security Old-Age, Survivors and Disability Insurance program. See the Annual Statistical Supplement to the Social Security Bulletin (SSA 2007) for details.

[6]In recent work, researchers have begun to introduce savings into dynamic programing models of retirement (e.g., French, 2005; van der Klaauw and Wolpin, 2005; Rust, Buchinsky and Benitez-Silva, 2001; French and Jones, 2007). Doing so requires treating savings as a continuous state variable and consumption as a continuous choice variable which significantly complicates estimation. In all these cases the authors have treated health as an exogenous discrete state variable. In contrast, we treat health as a continuous state variable and allow for the potential endogenous reporting of health status, but ignore savings. While adding savings as a state variable and consumption as a choice variable is possible from a conceptual standpoint, in practice this change in our specification would make our model intractable at its current level of approximation quality. Given our interest in the interplay among health, financial resources, and the labor market behavior, we believe our choice to carefully model health was a reasonable one. In fact, those in poor health tend to have relatively little in the way of savings. Thus, for example, amongst those that identify themselves as having a limitation that effects their capacity for work in our sample, median non-housing wealth is under eleven thousand dollars at baseline.

entitlements as sources of income and attempt to make reasonable assumptions about the timing of the income from these sources in cases where the timing is not immediately obvious from institutional details.

Our specification of the opportunity set implies that an individual applying for Social Security Disability benefits will incur financial costs. For the year of application he will forgo all earnings. Further, if his application is rejected, he will not be able to return to his previous job which will tend to represent a loss of income since earnings on bridge jobs are typically lower than earnings in career jobs. These costs vary across the population. Those with little in the way of income outside of earnings will lose a greater proportion of their total income during the year they apply for disability benefits, and, as a result, will suffer a larger loss in utility if utility is not linear in consumption. In addition, those in high paying jobs stand to lose more by applying both because disability benefits are paid on a progressive schedule and because those with high paying jobs are likely to suffer a larger loss if they give up their career job for a bridge job.

2.2 Current Period Rewards

The current period reward in any year t, $R^j(t)$, contains all of the benefits and costs associated with alternative j; $R^j(t)$ is the sum of the utility from consumption, $U^j_{cons}(t)$, and the non-pecuniary utility, $U^j_{np}(t)$, that the person receives from option j at time t.

2.2.1 *Utility from Consumption,* $U^j_{cons}(t)$

Defining $Y^j(t)$ to be the person's total income net of expenditures on health care if he chooses option j at time t, the individual's utility from consumption is assumed to be of the form

$$(1)\quad U^j_{cons}(t) = \tau \frac{Y^j(t)^{1-\theta}}{1-\theta} \quad \text{for } j \in \{C,B,N,A\}$$

where θ determines the level of risk aversion and τ (along with parameters in the non-pecuniary utility equation (2) that will be discussed in Section 2.2.2) is used to determine the importance of utility from consumption relative to non-pecuniary utility.

A benefit of the HRS is that it allows us to capture in detail how expenditures on health care and income vary across the possible options j. For option j at time t, $Y^j(t)$ is the sum of income from earnings, Social Security entitlements, defined benefit and defined contribution pension plans, Disability Insurance payments, non-pension wealth, food stamps,[7] Supplemental Security Income, and other exogenous sources of income (such as veteran benefits) minus expenditures on health care. In Appendix A.2 we discuss in general terms our timing assumptions related to the receipt of income from these sources and later in Section 2.3.2 (and in Appendix A.3) we provide detail about modelling and computation issues related to these sources. At this point it is worth noting that the reason that we can take full advantage of the detail about these incomes sources in the HRS is that, unlike much other work using the types of models employed here, we solve our model separately for each person in our sample.

All incomes are converted to 1992 dollars. The concept of income we use is after tax income. To this end we subtract off from our estimate of gross income both the worker's share of the payroll tax and a piecewise linear approximation of the individual's federal income tax obligations.[8]

While the approach we use for modeling the effect of the availability of health insurance is now common in the literature (Rust and Phelan, 1997; French and Jones, 2007; Blau and Gilleskie, 2006), using it implies that we are treating health care utilization as exogenous.[9] Indeed, the observed difference in out of pocket health care expenditures between those that do and do not have health insurance almost surely represents an underestimate of the value individuals put on the availability of health insurance benefits. In an attempt to mitigate the bias that this underestimation introduces in the estimates of our behavioral equations in Section 2.2.2, we allow the availability of employer provided health insurance benefits to have a direct effect on an individual's utility.

[7]When imputing food stamp benefits, we use our estimate of what the individual would be eligible for. This effectively puts a floor on income.

[8]Using 1992 dollars, we assume the income tax rate is 0 up to \$5280, 0.1851 between \$5280 and \$34600, 0.3354 between \$34600 and \$80863, 0.3689 between \$80863 and \$127600, 0.4215 between \$127600 and \$214000 and 0.4636 above \$214000. For the purpose of computing taxes we use only a portion of the Social Security Retirement or Disability benefits a person is receiving.

[9]As far as we know the only researchers who have tried to endogenize health care utilization decisions within the context of retirement models are Blau and Gilleskie (Forthcoming). Similar to what we find, they find that the availability of retiree health insurance seems to have little effect on retirement behavior.

2.2.2. Non-Pecuniary Utility, $U_{np}^{j}(t)$

We assume that the nonpecuniary utility $U_{np}^{j}(t)$ associated with an option j is a linear function of a person's time t health η_t, an indicator $HI^j(t)$ of whether the person has either private health insurance or medicare at time t if he chooses option j, exogenous observable characteristics of the individual X(t), and a set of other transitory factors (ϵ_t^j) unobserved by the econometrician (but known to the individual in the current period) that measure the person's particular circumstances and outlook in year t. In addition, we allow individuals to have unobserved, permanent differences in their preferences for work by including a person-specific, permanent heterogeneity term κ that enters the non-pecuniary utility associated with the work options C and B, i.e., $\kappa^C=\kappa^B=\kappa$ and $\kappa^N=\kappa^A=0$.

$$U_{np}^{j}(t) = \lambda_X^j X(t) + \lambda_{HI} HI^{j}(t) + \kappa^j + \lambda_\eta^j \eta_t + \epsilon_t^j \quad \text{for } j \in (C,B,N,A). \tag{2}$$

Thus, in addition to allowing health to have effects on net income, our model allows decisions to have non-pecuniary costs which depend on a person's health.[10]

We choose N as the base case of our discrete choice model which implies that we normalize the coefficients λ_X^N and λ_η^N to zero and interpret λ_X^j, j=C,B,A as the effect of X on the utility of option j relative to option N and interpret λ_η^j, j=C,B,A as the effect of η_t on the utility of option j relative to option N. To summarize,

$$R^{j}(t) = U_{cons}^{j}(t) + U_{np}^{j}(t) = \tau \frac{Y^{j}(t)^{1-\theta}}{1-\theta} + \lambda_X^j X_t + \lambda_{HI} HI^{j}(t) + \kappa^j + \lambda_\eta^j \eta_t + \epsilon_t^j, \quad j \in (C,B,N,A) \tag{3}$$

with $\lambda_X^N = 0$ and $\lambda_\eta^N = 0$.

2.3 Discounted Expected Utility - Value Functions

2.3.1 Specification of Value Functions

[10]Given the way we have specified a person's alternatives, a person who applies for DI will receive non-pecuniary utility of $U_{np}^{A}(t)$ in the year that he applies. If accepted he is assumed to remain out of the workforce and receive non-pecuniary utility of $U_{np}^{N}(t)$ for the remainder of his life. Thus, for example, λ_η^A indicates, in part, how the cost or stigma of taking part in the DI application process varies with health.

Letting S(t) represent the set of all state variables at time t, the expected present value of lifetime rewards associated with any option $j \in \{C,B,N,A\}$ that is available at time t can be represented by a standard Bellman equation (Bellman 1957):

(4) $V_j(t, S(t))=R^j(S(t)) + \beta(S(t))\cdot E[V(t+1,S(t+1))|S(t),d^j(t)=1]$

where $V(t+1,S(t+1)) = \max\{V_k(t+1,S(t+1)) \ : \ k \in D_{t+1}(d^j(t)=1,S(t+1))\}$.

We have written D_{t+1} as a function of $d^j(t)$ and S(t+1) because, as discussed earlier, a person's choice set at time t+1 depends on the person's choice at t, the person's age at time t+1, whether the person becomes displaced from his career job between time t and time t+1 if he was working in his career job at time t, and whether the person has been approved for Disability Insurance at any time in the past.

β is the one period discount factor which varies across people and across time for a particular person. Specifically, we assume that for person i at time t, β depends on a factor β^{Common} that is common across people and on the probability that person i will be alive at time t+1:

(5) $\beta=\beta^{Common}\cdot Pr(\text{Alive at } t+1|\text{Alive at } t)$.[11]

We assume that the probability of dying between t and t+1 depends on the respondent's age and his health at t, η_t. This probability is computed using a discrete-time proportional hazard model. The baseline hazard, which represents the probability of dying at a particular age conditional on not dying before that age, is computed using life table survival probabilities for U.S. men obtained from the Social Security Administration. Health shifts the baseline hazard in a proportional fashion.[12]

2.3.2 State Variables

The set of state variables at time t, S(t), includes all variables that provide information about the set of choices that will be available in the current and future periods, the discount factor, or the utility associated with all choices that may be available in the current and future periods. The information

[11]This specification implies that the value of death is zero so that people have no bequest motive.

[12]The amount of the shift is estimated outside of our behavioral model using information on the subsequent mortality of our HRS sample together with a health index that is the same as the one used in our behavioral models. See Appendix A.4 for details.

that influences future choices and the discount factor was described in Sections 2.1 and 2.3.1 respectively. In the next two subsections we focus on the state variables that influence either non-pecuniary or pecuniary utility - $U^j_{np}(t)$ or $U^j_{cons}(t)$.

State variables that influence non-pecuniary utility

Equation (2), indicates that non-pecuniary utility at time t is determined by X(t), $\epsilon(t)$, κ, η_t, and HI(t) where $\epsilon(t) \subseteq \{\epsilon^C_t, \epsilon^B_t, \epsilon^N_t, \epsilon^A_t\}$ is the vector of ϵ's from all of the current period utility equations that are relevant in time t given a person's choice set. X which includes, for example, a constant and a person's educational level, is predetermined and known to the agent and econometrician for all periods. The permanent, person-specific, unobserved heterogeneity value, κ, is known to the individual but is unobserved to the econometrician. We assume that, in the population, $\kappa \sim N(0,\sigma^2_\kappa)$ where σ_κ, which determines the importance of unobserved heterogeneity, is a parameter to be estimated. $\epsilon(t)$ is observed by the individual but not by the econometrician at time t. Both the econometrician and individual know the distribution of ϵ in future periods. We assume that $\epsilon_t^j \sim N(0,1)$, j=C,B,A,N and that $E(\epsilon_t^j, \epsilon_r^k)=0$ if $j \neq k$ or $t \neq r$.[13]

A person's health, η, is exogenously determined but correlated across time. We assume that health at time t depends on demographic characteristics in X(t), including a person's age. Based on evidence in Bound et al. (1999), we assume that the portion of health that remains after removing the effect of X(t) in each period follows an AR(1) process:

(6) $$\eta_{t+1} = \rho(\eta_t - \pi X_t) + \pi X_{t+1} + \xi_{t+1}$$

where $\xi_{t+1} \sim N(0,\sigma^2_\xi)\ \forall t$. Given a current period value of health, both the agent and econometrician can use equation (6) to compute the distribution of health in all future periods. However, while the agent knows his current health, the econometrician observes only a noisy, self-reported health measure. The manner in which we deal with this data problem is an estimation issue which we discuss in detail in Section 3.

[13] Normalizing the variance implies that a constant can be estimated as part of $\lambda^j_X X_t$ in equation (3).

Finally, at time t a person's beliefs about his health insurance status at time t+1, HI(t+1), is determined by the health insurance characteristics of his career job (which we denote HI^C), the health insurance characteristics of his bridge job at time t if he is working in a bridge job at time t (which we denote $HI^B(t)$), and the person's age at t+1. We identify the health insurance associated with the career job to be one of three types: $HI^C=3$ if the insurance plan covers the worker while he is working on his career job and also provides retiree health insurance which covers him after he leaves the job; $HI^C=2$ if the insurance plan covers the worker while he is working on his career job but does not provide retiree coverage; $HI^C=1$ if the person has no health insurance on his career job. Primarily for computational reasons, we assume that bridge jobs do not have retiree health insurance. Thus, there are only two possible characterizations for $HI^B(t)$: $HI^B(t)=2$ if the insurance plan covers the worker while he is working on his bridge job and $HI^B(t)=1$ if a person does not work in a bridge job in time t or works in a bridge job that does not have health insurance.

We assume that at time t a person believes he will have health insurance at future time t+1 if any of the following conditions are true: 1). $HI^C=3$; 2). $d^C(t+1)=1$ and ($HI^C=2$ or $HI^C=3$); 3). $d^B(t+1)=1$ and $HI^B(t)=2$; 4). $Age(t+1)\geq 65$; or 5) DI(t)=1. The first condition indicates that a person with retiree health insurance on his career job believes that he will always have health insurance. The second condition identifies a person who is still working in a career job which has health insurance. The third condition indicates that a person who has health insurance in a bridge job imagines that he will continue to have health insurance if he remains in a bridge job in the next period. The fourth condition is present because everyone who has turned 65 years of age receives medicare. The last condition shows that, if approved for Disability insurance, individuals begin to receive medicare benefits (after a waiting period).[14] In addition, we assume that a person who is working in a bridge job without health insurance at time t or has chosen an option other than the bridge option at time t believes that there is

[14]This is a rough approximation of the reality that individuals receive Medicare benefits 24 months after starting Disability Insurance benefits.

some probability that the bridge offer he receives in time t+1 will include health insurance. [15] In addition, we allow any person who has employer provided health insurance at time t but not at time t+1 to buy COBRA insurance at time t+1.[16]

Then, to summarize, at time t the state variables that influence non-pecuniary utility are {X(t), ϵ(t), κ, η_t , HI^C, and HI^B(t)}.

State variables that influence income

Some of the variables that influence non-pecuniary utility also provide information about current and future income levels Y^j, j=C, B, N, A. For example, the specification of the health expenditure equation in Appendix A.3.1 indicates that X(t), η_t, HI^C, HI^B(t), and DI(t) all influence health expenditures at time t.

In addition, some new state variables are needed to represent a person's information about income. For example, income calculations depend in part on a set of baseline variables, β, that describe everything about a person's financial situation, previous work history, and earnings potential when the person arrives at t=1. This set of baseline variables describes exogenous sources of income (such as veterans benefits) and also contains information about a person's wealth at time t=1. In addition, because β contains information about a person's complete SS earnings history as of time t=1 and the specific details that characterize an individual's defined benefit and defined contribution pension plans, it also plays an important role in determining the income that would be received from the remaining sources of income described in Section 2.2.1: earnings, the SS and DI systems, and DB and DC pension plans. Below we describe the state variables that are needed (in addition to β) to characterize what a person knows about the income from each of these sources. More detail on modelling and computation issues related to these sources is presented in Appendix A3.

The earnings equations are given in Appendix A.3.2. Earnings in career jobs depend on a fixed effect (which can be viewed as an element of β) and a transitory component ψ^C_t. Earnings in bridge

[15]We assume that this probability is .20 for a person who is working in a bridge job without health insurance at time t and is .67 for a person who has chosen an option other than the bridge job option in time t.

[16]We assume that the cost of this coverage is $1000.

jobs are allowed to vary with a person's baseline earnings in his career job W^C_O (which is contained in θ), a transitory component ψ^B_t, and the age at which a person left his career job. Notationally, we let $\psi(t) \subset \{\psi^C_t, \psi^B_t\}$ be the vector of transitory earnings shocks that are relevant at time t given a person's choice set. Sufficient for knowing the age at which a person left his career job is the person's age at baseline (contained in X) and the number of years of experience that the person worked in his career job as of time t (which we refer to as EXC(t)).[17]

A person's SS benefits at some future year t* depend on his 35 highest years of labor earnings, the age when he began receiving SS benefits, and details about any earnings that were received after beginning benefits. Sufficient for providing this information is the person's earnings history as of time one (which is contained in the baseline characteristics θ), and his complete earnings history between time t=1 and time t*-1. Unfortunately, a specification which requires the agent to keep track of a complete earnings history is not tractable since it requires that a person's entire histories of the ψ^C's and the ψ^B's be treated as state variables in the model. Our model is made tractable through an assumption that an individual considers expected future earnings rather than actual future earnings when thinking about future SS benefits.[18] In this case, sufficient for computing the SS benefits that a person will receive in some future year t* is the person's earnings history as of time t=1 (which is contained in θ), the number of years that he will work in his career job after time zero and before time t* (which we denote EXC(t*)), the number of years that he will work in his bridge jobs after time zero and before time t* (EXB(t*)), and a variable which keeps track of all relevant information about what years the person worked after age 62 and before time t* (which we denote SSEX(t*) and describe in

[17]Because the person cannot return to his career job after leaving, EXC(t) represents the year at which a person who is working in a bridge job left his career job.

[18]While this assumption is made primarily for computational reasons related to the size of the state space, it will only tend to be restrictive if yearly randomness in career job earnings (associated with the ψ^C's) generates a large amount of variation in the defined benefit and defined contribution payments or if yearly randomness in career job or bridge job earnings (associated with the ψ^C's and ψ^B's) generates a large amount of variation in Social Security or DI benefits. There are several factors which mitigate the influence of this assumption. First, given our fixed effects specification for career job earnings, the variation of the unobservables ψ^C_t in Appendix equation (A.2) is relatively small. Second, over several years, positive shocks to earnings in some years tend to be offset by negative shocks to earnings in other years and this tends to have an offsetting effect on pension benefits and SS benefits. Finally, for many people, a large proportion of DB, DC, SS, and DI benefits are already determined by the time they reach the later parts of their working lives.

more detail in Appendix A.3.3). These three state variables are endogenously determined within the model.[19]

As with the SS calculation, we assume that individuals consider expected future earnings when thinking about payments from DB pensions, DC pensions, and the DI system. In this case, a person can compute the DB payment he will receive from his career job at some future time t* if he knows the details of the pension plan and his earnings history as of t=1 (which are both contained in the set of baseline information θ) and the year that he left his career job, as described by EXC(t*).With respect to defined contribution plans, future payments will depend on details of the plan, past contributions, and future contributions. We assume that an individual will continue to contribute to the DC plan at his career job at the same rate as he has contributed in the past. In this case, as with DB benefits, sufficient to characterize DC benefits at some future t* is information in θ and EXC(t*). Disability Insurance benefits are a part of the Social Security system and, with the exception of differences that arise because DI benefits are not age-restricted, are determined in a manner similar to SS payments. This implies that an individual can compute the DI payment he would receive at some future time t* if he knew the baseline information θ, EXC(t*), EXB(t*), and whether he has been approved for benefits as of time t*, DI(t*). We assume that a person who applies for Disability Insurance benefits at some time t is approved for benefits if

(7) $A_1^{DI} + A_2^{DI}\eta_t + e^{DI} > 0$

where A_1^{DI} and A_2^{DI} are a constant and slope coefficient and e^{DI} is a random component that is normally distributed.

Then, to summarize, the time t state variables in the model are $S(t)=\{\theta, X(t), EXC(t), EXB(t), SSEX(t), DI(t), \epsilon(t), \psi(t), L(t), \kappa, \eta_t, HI^C, \text{and } HI^B(t)\}$.

---

[19]Essentially, these variables are sufficient to characterize the entire earnings history that is relevant for the SS calculation if the person thinks about expected earnings in the future.

2.3.3 Solving value functions

The expected value in equation (4) is a multi-dimensional integral over the stochastic elements of S(t+1) whose realizations are not known at time t given the decision to choose j. For illustration, consider a person who is working in his career job in time t, $d^C(t)=1$. In this case, the stochastic elements of the state space whose time t+1 realizations are not known are L(t+1), $\epsilon(t+1)$, $\psi(t+1)$, η_{t+1}, and $HI^B(t+1)$. L(t+1) and $HI^B(t+1)$ are discrete random variables so the expected value involves summing over the probability functions of these variables and integrating over the density functions of the remaining continuous variables.[20]

Researchers have often relied on convenient distributional assumptions to reduce the burden of evaluating integrals of the type described in the previous paragraph. For example, as shown in Rust (1987) if one specifies the choice specific transitory shocks (i.e., $\epsilon(t)$ in our case) to be iid extreme value, the expected value in equation (4) has a closed form solution conditional on the values of the other state variables.[21] However, in this application, the Section 2.3.2 normality assumption for $\epsilon(t)$ is driven by practical considerations related to the importance of allowing certain correlations that will be discussed in detail in Section 3. This assumption, along with the equation (6) assumption about the distribution of η_{t+1} given η_t and the fact that we wish to avoid functional form assumptions related to earnings by taking advantage of the empirical distribution of ψ, implies that the expected value in equation (6) does not have a closed form solution. In Appendix B we describe in detail our method for approximating the integrals involved in this expectation. This method involves a combination of Gaussian quadrature and simulation methods. Of importance from the standpoint of estimation

---

[20]If the person has chosen a time t option other than his career job, L(t+1) is not relevant and the dimensions of $\epsilon(t+1)$ and $\psi(t+1)$ are reduced by one. In addition, if a person is working in a bridge job with health insurance in time t, no uncertainty exists about $HI^B(t+1)$ because the person believes that he will have bridge insurance in time t+1. Finally, if a person applies for disability insurance at time t, uncertainty exists about whether he will be approved in the next period.

[21]A recent example that takes advantage of the extreme value assumption is Diermeier, Keane and Merlo (2002) who are able to estimate a dynamic programming model of the decisions of congressional members with a very large state space by taking advantage of extreme value errors. Keane and Wolpin (1994) explore approximations based on simulation approaches that are useful in cases where closed form solutions do not exist.

feasibility, the derivatives of the expected value "approximator" are continuous with respect to model parameters.

The recursive formulation of value functions in equation (4) motivates a backwards recursion solution process of the general type that is standard in finite horizon, dynamic, discrete choice models. The most basic property of the algorithm is that in order to solve all necessary value functions at time t, it is necessary to know value functions at time t+1 for each combination of the state variables in S(t+1) that could arise at time t+1. In Appendix B we discuss computational issues that arise when implementing the backwards recursion solution process in our particular application, including the modification that is needed to deal with our continuous, serially correlated health variable.

Section 3. Estimation

Individuals make choices by comparing the values of the various options that are available. Generally speaking, our estimation approach is to choose parameters that maximize the probability of observed choices. However, as discussed at the beginning of Section 2, we would like to address two issues during estimation. First, although our model posits that individuals make decisions based on actual health, it is self-reported health that is observed in our data. Second, the group of individuals that are working at baseline is a select group of individuals. In sections 3.1 and 3.2 we discuss these two issues in turn and then in Section 3.3 and Appendix C we describe our Simulated Maximum Likelihood estimation approach.

3.1 Health

Because true health is unobserved, we use a latent variable model to construct an index of health (Bound 1991, Bound et al. 1999). Specifically, we imagine that health in time t is a linear function of exogenous factors (e.g. age and education), X_t; detailed health measures (i.e., physical performance measures), Z_t ; and other unobserved factors ν_t.

(8) $$\eta_t=\pi X_t+\gamma Z_t+\nu_t$$

We assume that ν_t is uncorrelated with both X_t and Z_t (this assumption is essentially definitional: ν_t is the part of health that is uncorrelated with X_t and Z_t). While we do not directly observe η_t, we do observe an indicator variable, h_t, of whether a person is work limited. Letting h_t* represent self-reported health at time t, the latent counterpart to h_t, we assume that h_t* is a simple function of η_t and a term reflecting reporting error

(9) $$h_t^* = \eta_t + \mu_t$$

We assume that μ_t and η_t are uncorrelated. Substituting equation (8) into equation (9), we get

(10) $$h_t^* = \pi X_t + \gamma Z_t + [\nu_t + \mu_t].$$

If $\nu_t+\mu_t$ is assumed to be normally distributed with a variance that is normalized to be one, equation (10) represents a probit model in which h_t* is greater than zero if the person reports that he is work limited. Estimates that use respondents rating of their own health on a five point scale as our measure of h_t* yield very similar results to those reported here. The relative size of var(ν_t) and var(μ_t) is not important for the estimation of π and γ in equation (10) but is important for other parts of the model because, for example, it is true health (i.e., the portion not including μ_t) that enters the utility equations.

The composite error term in equation (10), $\nu_t+\mu_t$, reflects a number of different factors. The ν_t component reflects aspects of health not captured by X_t and Z_t, while the μ_t component reflects reporting errors. These errors reflect differences in reporting behavior across individuals and across time for the same individual. The presence of μ_t introduces a number of biases in our estimates if we were to use h_t* directly when estimating the impact of health on labor market outcomes. If μ_t were completely random, it would represent classical measurement error, which will attenuate the estimated effect of health on labor market outcomes. If, however, people use health as a way to rationalize labor market behavior, then one would expect μ_t 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. This consideration suggests that our specification should allow for the possibility that the reporting error μ_t is correlated with each of the shocks ϵ_t^C, ϵ_t^B, ϵ_t^N, ϵ_t^A in the behavioral equations. For identification reasons similar to those that require us to set the equation (3) current period utility coefficients λ_X^N and λ_η^N in the base case to zero, we normalize the covariance between the reporting error and the utility unobservable in the base case to be zero (i.e., $COV(\mu_t,\epsilon_t^N)=0$) and estimate the three covariance parameters $COV(\mu_t, \epsilon_t^j)$, j=C,B,A.

Essentially, our latent variable model uses the detailed health information available in the HRS (the *Z's*) to instrument the potentially endogenous and error-ridden work limitation measure, h_t^*. The validity of this approach for estimating the effects of health on labor force withdrawal depends critically on the assumptions that the reports on the detailed health information available in the HRS are exogenous with respect to labor force status. In Bound et al (1998), we test this assumption by comparing the performance of our preferred health model to health models estimated using a sparser and arguably more clearly exogenous set of measures from the HRS and find no evidence that the physical performance measures we are using are endogenous to labor market status. There are a number of reasons we do not simply use the more detailed performance measures directly in our behavioral equations. Among these, the measures reflect only a component of health and our latent variable model substantially reduces the number of parameters we need to estimate. Substituting equation (8) and equation (3) into equation (4) shows that the value functions at time t can be rewritten as

$$(11)\quad V_j(t,S(t))= \tau\frac{Y^j(t)^{1-\theta}}{1-\theta}+ \lambda_X^j X_t + \lambda_{HI} HI^j(t) + \kappa^j + \lambda_\eta^j[\pi X_t + \gamma Z_t] + \lambda_\eta^j \nu_t + \epsilon_t^j$$
$$+ \beta E[V(t+1,S(t+1))|S(t),d^j(t)=1] \quad j=C,B,N,A$$

3.2 Initial Conditions

Although the choices we have been considering are all conditional on a person being employed at time t=0, this group will be a non-random sample of the population of people working at time t=0. To account for this we include in our estimation a reduced form initial conditions equation. In particular, we imagine a latent variable I* that is greater than 0 if the individual is working as of t=0 where

$$(12)\quad I^*=\Pi_1 X_0+ \Pi_2 Z_0 +\epsilon^{I.}$$

We assume that $\epsilon^I \sim N(0,1)$ in which case equation (12) is a probit model.

In this reduced form specification, ϵ^I captures both the portion of true health at baseline and the portion of preferences for work at baseline that are not captured by observed characteristics (i.e., not captured by demographic characteristics at t=0, X_0, and physical performance measures at t=0, Z_0). The former suggests that ϵ^I may be correlated with the unobserved portions of health ν_t, t=1,2,.... Equation (6) implies that for t>1, COV(ϵ^I,ν_t) is a function of COV(ϵ^I,ν_1) and ρ, and we estimate COV(ϵ^I,ν_1). The latter suggests that ϵ^I may be correlated with unobserved preferences to work which influence behavioral decisions in t=1,2,...,T. To allow for this possibility we estimate COV(ϵ^I,κ), the covariance between the initial conditions equation and the permanent unobserved heterogeneity term.

Credible identification of the covariance between the initial condition and the behavioral equations depends crucially on exclusion restrictions. In particular, some variable or variables must influence the initial condition, but have no direct effect on subsequent behavior. In our case we have assumed that, while health at t=0 affects whether or not one works at t=0, it does not have a direct effect on subsequent behavior after conditioning on health at t=1. We believe this assumption is a natural one. Current health affects current behavior directly by affecting the utility that a person derives from work and also affects behavior through the role that is plays in determining individuals' expectations about future health. After conditioning on current health, it seems reasonable to believe that the primary avenue through which past health would influence current behavior is that decisions made in the past (which are influenced by past health) have an implication for the set of choices that are available to the person in the current period. In this case, after conditioning on a person's opportunity set and his current health, it does not seem that past health should have much of a direct impact on behavior. In addition, for this exclusion restriction to be valid, health must be exogenous to retirement and must follow a Markov process. Like much other research in this area, we maintain these assumptions throughout. In earlier work (Bound, Schoenbaum, Stinebrickner and Waidmann, 1999 we found the Markov assumption to be a reasonable one. In current work (Bound and Waidmann, 2007) we test the exogeneity of health to retirement by testing to see if there are identifiable changes in health in response to exogenous retirements. We found no evidence of such changes and concluded that the assumption that health is exogenous to retirement to be a reasonable one.

3.3 The Likelihood Function

Estimation proceeds by evaluating the joint probability of the simultaneous conditions that must be satisfied for a person who is working at our baseline t=0 (i.e., is in our behavioral sample) or is not working at our baseline period t=0 (i.e., is not in our behavioral sample). The set of simultaneous conditions that must hold can be written in terms of the simultaneous equations (10), (11), and (12) that define our model and contain the parameters to be estimated. We describe the likelihood function and the methods we use to compute the likelihood function in Appendix C.

Section 4. Results

4.1 Data

Data for this research come from the Health and Retirement Study (HRS), which contains both the detailed health data necessary to implement our latent variable framework and the the labor force and economic data necessary to accurately model the choice sets faced by individuals. The first wave (wave 1) of the survey was conducted in 1992/93; respondents were re-interviewed in 1994 (wave 2) and at two-year intervals since. 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 estimation of our model uses the public release versions of the first four waves of data, supplemented by confidential matched data from the Social Security Administration giving earnings histories and from employers giving details of private pension plans in which respondents are enrolled. Once we limit ourselves to single men with valid data, we end up with a sample of 328 individuals in our initial conditions sample and 196 (working) individuals in the behavioral sample who contribute a total of 837 person-year observations. More detail about the HRS and the composition of our sample is presented in Appendix A.1.

4.2 Timing of Behavioral Choices and Health

As described in Section 2.1, the decision periods in our model are one year in length. Assumptions are required to map the continuous work histories that can be constructed from the data to a sequence of yearly decision periods, each characterized by a single behavioral decision. A consideration particularly relevant for making these assumptions is that a person's economic incentives tend to vary with the person's age. For example, changes in the payment amounts that a person would be eligible to receive from the social security system if he were to retire typically take place on a persons' birthdays. This motivates our desire to have each yearly period in our model correspond to a birthday year (i.e., the period that a person is a particular age). From a practical standpoint, in order to assign a single behavioral decision to a birthday year it is necessary to choose the point in time during the birthday year at which the behavioral decision will be determined from the data. In order to take into account that a person may not always make decisions immediately after his economic incentives change, we choose the point in time to be close to the end of the birthday year.[22] In the typical case where there are zero or one transition during a birthday year, this approach allows the behavioral choice for a particular age to reflect whether a transition has taken place at any time during which the financial incentives related to that age are relevant.

While a person's activity status can be ascertained at any point during the sample period, a person's health measures are available only at the HRS interview dates. We map the health information to a decision period using the person's age at the time of the interview. Given that interviews take place approximately two years apart, this implies that health information is only observed for a subset of the yearly decision periods for which we have determined a behavioral decision. As discussed in Appendix C, we address this issue by integrating over the joint probability of the missing health values as suggested by the missing data literature.

Figure 1 gives an example of how surveys are used to establish timing of health status measurements and job status transitions. For a hypothetical individual born on 15 September 1937,

[22]The HRS masks the full date of birth information to protect confidentiality. We measure behavioral choices on the first day of the month preceding the respondent's birth month.

the behavioral year runs from 15 September to 14 September the following year. For such a person, we will measure work status as of 1 August in each year. We set t=0 at the first such point following the baseline interview. In this example, the baseline interview took place in July, 1992, making 1 August 1992 correspond to t=0. For this individual, age at the interview (and at t=0) is 54. Depending on the spacing of the interviews, there may be one, two, or three possible transition points between survey waves. In the case illustrated, behavior at t=1 is defined using the second wave survey conducted in June 1994. Behavior at t=2 and t=3 is defined using the third wave survey from June 1996, and behavior at t=4, 5, and 6 is defined by the fourth wave survey in January 1999. In this example, health status is observed at t=0,2,4 & 7.

4.3 Descriptive Statistics

We are interested in understanding how the availability of economic resources and health affect economic behavior. Table 1 presents descriptive information on the incomes sources in 1991 for age-eligible men in the HRS. Results are stratified according to whether or not the man was working as of the date of his wave 1 interview and whether or not he identified himself as suffering from health conditions that might limit his capacity for work. The table is limited to those who report no change in employment or disability status between January 1991 and the date of their wave 1 interview. This restriction was imposed to ensure that the incomes reported represent incomes commensurate with the data we use to stratify the sample.

While only 27% of the overall sample report work limitations, more than 75% of those out of work report work limitations. Focusing on those not working, income sources differ substantially depending on the respondent's self-reported work limitation status. For example, while 40% of the men without work limitations report pension income, less than 14% of those with work limitations report pension income. In contrast, roughly 68% of those with limitations report receiving income from one of the major federal disability programs, Supplemental Security Income (SSI) and Social

Security Disability Insurance (DI).[23] Crudely put, Table 1 suggests that men are not likely to leave the labor force before the age of 62 unless they have income sources on which they can rely, but that the composition of the income sources that are used to support an exit from the labor force varies dramatically with health status. Not surprisingly, those who are working and report work limitations have lower incomes than those who are working and do not report work limitations. However these differences may have preceded the work limitation.

Table 2 presents incomes and income sources as of 1991 and 1999 for age-eligible men working as of wave 1, stratified by behavior as of wave 4. Here we see, for example, that while almost 90% of those men who continued to work in their career jobs as of wave 4 (1998) had earnings in 1999, only 70% of those who had changed jobs between wave 1 and wave 4 and only 19% of those who had applied for disability benefits had earnings in 1999. At the median, household incomes rose by 65% between 1991 and 1999 for those that stayed with the same employer. In contrast, the drop in median income for those that retired was about 25%, and for those that applied for disability benefits median income declined by about 7%.

Table 3 shows descriptive statistics for both the behavioral sample (working in 1992) and the group of individuals who are not working. The first column in this table shows that the average age at the last survey for individuals in our behavioral sample is 60.6. Approximately 10% report that they suffer a work limitation. The second column of Table 3 shows that men in our sample who are not working at baseline tend to be slightly older and are approximately three times as likely than those working to report having a work limitation.

Recall that the choice data used to identify the behavioral portion of the model come from the activity status of our behavioral sample at approximately yearly intervals. The third through sixth columns of table 3 report descriptive statistics on our behavioral sample broken down by whether

[23]In Table 1, DI payments and regular SS payments are grouped under the category Social Security. Although not shown explicitly in Table 1, approximately 44% of those not working and reporting work limitations are receiving income from the DI program.

they chose C, B, A, or N in the final survey.[24] There are several things to note. Those who retire (i.e., choose option N) are more likely to be eligible for a private defined benefit pension and more likely to have reached age 62 by wave 4 than those who did not. What is more, those who retire (N)--and especially those who applied for DI benefits (A)--were no more likely to be in poor health or report a work limitation at their wave 1 interview, but were much more likely to report health problems as of the final survey. These patterns make considerable sense.

A person in our behavioral sample would receive an average of $12,820 in SS benefits at age 65 (based on the contributions made as of the baseline interview) and would receive DI benefits of the same amount if he is approved for the program. On average, the expected career earnings and bridge earnings at the final survey based on estimates of Appendix equations (A.2) and (A.3) are $32,854 and $13,098 respectively.

4.4 Model Estimates and Simulations

The parameters that enter our model are:

1). the parameters of the DI approval equation (7);

2). the parameters π, γ, and σ^2_ν from the health equation (8) and the parameters ρ and σ^2_ξ from equation (6);

3). the parameters in the earnings equations (A.2 and A.3 in Appendix A);

4). the parameters of the health expenditure equation (A.1 in Appendix A);

5). τ, λ_{HI}, and $\{\lambda^j_X, \lambda^j_\eta$:j=C,B,A} from the current period utility equation (3);

6). the parameters Π_1 and Π_2 from the initial conditions equation (12);

7). the standard deviation of unobserved heterogeneity σ_κ;

8). The covariance parameters {COV(μ_t, ϵ_t^j), j=C,B,A} discussed after equation (10);

9). the covariance parameters {COV(ϵ^I,ν_1) and COV(ϵ^I ,κ), j=C,B,A} discussed after equation (12);

10). the parameter β^{Common} from the discount factor equation (6); and

[24] Persons are classified as having applied for disability benefits are those who apply at any time between the 1992 and 1998 surveys. Persons who have not applied for disability benefits and are still working but who have changed jobs since the baseline survey are classified as "Bridge." Those alive and not working in the final survey who have not applied for disability benefits are classified as "Retired."

11). the parameter θ from the pecuniary utility function in equation (1).

The identification of the DI approval equation (10) is made difficult in practice by the reality that the DI approval decision is only observed for those who apply for benefits and virtually all DI applicants have poor or fair self-reported health. While our model has features that in theory can address this type of problem, the reality that only a relatively small number of individuals apply for DI during our sample period makes identification difficult in practice. For our structural estimation we set A_1^{DI}=.08 and $A_2^{DI}=Var(e^{DI})=1$ in equation (7). The assumption that $A_2^{DI}=Var(e^{DI})=1$ amounts to assuming that DI award decisions are about as equally reliable indicators of disability status as are the global self-reported measures available in the HRS, which is consistent with the work of Benitez-Silva, Buchinsky and Rust (2006). We chose our estimate of A_1^{DI} so that the fraction of those that apply for benefits who are awarded them implied by our model matches the fraction in the data, roughly 2/3.

In Appendix D we describe the technical/computing steps used to address a very time-intensive estimation problem. In part due to the importance of having a likelihood function with certain properties that make estimation feasible and in part as a concession to the size of the computing task we reduced the number of parameters in the model by estimating the parameters of the earnings equations (#3 above with estimates shown in Appendix A.3.2) and health equations (#2 above with estimates shown in Table 4) outside of the behavioral model.[25] To estimate the health parameters (# 2 above) using Maximum Likelihood, we specify the likelihood contribution for a person in a way similar to equation (C.7) in Appendix C but we include in the probability expression only the conditions involving aggregate unobservables of the form $\nu_t+\mu_t$ which are the terms related to self-reported health (i.e., in illustrative equation (C.7) in Appendix C this would involve the first three terms in the intersection, $\nu_2+\mu_2<-\pi X_2-\gamma Z_2$, $\nu_3+\mu_3<-\pi X_3-\gamma Z_3$, $\nu_5+\mu_5>-\pi X_5-\gamma Z_5$). π and γ enter these

[25]A primary motivation for estimating the parameters of the earnings equation outside the structural model is that Defined Benefit pension payments are calculated using a computer program provided by the Health and Retirement Study. The reality that this program cannot be used interactively with our estimation program implies that all DB payments, which are a function of individual's earnings (or, equivalently, years of career experience in our model), must be calculated and stored prior to estimation. Estimating the health equation outside the model has certain benefits from the standpoint of ensuring that our likelihood function has properties that are necessary to use derivative-based updating algorithms

conditions directly. The remainder of the health parameters (in #2 above) affect the likelihood contribution through their effect on the covariance matrix of the aggregate unobservables. Given the discrete nature of the self-reported health variable, once we make the normalization $var(\nu_t+\mu_t)=1\ \forall t$, it is possible to identify the covariance between each pair of aggregate unobservables. In addition, the fact that we observe information about h_t^* at multiple times implies that it is possible to separately identify $var(\nu_t)$ and $var(\mu_t)$. What allows the separate identification is that fact that μ_t has only a transitory effect on measured health, while ν_t has an effect which dies out over time only slowly.

An additional concern was the difficulty of credibly identifying and estimating β^{Common} and θ. In response to this concern we take the approach of seeking guidance from recent literature in order to choose reasonable values of these two parameters, and we estimate the thirty-six remaining parameters described above.[26] We begin by estimating a "baseline" specification in which $\beta^{Common}=0.90$ and $\theta=1$ which implies that $Y^j(t)^{1-\theta}/(1-\theta)=\ln(Y)$.[27]

The first column of Table 5 shows the estimates of the behavioral equations for the baseline specification.[28] The estimate of τ indicates that the amount of consumption available from a particular option plays a statistically significant role in the utility that is derived from that option. The estimates in the first column of Table 5 also indicate that health plays a statistically significant role. Given that larger values of health represent worse health, the negative estimates of λ_η^C and λ_η^B indicate that individuals in bad health get less utility from the work options (relative to the option N) than individuals in better health. The positive estimate of λ_η^A indicates that individuals in bad health get higher non-pecuniary benefits from applying for Disability Insurance (relative to the option N)

[26] See Magnac and Thesmar (2002) for a discussion of the difficulties related to the estimation of the discount factor. Given these difficulties, researchers often fix the discount factor at a seemingly reasonable number (Berkovec and Stern 1991). As a concession to the small number of individuals who apply for DI at t=1, we make the assumption that the covariance terms described above are the same regardless of the reason that a person is out of the workforce. That is, we assume that $COV(\mu_t, \epsilon_t^A)=COV(\mu_t, \epsilon_t^N)=0$.

[27] In a longer version of this paper (Bound, Stinebrickner & Waidmann 2007), we also examine the robustness of our results to changes in β^{Common} and θ.

[28] Although not shown in Table 5, we also included two dummy variables characterizing a person's education level (less than high school and more than high school).

than individuals in better health. In addition, the estimates imply that individuals, especially those in decent health, face a very significant non-pecuniary cost associated with applying for DI benefits.[29] Our estimates imply that, over and above the effect of health insurance on disposable income, health insurance has a positive relationship with well-being in the sample, but the estimated effect is quite small and not statistically significant. Among the variance/covariance estimates, most striking is the importance of unobserved heterogeneity; the point estimate (standard error) of the standard deviation σ_κ is 1.423 (.322).

In order to quantify the roles that economic resources and health play in determining labor decisions we begin by performing simulations using an "illustrative person." We construct a person who has a college education and has career earnings, bridge earnings, SS benefits, and potential DI benefits that are close to the average for people in our sample, but has no private pension wealth or other sources of wealth.[30] We first assume that the illustrative person has true health η_1 at time t=1 that is equal to the average true health of the individuals in our sample. The first column of Table 6 shows simulated choice probabilities at t=1 for an illustrative person at 55, 60, 62, 64, and 65 years of age, respectively. Since the simulated individuals would have been employed at t=0, these simulated probabilities can be thought of as one year labor force exit rates.

The simulated choice probabilities at age 55 and age 60 are quite similar. At these ages, the illustrative person's only economic resources if he leaves the workforce come from assistance programs such as the food stamp program. This reality, combined with the fact that being in average health implies that it is not particularly unenjoyable to work and that applying for Disability Insurance is not particularly worthwhile, implies that the person at age 55 and age 60 has a very low probability of leaving the workforce for either the non-work option N (0.02) or for the option of

---

[29] In our behavioral sample the mean of health is -0.904, while the standard deviation is 0.802. These numbers, together with the estimates reported in Table 5 imply that virtually all individuals in our sample face a non-pecuniary cost associated with apply for DI. These costs hit zero when an individuals health is 4.2 standard deviations worse than the average in the sample!

[30] Our illustrative person would receive $15,588 in SS benefits at age 65 (given amount of contributions as of time t=1), and would receive DI benefits of $15,588 if approved for the program.

applying for Disability Insurance A (0.001). Evidence regarding the effect of economic resources on behavior can be seen by comparing the choice probabilities at the age of 60 to the choice probabilities at the age of 62 at which time the person becomes eligible for Social Security retirement benefits. The consumption increase in the non-work option (N) causes the probability of choosing this option to increase by a factor of approximately two (from 0.024 to 0.046). The fact that the probability of applying for DI remains extremely small (0.003) for the average health person even when SS benefits become available is evidence of the very strong importance of health in the DI application decision.[31] Delaying retirement past the age of 62 increases a person's Social Security benefits. Comparing simulated choice probabilities between the age of 62 and the ages of 64 and 65 reveals that this increase in Social Security benefits has a relatively small effect on retirement decisions.

The second column of Table 6 shows choice probabilities at different ages for the illustrative person under the assumption that his health at t=1 is one standard deviation below average.[32] As before, the choice probabilities are fairly similar at ages 55 and 60 for this person. However, comparing the choice probabilities for this person at ages 55 and 60 to the choice probabilities for the person in average health at ages 55 and 60 indicates that health has a very important effect on the probability that a person will transition out of the workforce at ages 55 and 60. For example, the total probability of leaving the workforce (N+A) at age 60 is 0.025 for the person in average health and is 0.106 for the below-average health person with the impact of worse health coming from both an increase in the probability of choosing the non-work category (N) and an increase in the probability of applying for DI (A). Comparing these two numbers with the choice probabilities of

---

[31]Before age 62, a person with no outside wealth who applies for DI must rely on social assistance such as food stamps while waiting for the approval decision. However, at age 62 a person can receive SS benefits while waiting for the DI approval decision.

[32]As is clear from the discussion in Appendix A.1, a large share of those in poor health as of the initial HRS survey year do not make it into our behavioral sample. Thus, a large share of those in poor health as of t=1, would not have been in poor health two years earlier and would have suffered a major negative health shock in the interim. Thus, while lagged health does not enter the behavioral model, it probably still makes sense to interpret the results in terms of the behavioral effects of the deterioration in health status, rather than the effects of permanently poor health.

the individuals in average and below-average health at age 62 indicates that the incentive effect of economic resources depends to some extent on health. At age 62 when SS benefits become available, the total probability of leaving the workforce increases by approximately 0.06 (from 0.106 to 0.168) for the person in below-average health but increases by only approximately 0.02 (from 0.025 to 0.049) for the person in average health.

The third column of Table 6 shows choice probabilities at different ages for the illustrative person under the assumption that his health is 1.5 standard deviations below average. Comparing the results of the second and third columns indicates that, once a person reaches poor health, an incremental worsening of health can have large effects on decisions. When compared to the person who has health 1 standard deviation below average, the person in worse health is at least twice as likely to apply for Disability Insurance at each age and has a probability of leaving work (N+A) that is approximately 10 percentage points higher at each age.

In a previous version of this work (Bound, Stinebrickner and Waidmann 2007), we also examined the robustness of these estimates to the assumptions made about time preferences and risk aversion. In general, varying these parameters produced results that were consistent with the baseline specification.

Does the treatment of health matter? - A comparison to a model trusting self-reported health

Our models deal with the measurement error and endogeneity problems that are potentially present in self-reported, survey health measures. Given that the implementation of the health portions of our models is non-trivial, and, as a result, has implications for the feasibility of estimating other aspects of the decision process, it is valuable to examine the extent to which our findings would differ from a model in which these potential problems were ignored.

To do this, we estimate a new version of our model in which our health framework has been replaced with a treatment of health that is consistent with the standard in the literature. Specifically, in the current period utility equation (2) of our baseline specification we replace our continuous measure of health, η_t, with a binary health variable that is equal to one if the person is in bad health and is equal to zero otherwise. We assume that the person's self-report of this binary health variable is accurate so that, unlike the continuous health measure that enters our models, the binary health measure that enters this new model is observed directly in the HRS data.[33] In order to compute the new value functions of the type in equation (4), we make the standard assumption that a person knows his current period binary health status, and, given this status, knows the probability of each possible health status arising in period t+1. To be consistent with the self-reported health measure that we use to estimate equation (10), we characterize a person to be in bad health if he reports that he has an impairment or health problem that limits the kind or amount of paid work he can do.[34]

The second column of Table 5 reports the estimates of this "traditional" model. The estimate of τ is similar to the estimate shown in the first column of Table 5 for the baseline specification of our continuous-health model. Consistent with the estimates of λ_{η}^{C}, λ_{η}^{B}, and λ_{η}^{A} in the continuous-health model, the estimated effects of the binary health variables in the traditional model are statistically significant in each of the current period utility equations. However, the t-statistics associated with the estimated effects of health are non-trivially different between the models. For example, the t-statistic associated with the effect of health in the option C is greater by a factor of more than three in the traditional health model than it is in the baseline specification of our baseline continuous-health model. While these differences in statistical significance raise the possibility that ignoring the measurement and endogeneity issues related to self-reported health may lead to different conclusions about the importance of health, knowing whether this is actually the case cannot be confirmed simply by examining individual coefficients in isolation. Therefore, we take the approach of comparing the health implications of the two models by examining whether simulated choice probabilities from our baseline model are different than simulated choice probabilities from the traditional model.

---

[33]To be more exact, a person's health status is observed in each survey year but is not observed in non-survey years. We deal with this issue using standard MLE missing data techniques which involve integrating over the joint probability function of the non-observed health values.

[34]In our sample, the fraction of individuals in good health who transitioned to poor health was .096, while the fraction in poor health who stayed in poor health was .684. We found no statistically significant evidence that these transition probabilities varied with the age of the respondent of the years of the survey used. We converted these two year transition rates into one year transition rates using the assumption that these transitions represented a Markov process.

To do this, we again make use of the illustrative person that was used to produce the simulations in Table 6. For the traditional model there are only two possible health values. In In the first two columns of Table 7 we show choice probabilities at different ages for the illustrative person who reports himself to be in good and bad health, respectively. A comparison of these two columns reveals that changing the self-reported health of the illustrative person from good to bad has very large effects on behavior. For example, the probability that the illustrative person will be out of the workforce at age 60 is .094 if the illustrative person reports being in good health and is .426 if the illustrative person reports being in bad health.

The goal is to compare the numbers from the traditional model to analogous numbers constructed from our continuous health model. It is not trivial to construct choice probabilities for the illustrative person who *reports* being in good or bad health in our continuous health model because this model is written in terms of true rather than reported health. To construct these probabilities, we need to account that, in our model, there is a continuum of health values, η, with a distribution that is characterized by equation (8) and that the connection between a particular value of continuous health, η, and the person's self-reported binary health is provided by equations (9) and (10) which determine the probability that a person with a particular value of η will report himself to be in bad health or good health. Thus to construct the analog to the first two columns, we integrate the choice probabilities that can be constructed for the illustrative person for a particular value of η over the person's distribution of η, where we weight the choice probabilities in the integral to reflect the probability that a person would report himself to be in bad health given η. The results of this exercise shown in the third column represent the average choice probabilities of the illustrative person over all health scenarios in which he would report good health. Similarly, the results in the fourth column represent the average choice probabilities of the illustrative person over all health scenarios in which he would report bad health.

A pair-wise comparison across the first four columns provides evidence of substantial differences between the traditional model and our model. For example, for the traditional health model, we see the difference between good and bad self-reported health is a .33 difference in the probability of not working at age 60. In our continuous health model, we see a .16 difference in the probability of working at age 60 between a person who reports himself to be in good health and a person who reports himself to be in bad health. Thus, our results suggest that dealing with the potential problems associated with self-reported health is important in this context.

It is not an easy task to pinpoint exactly why these differences exist. Somewhat generically, the difference between the two models arises to a large extent because our continuous health model is able to take into account that reported differences in health should be attributed to both true differences in health and a non-trivial amount of reporting error. More concretely, we find evidence that a large part of the difference between the traditional model and our model operates through differences in the importance of unobserved heterogeneity between the models. Specifically, we find in the last two columns of Table 7 that if we reestimate and resimulate our continuous model holding the standard deviation of heterogeneity, σ_κ , constant at its value from the traditional model, a little more than half of the aforementioned difference between the traditional model and our model disappears. For example, we see that there is now a .26 difference in the probability of working at age 60 between a person who reports himself to be in good health and a person who reports himself to be in bad health. It is necessary to leave a detailed exploration of why substantial differences exist in the importance of heterogeneity between the traditional and continuous models to future work. What we note here is that a stark implication of moving from the traditional model (in which true health is observed exactly) to the continuous model (which recognizes that only the distribution of true health can be observed) is that σ_κ is estimated less precisely. This finding highlights the importance of work such as Kasahara and Shimotsu (2007) that focuses on understanding the identification of finite mixture models in dynamic, discrete choice models.

Potential Changes in Policy

The simulations involving the illustrative person suggest that changes in policy that influence economic resources may have substantial effects on individual behavior and that these effects may vary across people with different health. Here we use our baseline model and our behavioral sample to simulate the effects of several potential changes in policy. Policy 1 examines the effect of removing the option of early SS benefits. Policy 2 examines the effect of a policy that has been implemented - changing the normal retirement age from 65 to 67. Policy 3 examines the impact of removing the Disability Insurance program entirely. Policies 4, 5, and 6 examine, respectively, a 25% reduction in SS benefits, a 25% reduction in DI benefits, and a 25% reduction in both SS and DI benefits.

To quantify the effects of these policy changes, we first perform a baseline simulation in which no policy change has occurred. For each person in our behavioral sample (i.e., those individuals working at time t=0), we condition on the person's information that is available at time t=0 and simulate a sequence of choices corresponding to the years that a choice is observed for the person in the data.[35] We then repeat this process 7200 times for each person in our behavioral sample. The resulting simulations can be used to compute the proportion of individuals that would choose each of the options in the set {C, B, N, A} at particular ages and the proportion that would choose each of the options when all of the ages are pooled. The results are shown in the first column of Table 8A.

The first entries in Column 1 represent the choice proportions that are generated if individual choices at all ages are pooled. [36] Thus, under the baseline specification, our model indicates that individuals will choose the work option (C+B) in 1-.152=.848 of the periods in the pooled decision periods, will choose the non-work option (N) in .134 of the pooled decision periods, and will choose to apply for Disability Insurance (A) in .019 of the pooled decision periods. To get a sense of the fit of our model, recall from Section 4.3 that the actual proportions in the data are, respectively, .830 (C+B), .153 (N), and .018 (A). The remainder of the entries in Column 1 reflect the choice

---

[35]Thus, we take the approach of simulating choices within the sample period. The likelihood contribution for a person in Section 3.3 and Appendix C assumes that a person who applies for DI and is accepted makes no subsequent decisions (i.e., he remains on DI forever). For these simulations we assume that a person makes choices in all data years (i.e., in essence for the simulations we assume that individuals who apply for DI benefits (A) in some year t do not get approved and we simulate choices for the person in years subsequent to the DI application. The motivation for this simulation approach is that it implies that the set of people making decisions at particular ages is identical across all policies that we simulate even if the policies influence how many people apply for DI benefits.

[36]Thus, we are creating these entries by using all behavioral sample members in all years that choices are observed. Recall that the youngest age at which any individual in the sample is observed making a choice is 51 and the oldest age at which any individual in the sample is observed making a choice is 66. Each individual contributes simulated choices at the subset of these ages which corresponds to the years choices are observed for him in the data.

proportions when choices are disaggregated by age (for select ages). We note that, while from an operational standpoint it would be possible to compare these disaggregated simulated proportions to the actual disaggregated proportions in the data, in practice the usefulness of a full comparison of this type is limited due to relatively small sample sizes at individual ages. However, it is certainly worth noting that the general message from such a comparison is that the age gradient in the simulations is substantially less steep than the comparable age gradient in the actual data. For example, while the proportion of men working falls from roughly .89 to .54 between the ages of 55 and 65 in the data, our simulations (using our model which does not allow a person's age to influence his non-pecuniary utility) show a decrease from .85 to .78 between these ages. Thus, roughly speaking, at least in our model, health and economic resources can explain about 20% of the decline in work between the age of 55 and 65.

For each policy change, the simulation process is repeated after modifying the model appropriately to reflect the change. Column 2 of Table 8A shows the proportions associated with Policy 1 in which no benefits are available from the Social Security system until a person reaches the age of 65. When a person is younger than 62, this policy change influences decisions only through its influence on future income. Because knowledge that Social Security benefits will not be available at the ages of 62-64 tends to reduce the value of each option in a somewhat similar fashion, it is perhaps not surprising that the policy change has little effect before the age of 62. For example, at age 55, the proportion choosing the non-work option (N) is .123 under the baseline simulation and .121 under the policy change. The policy change leads to an increase in work at the age of 62 when the amount of current period consumption that a person receives in the non-work and DI options is reduced relative to the baseline case; at age 62 the proportion choosing the non-work option falls by approximately 15% (from .155 under the baseline simulation to .132 under the policy change). A similar effect (.176 versus .158) is shown in the table for age 64. The policy change leads to little change in the number of DI applicants. Thus, the decrease in the proportion choosing N is accompanied by an increase of similar size in the proportion choosing C or B. At 65 when Social Security benefits become the same in the baseline simulation and policy change simulation, the proportion choosing option N becomes similar under the baseline and policy change (.220 vs. .214).

We also construct the analog to Table 8A for individuals with health that is one standard deviation or more below the average for the sample (hereafter referred to as "bad" health) and for the remaining individuals in our sample (hereafter referred to as individuals in "good" health) and show these proportions in Table 8B and Table 8C, respectively.[37] Consistent with our illustrative person simulations, we find that individuals in bad health are much less likely to be working in all periods than individuals in good health. Also consistent with our illustrative person simulations, we find some evidence that the effect of this policy change varies with a person's health. For example, a comparison of the first and second columns of Table 8B reveals that, for individuals in our sample with bad health, the policy causes the proportion choosing N at age 62 to fall by approximately .04 (from .351 under the baseline to .310 under Policy 1), while the proportion applying for DI rises by 0.014. Thus, our model implies that among those in poor health affected by the elimination of early retirement benefits, roughly one third would apply for DI, while two thirds would continue to work. A comparison of the first and second columns of Table 8C reveals that, for individuals in our sample with good health, the policy causes the proportion choosing N at age 62 to fall by approximately .02 (from .137 under the baseline to .116 under Policy 1). For this group, there is essentially no effect of the policy change on the number of men applying for DI.

The third column in this set of tables (Policy 2) shows the proportions associated with a second policy which influences Social Security benefits in a less drastic way than the first policy. Specifically, the policy change involves increasing the normal retirement age from 65 to 67 and allowing individuals to apply for DI benefits until the age of 67. Under this policy, while individuals still become eligible for Social Security benefits at the age of 62, the amount of the

[37]True health is not observed in our data. However, our model produces the distribution of health for a particular person. This can be used to compute the probability that a person is in "good" or "bad" health.

benefits that is received if one retires at age 62 is reduced by 12.5%.[38] This change in the normal retirement age is currently being phased in. A comparison of Column 1 and Column 3 of Table 8A reveals that this policy will have relatively small effects on individual decisions. The change has virtually no effect before age 62. At ages 62 and 64, the policy causes a decrease in the proportion choosing the non-work option (N) of one percentage point or approximately seven percent for people in our sample.[39]

The fourth column of Table 8A (Policy 3) shows the proportions associated with a third policy change in which the Disability Insurance program is removed entirely. Comparing the "Pooled Ages" entries in Column 1 to those in Column 4 reveals that, when the DI program is removed, slightly less than half of the individuals who were DI applicants under the baseline simulation choose to work.[40] Evidence about the differential effect of the policy change by health status can be seen in Tables 11B and 11C. The "Pooled Ages" entry in Column 1 of Table 8C shows that, under the baseline simulation, individuals in good health choose to apply for DI benefits (A) in only .007 of the pooled decision years. Thus, removing DI benefits has very little effect on individuals in good health. The "Pooled Ages" entry in Column 1 of Table 8B shows that, under the baseline simulation, individuals in bad health choose to apply for DI (A) benefits in .104 of the pooled decision years. For these individuals, comparing the "Pooled Ages" entries under the baseline (Column 1 of Table 8B) to that under Policy 3 (Column 4 of Table 8B) shows that the removal of the DI option results in an increase in the proportion choosing N from .296 to .362 and a smaller change in the proportion choosing C+B (from .600 to .637). Thus, to a large extent, the differential impact of this policy change by health status comes from the fact that many individuals in bad health who

---

[38]For cohorts born before 1938, for whom the normal retirement age was 65, benefits were reduced by 5/9 of 1% for each month prior to age 65 a person retired, for a maximum of 20% for a person who retired when they reached the age of 62. With the rise in the normal retirement age to 67, the reduction will rise to 30% for someone retiring at the age of 62 (SSA 2007).

[39]French and Jones find similar small effects of the increase in the Social Security Normal Retirement age.

[40]The proportion of DI application in the baseline is .019. When the DI program is removed, C+B increases from .848 to .856 and N increases from .134 to .144.

apply for DI benefits under the baseline remain out of the workforce but no longer receive income from DI benefits when the program is removed.

The last three columns of Tables 11A, B & C report results from simulating what would happen if first just DI benefits, then just Retirement benefits, and then both together were reduced by 25%. Although not shown, the results from simulating what would happen were benefits increased by 25% were very similar. Our estimates suggest that a 25% reduction in DI benefits would reduce applications for DI by a little over 20%.[41] The effect of reducing Retirement benefits by 25% varies by age. For those who have reached the Social Security early retirement age of 62, there is about a 7% reduction in the fraction of men simply leaving the labor force (N).[42] One can also see in this table a clear evidence of interaction effects between the Social Security retirement and Disability Programs; when retirement but not disability benefits are reduced, the application for disability benefits increases, with the reverse being true when disability, but not retirement benefits are reduced.

[41]The implied application elasticity is quite close to Kreider's (1999b) estimate, but somewhat larger than early estimates based on time series data. See Bound and Burkhauser (1999) for a discussion of existing estimates.

[42]Van der Klaauw and Wolpin (2005) also report only modest effects of a 25% change in benefits.

Section 5. Conclusion

In this paper, we report estimates of a dynamic programming model that addresses the interplay among health, financial resources, and the labor market behavior of men nearing retirement age. A significant contribution of our work is the use of a latent variable model for health that we believe is robust to concerns about the endogeneity of self-reported health. Our comparison of results obtained using this health model to those obtained using a more traditional measure of health suggest that the manner in which one models health can have a substantial effect on conclusions about the behavioral effects of poor health.

Our estimates imply that individuals in good health are unlikely to retire unless they have generous financial resources available to them. On the other hand, our estimates imply that a man in poor health is quite likely to leave the workforce even when he is not yet eligible for any kind of pension benefits. In fact, our simulations show that a typical individual in poor health is 10 times more likely than a similar person in average health to retire before becoming eligible for pension benefits. These estimates underline the importance that health plays in determining early retirement behavior.

Strikingly, our estimates imply that changes in the Social Security Retirement Program are likely to have quite small effects on applications for the Disability Insurance Program.[43] We suspect that the reason for this has to do with the fact that those potentially eligible for DI are a quite a distinct population. Our findings have strong predictions about the patterns we might see in the application for DI benefits as the age for normal retirement under Social Security rises over the next decade. Despite the fact that this change will substantially increase the financial rewards associated with receiving DI rather than early retirement benefits, our estimates suggest that the number of individuals over the age of 62 who apply for DI will not rise by much.

[43]As far as we know the only other attempt to estimate the impact of changes in the Social Security Retirement Program on Disability participation is the work by Mitchell and Phillips (2000). Mitchell and Phillips use a conditional logit framework to study the effect of financial incentives on the probability that a person will retire early or apply for disability insurance. In qualitative terms, Mitchell and Phillips results are similar to ours. They estimate that eliminating Social Security early retirement benefits would have only modest effects on behavior.

References

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, 375-380.

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.

Angeletos, G-M., D. Laibson, A. Repetto, J. Tobacman, and S. Weinberg (2001): "The Hyperbolic Buffer Stock Model: Calibration, Simulation, and Empirical Evaluation," *Journal of Economic Perspectives*, 15(3), Summer, 47-68.

Baltes, P.B., and M.M. Baltes (1990): "Psychological Perspectives on Successful Aging: a Model of Selective Optimization with Compensation," in *Successful Aging: Perspectives from the Behavioral Sciences* , ed. by P.B. Baltes and M.M. Baltes,(pp. 1-34) New York: Cambridge University Press.

Bellman, R. (1957): *Dynamic Programming.* New Jersey: Princeton University Press.

Benitez-Silva, H., M. Buchinsky, and J. Rust (2006): "How Large Are the Classification Errors in the Social Security Disability Award Process?" Working paper, Department of Economics, Stony Brook University, http://ms.cc.sunysb.edu/~hbenitezsilv/research.html.

Berkovec, J., and S. Stern (1991): "Job Exit Behavior of Older Men," *Econometrica*, 59(1), 189-210.

Blau, D. and D. Gilleskie (2001a): "Retiree Health Insurance and the Labor Force Behavior of Older Men in the 1990s," *Review of Economics and Statistics*, 83(1), 64-80.

Blau, D.M., and D.B. Gilleskie (2001b): "The Effect of Health on Employment Transitions of Older Men"in *Worker Wellbeing in a Changing Labor Market, Research in Labor Economics*, 20, ed. by S. Polachek, Amsterdam:JAI Press.

Blau, D.M., and D.B. Gilleskie, (2006), "Health Insurance and Retirement of Married Couples," *Journal of Applied Econometrics,* 21:935-953.

Blau, D.M., and D.B. Gilleskie (Forthcoming): "The Role of Retiree Health Insurance in the Employment Behavior of Older Men," *International Economic Review*.

Blinder, A.S., R.H. Gordon, and D.E. Wise (1980): "Reconsidering the Work Disincentive Effects of Social Security," *National Tax Journal*, 33(4), 431-442.

Bound, J. (1989): "The Health and Earnings of Rejected Disability Insurance Applicants," *American Economics Review*, 79, 482- 503.

Bound, J. (1991): "Self-reported Versus Objective Measures of Health in Retirement Models," *The Journal of Human Resources*, 26, 106-138.

Bound, John and Richard Burkhauser, 1999, "Economic Analysis of Transfer Programs Targeted on People with Disabilities." Chapter in the *Handbook of Labor Economics*, V. 3, eds. D. Card and O. Ashenfelter, pp. 3417-3528.

Bound, J., M. Schoenbaum, T. Stinebrickner, and T. Waidmann (1998): "Modeling the Effect of Health on Retirement Behavior," paper presented at the National Bureau of Economics Summer Institute.

Bound, J., M. Schoenbaum, T. Stinebrickner, and T. Waidmann (1999): "The Dynamic Effects of Health on the Labor Force Transitions of Older Workers," *Labour Economics*, 6(2), 179-202.

Bound, J., M. Schoenbaum, and T. Waidmann (1997). "Health Limitations and Early Retirement," HRS-AHEAD Research Brief #004. Ann Arbor: Institute for Social Research.

Bound, J. and T. Waidmann (1992): "Disability Transfers and the Labor Force Attachment of Older Men: Evidence from Historical Record,". *Quarterly Journal of Economics*, 107, 1395-1419.

Bound, J., and T. Waidmann (2002): "Accounting for Recent Declines in Employment Rates among Working-Aged Men and Women with Disabilities," *Journal of Human Resources*; 37(2), 231-50.

Bound, J., and T. Waidmann (2007): "Estimating the Health Effects of Retirement" paper presented at the , The 9th annual conference of the Retirement Research Consortium (RRC), Washington DC.

Brim, O.G. (1988): "Losing and Winning," *Psychology Today*, 48-52.

Burkhauser, R.V. and J.F. Quinn (1983): "Is mandatory retirement overrated? Evidence from the 1970s," *Journal of Human Resources*, 18(3), 337-358.

Chirikos, T.N. (1993): "The Relationship Between Health and Labor Market Status," *Annual Review of Public Health*, 14, 293-312.

Currie J., and B.C. Madrian (1999): "Health, Health Insurance, and the Labor Market," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, New York: Elsevier Science Publishing Company.

Curtin, R.T., J. Lamkin, and R. Peticolas. (1998): "Pension Estimation Program Documentation." HRS Documentation Report DR-004 http://hrsonline.isr.umich.edu/docs/userg/dr-004.pdf.

Diamond, P. A. and Hausman, J. H. (1984): "The Retirement and Unemployment Behavior of Older Men." In *Retirement and Economic Behavior* ed. by H. Aaron and G. Burtless, Washington, D. C.: The Brookings Institution.

Diermeier, D., M.P. Keane, and A. Merlo (2002): "The Career Decisions of Politicians," Unpublished Manuscript.

Disney, R., C. Emmerson and M. Wakefield (2006): "Ill Health and Retirement in Britain: A Panel Data-based Analysis," *Journal of Health Economics*, 25:621-649.

French, E. (2005): "The Effects Of Health, Wealth, And Wages On Labor Supply And Retirement Behavior,"*Review of Economic Studies, 72:395-427.*

French, E., and J. B. Jones (2007): "The Effects of Health Insurance and Self-Insurance on Retirement Behavior,"Chicago Fed WP2001-19.

French, E., and J.B. Jones (2004): "On the Distribution and Dynamics of Health Costs" *Journal of Applied Econometrics*, 19:705-721.

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.

Gordon, R. H. and Blinder, A. S. (1980): “Market Wages, Reservation Wages and Retirement Decision,” *Journal of Public Economics*, 14, 277-308.

Gruber, J., and D. Wise (eds.) (1999): *Social Security and Retirement Around the World*, University of Chicago Press: Chicago.

Gruber, J., and D. Wise (eds.) (2004): *Social Security Programs and Retirement Around the World: Micro Estimation*, University of Chicago Press: Chicago.

Gustman, A. L. and Steinmeier, T. (1986): “A Structural Retirement Model,” *Econometrica*, 54(3), 555-584.

Gustman, A. L. and Steinmeier, T. L. (1994): “Employer Provided Health Insurance and Retirement Behavior,” *Industrial Labor Relations Review,* 48, 124-140.

Gustman, A. L. and Steinmeier, T. L. (2000): “Retirement in Dual-Career Families: A Structural Model,” *Journal of Labor Economics,* 18(3), 503-545.

Gustman, A. L. and Steinmeier, T. L. (2005): “Social Security Early Entitlement Age in a Structural Model of Retirement and Wealth,” *Journal of Public Economics,* 89, 441-463.

Hajivassiliou, V. (1990): “Smooth Simulation Estimation of Panel Data LDV Models,” Unpublished Manuscript.

Halpern, J., and J.A. Hausman (1986): “Choice under Uncertainty: A Model of Applications for the Social Security Disability Insurance Program,” *Journal of Public Economics*; 31(2), 131-161.

Haveman, R., P. de-Jong, and B. Wolfe (1991): “Disability Transfers and the Work Decision of Older Men,” *Quarterly Journal of Economics*; 106(3), 939-49.

Hill, D.H. (1993): “Respondent Rules and the Quality of Net Worth Data in the HRS,” PSC Publications. HRS Working Paper 94-002.

Honig, M. (1985): “Partial Retirement Among Women,” *Journal of Human Resources*, 20, 613-621.

Honig, M and G. Hanoch (1985): “Partial Retirement as a Separate Mode of Retirement Behavior,” *Journal of Human Resources*, 20, 21-46.

Honig, M., and C. Reimers (1987): “Retirement, Re-entry, and Part-time Work,” *Eastern Economic Journal*, 13, 361-371.

Juster, F.T., and R. Suzman (1995): “An Overview of the Health and Retirement Study,” *Journal of Human Resources*, 30 Suppl., S7-S56.

Kasahara, H. and K. Shimotsu. “Nonparametric identification of finite mixture models of dynamic discrete choices,” revised and resubmit, Econometrica, 2007.

Keane, M. P. (1994): "A Computationally Practical Simulation Estimator for Panel Data," *Econometrica*, 62(1), 95-116.

Keane, M, and R. Moffitt (1998): "A Structural Model of Multiple Welfare Program Participation and Labor Supply," *International Economic Review*, 39 (3), 553-89.

Keane, M.P., and K.I. Wolpin (1994): "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *Review of Economics and Statistics*, 76, 648-672.

Keane, M., and K. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy,* 105, 473-522.

Kreider, B. (1999a): "Latent Work Disability and Reporting Bias," *Journal of Human Resources*; 34, 734-769.

Kreider, B. (1999b): "Social Security Disability Insurance: Applications, Awards and Life Time Income Flows," *Journal of Labor Economics*, 17:784-827.

Kreider, B., and R. Riphahn (2000): "Explaining Applications to the U.S. Disability Program: A Semiparametric Approach," *Journal of Human Resources*, 82-115.

Lavy, V., M. Palumbo, and S. Stern (1998): "Simulation of Multinomial Probit Probabilities and Imputation," in *Advances in Econometrics*, ed. by T. Fomby and R. Carter Hill, JAI Press.

Lindeboom, M., and M. Kerkhofs, (2004): "Health and Work of the Elderly," unpublished manuscript, Free University of Amsterdam.

Lumsdaine R.L., O.S. Mitchell (1999): "Retirement," in *Handbook of Labor Economics,* ed. by O. Ashenfelter and D. Card, New York: Elsevier Science Publishing Company.

Magnac, T., and D. Thesmar (2002): "Identifying Dynamic Discrete Decision Processes", *Econometrica*, 70, 801-816.

Mitchell O. and J.W.R. Phillips (2000): "Retirement Responses to Early Social Security Benefit Reductions," NBER Working Paper 7963.

Myers, R.J., (1982): "Why Do People Retire Early from Work?" *Aging and Work*, 5,83-91.

Olin, G.L. and S.R. Machlin (2004): *Chartbook #11: Health Care Expenses in the Community Population, 1999.* February 2004. Agency for Healthcare Research and Quality, Rockville, MD. http://www.meps.ahrq.gov/papers/cb11/cb11.htm.

Parsons, D.O. (1980): "The Decline of Male Labor Force Participation," *Journal of Political Economy*, 88, 117-134.

Parsons, D.O. (1982): "The Male Labor Force Participation Decision: Health, Reported Health, and Economic Incentives," *Economica*, 49, 81-91.

Quinn, J. F. (1977): "Microeconomic Determinants of Early Retirement: a Cross-sectional View of White Married Men," *Journal of Human Resources*, 12(3), 329-347.

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*, ed. by D. Salisbury. Employee Benefit Research Institute, 25-39.

Quinn, J.F. (2000): "New Paths to Retirement,"in *Forecasting Retirement Needs and Retirement Wealth*, ed. by B. Hammond, O. Mitchell, and A. Rappaport, Philadelphia, PA: University of Pennsylvania Press, 13-32.

Quinn, J.F., and M. Kozy (1996): "The Roles of Bridge Jobs in the Retirement Transition: Gender, Race, and Ethnicity," *The Gerontologist*, 36 (3), 363-372.

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.

Ruhm, C.J. (1990): "Bridge Jobs and Partial Retirement," *Journal of Labor Economics*, 8, 482-501.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999-1033.

Rust, J. (1997): "Using Randomization to Break the Curse of Dimensionality," *Econometrica*, 65, 781-832.

Rust, J., and C. Phelan (1997): "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets," *Econometrica*, 65(4), 781-831.

Rust, J., M. Buchinsky and H. Benitez-Silva (2001): "An Empirical Model of Social Insurance at the End of the Life Cycle." ftp://gemini.econ.umd.edu/pub/johnrust/disability/nsfpro.pdf.

Smith, J. P. (1995): "Racial and Ethnic Differences in Wealth in the Health and Retirement Study," *Journal of Human Resources*, Vol. 30, S158-83.

Social Security Administration. (1995): *Social Security Bulletin: Annual Statistical Supplement, 1995.* Washington, DC: Social Security Administration.

Social Security Administration. (2007): *Annual Statistical Supplement to the Social Security Bulletin, 2006.* Washington, DC: Social Security Administration.

Stinebrickner, T.R (1999): "Using Latent Variables in Dynamic, Discrete Choice Models: The Effects of School Characteristics on Teacher Decisions," *Research in Labor Economics*, 18:141-176.

Stinebrickner, T.R. (2000): "Serially Correlated State Variables in Dynamic, Discrete Choice Models," *Journal of Applied Econometrics*, 15, 595-624.

Stock, J., and D. Wise (1990): "Pensions, the Option Value of Work and Retirement," *Econometrica*, 58, 1151-1180.

Swann, C. (2002): "Maximum Likelihood Estimation Using Parallel Computing: An Introduction to MPI," *Computational Economics*, 19(2), 145-178.

Tauchen, G., and R. Hussey (1991): "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica*, 59, 371-396.

van der Klaauw, W. and K. Wolpin, (2005) “Social Security, and the Retirement and Savings Behavior of Low Income Households,” unpublished manuscript, Department of Economics, University of Pennsylvania, http://www.econ.upenn.edu/pier.

Waidmann, T., J. Bound, and M. Schoenbaum (1995): “The Illusion of Failure: Trends in Self-reported Health of the U.S. Elderly,” *The Milbank Quarterly*, 73, 253-87.

Figure 1 Example time-line for measurement of health and work status

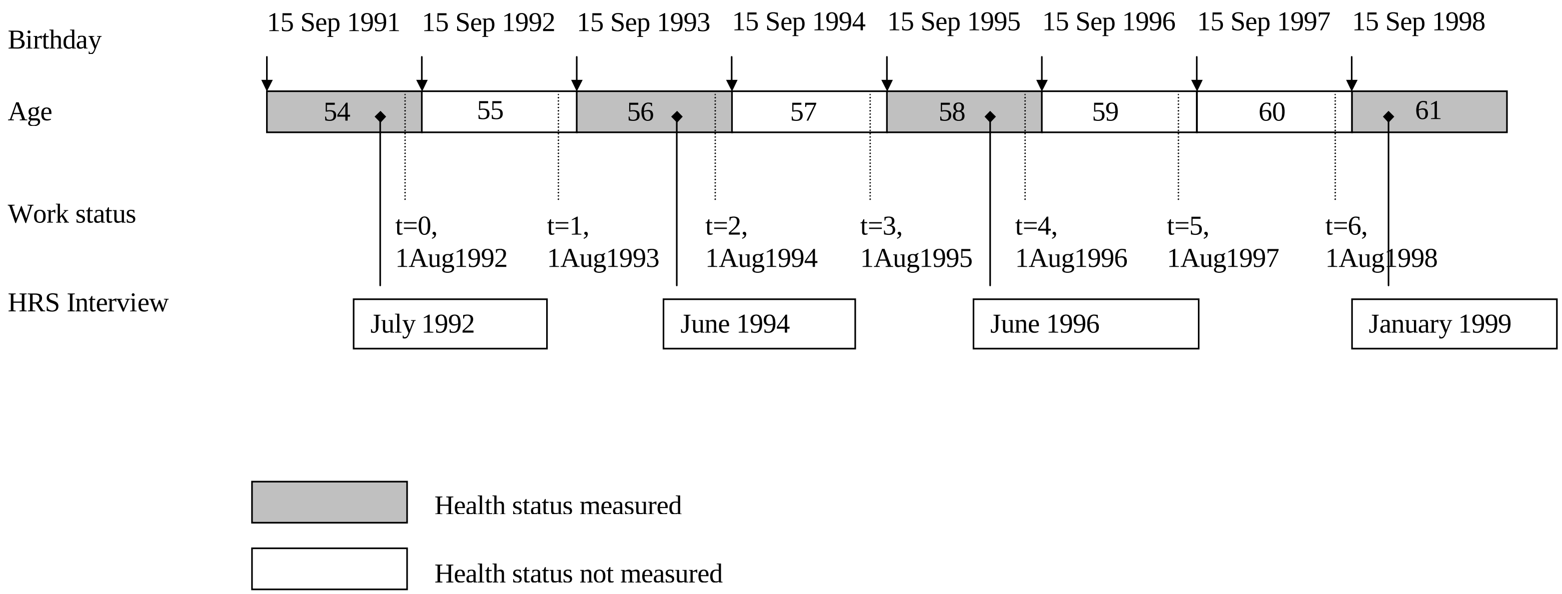

Table 1 Income Sources at Baseline

Work Status	**Not Working**		**Not Working**		**Working**		**Working**	
Limitation Status	**No Limitation**		**Limited**		**No Limitation**		**Limited**	
N	**40**		**122**		**377**		**31**	
Own								
Earnings	15.0%	1,590	3.3%	762	94.2%	33,174	90.3%	20,555
Unemployment Insurance	2.5%	40	0.8%	25	5.3%	90	9.7%	333
Worker's Comp	0.0%	-	2.5%	105	1.9%	53	0.0%	-
Veteran's Benefits	12.5%	1,831	8.2%	592	3.7%	420	9.7%	1,329
Pensions	40.0%	5,894	13.9%	1,868	2.7%	400	6.5%	365
Annuities	0.0%	-	3.3%	544	0.0%	-	3.2%	258
SSI	7.5%	173	35.2%	1,749	0.0%	-	0.0%	-
Social Security	10.0%	558	37.7%	2,462	0.0%	-	6.5%	646
Welfare	5.0%	85	12.3%	242	0.5%	5	0.0%	-
Total		10,172		8,350		34,141		23,485
Household								
Business/Royalties/Trusts	0.0%	-	0.0%	-	9.0%	3,325	9.7%	1,355
Unearned Income	35.0%	488	9.8%	216	35.3%	1,875	41.9%	1,376
Alimony	2.5%	98	0.8%	15	0.3%	3	0.0%	-
Food Stamps	15.0%	185	32.0%	353	1.9%	23	3.2%	108
Total		770		583		5,226		2,839
Total, all sources		10,941		8,933		39,367		26,324

Table 2 Income Transitions: Percent with each income source and mean among those with any

1992-1998 Choice	Bridge				Career				Apply for DI				Retire			
N	63				86				27				77			
	1991		1999		1991		1999		1991		1999		1991		1999	
	%	mean	%	mean	%	mean	%	mean	%	mean	%	mean	%	mean	%	mean
Own																
Earnings	96.8%	$31,469	69.8%	$29,158	98.8%	$29,936	89.5%	$36,200	100.0%	$20,802	18.5%	$2,951	100.0%	$30,293	10.4%	$1,506
Unemployment Ins.	4.8%	106	7.9%	297	8.1%	262	4.7%	127	7.4%	326	0.0%	-	11.7%	182	1.3%	104
Workers' Comp	1.6%	63	1.6%	14	1.2%	2	2.3%	264	0.0%	-	3.7%	185	2.6%	21	0.0%	-
Veteran's Benefits	4.8%	629	4.8%	620	4.7%	497	5.8%	710	0.0%	-	0.0%	-	6.5%	979	6.5%	1,255
Pensions	1.6%	206	60.3%	7,962	4.7%	286	44.2%	2,133	0.0%	-	29.6%	1,785	3.9%	676	79.2%	15,986
Annuities	3.2%	317	3.2%	143	0.0%	-	2.3%	-	3.7%	370	0.0%	-	1.3%	117	5.2%	320
SSI	0.0%	-	1.6%	133	0.0%	-	0.0%	-	3.7%	148	18.5%	655	0.0%	-	2.6%	17
Social Security	1.6%	111	36.5%	3,447	1.2%	82	18.6%	1,507	0.0%	-	70.4%	6,828	1.3%	56	64.9%	5,874
Welfare	1.6%	38	0.0%	-	0.0%	-	0.0%	-	3.7%	69	0.0%	-	0.0%	-	1.3%	8
TotalOwn		**$32,940**		**$41,774**		**$31,064**		**$40,940**		**$21,715**		**$12,404**		**$32,323**		**$25,071**
Spouse																
Earnings	0.0%	-	11.1%	2,333	0.0%	-	14.0%	4,360	0.0%	-	11.1%	3,963	0.0%	-	20.8%	4,983
Unemployment Ins.	0.0%	-	0.0%	-	0.0%	-	0.0%	-	0.0%	-	0.0%	-	0.0%	-	0.0%	-
Workers' Comp	0.0%	-	0.0%	-	0.0%	-	1.2%	47	0.0%	-	0.0%	-	0.0%	-	0.0%	-
Veteran's Benefits	0.0%	-	0.0%	-	0.0%	-	0.0%	-	0.0%	-	3.7%	44	0.0%	-	0.0%	-
Pensions	0.0%	-	4.8%	91	0.0%	-	2.3%	195	0.0%	-	3.7%	281	0.0%	-	7.8%	236
Annuities	0.0%	-	1.6%	105	0.0%	-	0.0%	-	0.0%	-	0.0%	-	0.0%	-	1.3%	-
SSI	0.0%	-	0.0%	-	0.0%	-	2.3%	99	0.0%	-	0.0%	-	0.0%	-	1.3%	-
Social Security	0.0%	-	3.2%	136	0.0%	-	0.0%	-	0.0%	-	7.4%	659	0.0%	-	2.6%	192
Welfare	0.0%	-	0.0%	-	0.0%	-	0.0%	-	0.0%	-	0.0%	-	0.0%	-	0.0%	-
Total Spouse		-		**$2,665**		-		**$4,701**		-		**$4,948**		-		**$5,412**
Household																
Business	4.8%	249	22.2%	13,743	8.1%	1,430	14.0%	3,079	3.7%	444	11.1%	2,169	0.0%	-	3.9%	91
Unearned Income	39.7%	778	61.9%	7,679	39.5%	1,583	62.8%	5,575	11.1%	5,850	37.0%	3,365	35.1%	1,327	62.3%	4,485
Alimony	1.6%	16	0.0%	-	0.0%	-	0.0%	-	0.0%	-	0.0%	-	0.0%	-	0.0%	-
Food Stamps	3.2%	4	0.0%	-	0.0%	-	2.3%	26	11.1%	141	29.6%	101	3.9%	35	1.3%	-
Total Household		**$1,047**		**$21,422**		**$3,013**		**$8,680**		**$6,436**		**$5,626**		**$1,362**		**$4,576**
Total		**$33,988**		**$65,862**		**$34,077**		**$54,322**		**$28,151**		**$22,978**		**$33,686**		**$35,059**
Distribution																
25th Percentile		**$18,000**		**$17,432**		**$15,500**		**$22,000**		**$9,000**		**$6,144**		**$18,000**		**$15,320**
Median		**$30,000**		**$27,764**		**$27,400**		**$45,140**		**$14,600**		**$13,560**		**$34,000**		**$25,728**
75th Percentile		**$45,000**		**$58,460**		**$48,000**		**$70,200**		**$29,000**		**$26,400**		**$42,169**		**$48,872**

Note: Columns represent choice in 1998 based on response to the wave 4 survey. All respondents were employed at baseline.

Table 3 Descriptive Statistics

	Working at t=0	Not working at t=0	Of those working at t=0, behavior last observation			
			Career	Bridge	Retired	Disability
Age at wave 4	60.6	61.1	59.8	61.1	62.0	59.6
< High School	.222	0.261	0.243	0.229	0.188	0.267
Some College	.170	0.183	0.135	0.114	0.208	0.333
College Grad	.193	0.191	0.243	0.114	0.167	0.200
Work limited at wave 1	.097	0.296	0.095	0.143	0.083	0.067
Work limited at wave 4	.188	0.365	0.108	0.057	0.208	0.867
Age 62 or older at wave 4	.369	0.435	0.243	0.457	0.583	0.133
Normal retirement benefit (assuming no further work)	$12,820	$11,271	$13,136	$12,327	$12,775	$11,711
Eligible for DB pension ever	.466	0.078	0.419	0.486	0.542	0.333
Eligible for DC pension ever	.369	0.122	0.405	0.371	0.375	0.267
Expected Career Earnings at last observation	$32,854	$6,048	$33,143	$32,904	$34,700	$23,761
Expected Bridge Earnings at last observation	$13,098	$2,294	$14,647	$12,253	$10,658	$13,793
With no employer health insurance	.176		0.095	0.314	0.125	0.467
With only current coverage	.290		0.297	0.286	0.292	0.200
With both current & retiree coverage	.534		0.608	0.400	0.583	0.333
Median Non-housing wealth	$18,900	$15,000	$20,250	$15,000	$24,000	$8,000
Median Housing wealth	$4,750	-	$250	-	$33,500	$10,000
Fraction Choosing	.605	0.395	0.430	0.204	0.279	0.087

Table 4 - Estimates of Health Equation and Initial Conditions Equation for Baseline Specification

	Health Equation (10)	Initial Conditions Equation (12)
X from equations (10) and (12)		
Constant	-.785 (.105)	.134 (.104)
Less than high school education	-.031 (.190)	.108 (.198)
College education	-.203 (.187)	-.155 (.194)
Age	.008 (.019)	-.059 (.022)
Z's — Respondent reports difficulty with the specified activity.*		
Jog one mile	.202 (.077)	.091 (.091)
Walk several blocks	.325 (.106)	-.0622 (.178)
Walk one block	-.244 (.140)	-.080 (.280)
Sit for about 2 hours	.081 (.100)	.003 (.102)
Get up from a chair after sitting long periods	-.024 (.140)	.045 (.137)
Get in and out of bed without help	.107 (.149)	-.131 (.191)
Go up several flights of stairs	.225 (.073)	-.249 (.134)
Go up one flight of stairs	.230 (.123)	-.207 (.218)
Lift or carry weights over 10 lbs	.307 (.105)	.154 (.164)
Stoop, kneel, or crouch	.157 (.086)	.128 (.131)
Pickup a dime from a table	.152 (.116)	.022 (.199)
Reach or extend your arms above shoulder level	.273 (.102)	.270 (.200)
Pull or push large objects like a living room chair	.200 (.105)	-.347 (.181)
ρ	.92 (.155)	
$var(\nu_t)$	.82 (.080)	

*Activities of Daily Living at time health is observed for first column and t=0 for second column. Coded as 1 if the man reports difficulty doing the activity, 0 otherwise.

Table 5 -Estimates of Model (behavioral equations and covariance terms)

	Baseline	Traditional model with Binary Health
	Estimate (Std. Error)	**Estimate (Std. Error)**
Pecuniary Utility		
τ	.336 (.060)	.325 (.060)
$U_{np}^{C}(t)$ **Non-Pecuniary Utility Career** (C)		
Constant	.546 (.338)	.772 (.193)
λ_{η}^{C} (Health, η)	-.776 (.397)	-1.590 (.242)
λ_{HI} (Has Health Insurance)	.063 (.043)	.131 (.102)
$U_{np}^{B}(t)$ **Non-Pecuniary Utility Bridge(B)**		
Constant	-.908(.486)	.242 (.197)
λ_{η}^{B} (Health, η)	-1.357(.351)	-2.978 (.394)
λ_{HI} (Has Health Insurance)	.063 (.043)	.131 (.102)
$U_{np}^{A}(t)$ **Non-Pecuniary Utility DI (A)**		
Constant	-2.102(.399)	-3.496 (1.690)
λ_{η}^{A} (Health, η)	.851 (.212)	2.900 (1.657)
λ_{HI} (Has Health Insurance)	.063 (.043)	.131 (.102)
$U_{np}^{N}(t)$ **Non-Pecuniary Utility Non-Work (N)**		
Constant	Normalized to zero	Normalized to zero
λ_{η}^{N} (Health, η)	Normalized to zero	Normalized to zero
λ_{HI} (Has Health Insurance)	.063 (.043)	.129 (.104)
Health Equation	See Table 4	N.A.
Initial Conditions Equation	See Table 4	Not Shown
Covariance Terms		
σ_{κ}	1.423 (.322)	.824 (.052)
COV(ϵ^{I}, ν_t)	-.309 (.106)	-.319 (.125)
COV(μ_t, ϵ_t^{C})	-.191 (.173)	N.A.
COV(μ_t, ϵ_t^{B})	.216 (.128)	N.A.
COV(ϵ^{I} , κ)	-.784 (.412)	.008 (.255)
Log Likelihood Function Value	-940.033	

Although not shown, each non-pecuniary equation also includes two dummy variables characterizing a person's education level (less than high school and more than high school). The effect of health insurance is constrained to be the same across choices.

Table 6 Choice Probabilities of Illustrative Person at Different Ages - Baseline Specification

Choice at t=1	A. Average Health	B. Health 1 Std. Dev. Below Average	C. Health 1.5 Std. Dev. Below Average
		AGE=55	
Labor Force Exit (N+A)	0.019	0.098	0.195
Non-Work (N)	0.018	0.057	0.084
Apply DI (A)	0.001	0.041	0.111
		AGE=60	
Labor Force Exit (N+A)	0.025	0.106	0.198
Non-Work (N)	0.024	0.082	0.127
Apply DI (A)	0.001	0.024	0.071
		AGE=62	
Labor Force Exit (N+A)	0.049	0.168	0.269
Non-Work (N)	0.046	0.138	0.209
Apply DI (A)	0.003	0.030	0.060
		AGE=64	
Labor Force Exit (N+A)	0.053	0.172	0.279
Non-Work (N)	0.050	0.146	0.207
Apply DI (A)	0.003	0.026	0.072
		AGE=65	
Labor Force Exit (N+A)	0.053	0.161	0.254
Non-Work (N)	0.053	0.161	0.254
Apply DI (A)	NA*	NA*	NA*

* Person cannot apply for Disability Insurance at age 65 or older.

Table 7 Choice Probabilities of Illustrative Person at Different Ages Using Alternate Models of Health

	Traditional Model w/ Binary S-R Health		Continuous Health Model		Continuous Health w/ σ_κ Fixed at Binary Value	
	Good Health	Bad Health	Good Health	Bad Health	Good Health	Bad Health
			AGE=55			
Labor Force Exit (N+A)	0.079	0.452	0.079	0.244	0.071	0.336
Non-Work (N)	0.068	0.111	0.067	0.096	0.061	0.149
Apply DI (A)	0.011	0.341	0.011	0.148	0.010	0.186
			AGE=60			
Labor Force Exit (N+A)	0.094	0.426	0.083	0.245	0.082	0.340
Non-Work (N)	0.088	0.183	0.074	0.14	0.074	0.207
Apply DI (A)	0.006	0.243	0.008	0.105	0.008	0.133
			AGE=62			
Labor Force Exit (N+A)	0.159	0.513	0.112	0.293	0.120	0.403
Non-Work (N)	0.147	0.286	0.102	0.191	0.110	0.259
Apply DI (A)	0.012	0.227	0.010	0.101	0.010	0.144
			AGE=64			
Labor Force Exit (N+A)	0.115	0.515	0.114	0.303	0.125	0.411
Non-Work (N)	0.107	0.158	0.105	0.19	0.114	0.251
Apply DI (A)	0.008	0.356	0.009	0.113	0.010	0.160
			AGE=65			
Labor Force Exit (N+A)	0.091	0.291	0.111	0.274	0.118	0.359
Non-Work (N)	0.091	0.291	0.111	0.274	0.118	0.359
Apply DI (A)	NA*	NA*	NA*	NA*	NA*	NA*

* person cannot apply for Disability Insurance at age 65 or older.

Table 8A Policy Simulations - All Health Combined

	Baseline	Policy 1 Remove SS ERB	Policy 2 SS NRA to 67	Policy 3 Eliminate DI	Policy 4 25% Decrease DI Benefits	Policy 5 25% Decrease SS Benefits	Policy 6 25% Decrease DI & SS Benefits
Pooled Ages							
LF Exit (N+A)	.152	.148	.151	.144	.150	.150	.148
Non-Work (N)	.134	.128	.130	.144	.135	.130	.132
Apply DI (A)	.019	.020	.021	0**	.015	.020	.017
AGE=55							
LF Exit (N+A)	.147	.147	.147	.136	.145	.147	.144
Non-Work (N)	.123	.121	.122	.136	.125	.122	.123
Apply DI (A)	.024	.026	.025	0**	.020	.025	.021
AGE=60							
LF Exit (N+A)	.154	.151	.153	.149	.153	.153	.151
Non-Work (N)	.142	.137	.140	.149	.143	.140	.140
Apply DI (A)	.012	.014	.013	0**	.010	.013	.011
AGE=62							
LF Exit (N+A)	.170	.147	.163	.163	.168	.162	.160
Non-Work (N)	.155	.132	.146	.163	.156	.145	.146
Apply DI (A)	.015	.016	.016	0**	.012	.017	.014
AGE=64							
LF Exit (N+A)	.205	.183	.197	.197	.204	.197	.195
Non-Work (N)	.176	.158	.166	.197	.180	.164	.167
Apply DI (A)	.029	.025	.031	0**	.024	.033	.028
AGE=65							
LF Exit (N+A)	.220	.214	.218	.216	.220	.209	.208
Non-Work (N)	.220	.214	.181	.216	.220	.209	.208
Apply DI (A)	0*	0*	.037	0*	0*	0*	0*

* Person cannot apply for Disability Insurance at age 65 or older except in Policy 2.
** Policy involves removing Disability Insurance program.
Policy 1 involves removing all SS benefits before age 65.
Policy 2 involves changing the normal retirement age from 65 to 67.
Policy 3 involves removing the DI program.
Policy 4 involves a 25% decrease in DI benefits without changing SS Retirement benefits.
Policy 5 involves a 25% decrease in SS Retirement benefits without changing DI benefits.
Policy 6 involves ad 25% reduction in both SS Retirement and DI benefits.

Table 8B Policy Simulations - Bad Health (1 standard deviation or more below average)

	Baseline	**Policy 1 Remove SS ERB**	**Policy 2 SS NRA to 67**	**Policy 3 Eliminate DI**	**Policy 4 25% Decrease DI Benefits**	**Policy 5 25% Decrease SS Benefits**	**Policy 6 25% Decrease DI & SS Benefits**
Pooled Ages							
LF Exit (N+A)	.400	.396	.399	.363	.390	.399	.389
Non-Work (N)	.296	.283	.283	.363	.307	.286	.297
Apply DI (A)	.104	.112	.117	0**	.084	.113	.092
AGE=55							
LF Exit (N+A)	.405	.406	.405	.359	.393	.406	.394
Non-Work (N)	.277	.272	.275	.359	.289	.273	.285
Apply DI (A)	.127	.133	.130	0**	.104	.133	.109
AGE=60							
LF Exit (N+A)	.384	.381	.383	.366	.380	.383	.377
Non-Work (N)	.324	.308	.317	.366	.332	.315	.322
Apply DI (A)	.060	.074	.066	0**	.048	.068	.055
AGE=62							
LF Exit (N+A)	.451	.423	.444	.425	.444	.444	.436
Non-Work (N)	.351	.310	.333	.425	.367	.326	.340
Apply DI (A)	.100	.114	.111	0**	.078	.118	.096
AGE=64							
LF Exit (N+A)	.481	.455	.472	.446	.473	.473	.463
Non-Work (N)	.301	.296	.274	.446	.329	.261	.290
Apply DI (A)	.180	.159	.198	0**	.144	.212	.173
AGE=65							
LF Exit (N+A)	.474	.465	.489	.457	.474	.461	.456
Non-Work (N)	.474	.465	.307	.457	.474	.461	.456
Apply DI (A)	0*	0*	.181	0*	0*	0*	0*

* person cannot apply for Disability Insurance at age 65 or older except in Policy 2.
** policy involves removing Disability Insurance program.
Policy 1 involves removing all SS benefits before age 65.
Policy 2 involves changing the normal retirement age from 65 to 67.
Policy 3 involves removing the DI program.
Policy 4 involves a 25% decrease in DI benefits without changing SS Retirement benefits.
Policy 5 involves a 25% decrease in SS Retirement benefits without changing DI benefits.
Policy 6 involves ad 25% reduction in both SS Retirement and DI benefits.

Table 8C Policy Simulations - Good Health (no worse than 1 standard deviation below average)

	Baseline	Policy 1 Remove SS ERB	Policy 2 SS NRA to 67	Policy 3 Eliminate DI	Policy 4 25% Decrease DI Benefits	Policy 5 25% Decrease SS Benefits	Policy 6 25% Decrease DI & SS Benefits
Pooled Ages							
LF Exit (N+A)	.118	.113	.116	.113	.117	.116	.114
Non-Work (N)	.111	.106	.109	.113	.111	.109	.108
Apply DI (A)	.007	.007	.007	0**	.006	.007	.006
AGE=55							
LF Exit (N+A)	.103	.103	.103	.098	.102	.103	.102
Non-Work (N)	.097	.095	.096	.098	.097	.096	.096
Apply DI (A)	.007	.007	.007	0**	.006	.007	.006
AGE=60							
LF Exit (N+A)	.125	.122	.123	.122	.124	.123	.122
Non-Work (N)	.119	.115	.117	.122	.119	.117	.117
Apply DI (A)	.006	.007	.006	0**	.005	.006	.006
AGE=62							
LF Exit (N+A)	.145	.123	.138	.140	.144	.137	.136
Non-Work (N)	.137	.116	.130	.140	.137	.129	.129
Apply DI (A)	.008	.007	.008	0**	.006	.008	.007
AGE=64							
LF Exit (N+A)	.171	.149	.163	.166	.170	.163	.161
Non-Work (N)	.161	.141	.152	.166	.161	.151	.151
Apply DI (A)	.010	.008	.010	0**	.009	.011	.010
AGE=65							
LF Exit (N+A)	.178	.172	.208	.175	.178	.167	.166
Non-Work (N)	.178	.172	.195	.175	.178	.167	.166
Apply DI (A)	0*	0*	.013	0*	0*	.013	0*

* person cannot apply for Disability Insurance at age 65 or older except in Policy 2.
** policy involves removing Disability Insurance program.
Policy 1 involves removing all SS benefits before age 65.
Policy 2 involves changing the normal retirement age from 65 to 67.
Policy 3 involves removing the DI program.
Policy 4 involves a 25% decrease in DI benefits without changing SS Retirement benefits.
Policy 5 involves a 25% decrease in SS Retirement benefits without changing DI benefits.
Policy 6 involves ad 25% reduction in both SS Retirement and DI benefits.

Appendix A. Data

A.1 Overview of Health and Retirement Study and sample composition

The HRS is described in additional detail in Juster and Suzman (1995). The survey 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 over-samples 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 in 1994; the second wave re-interviewed 11,317 of the original respondents, representing 91% of the original sample. The third wave (1996) reinterviewed 10,681 of the original sample, and the fourth (1998) reinterviewed 10,242, representing 86 percent of the original respondents. By 1998, 805 (6.4%) of the original respondents were deceased.

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, of whom 4,522 are men, of whom 733 are not married/partnered. Table A.1 describes the effects of sample exclusions necessary for our analysis. From this group, we exclude 206 respondents who did not have a Social Security earnings history, either because the respondent refused the HRS permission to access their records or because they had no covered earnings between 1951 and 1991. We exclude 31 respondents who claim to have a private pension on their current job, but who have no corresponding record provided by their employers. We then exclude 153 respondents who were not eligible to receive both retirement (old age) and disability payments from the Social Security system if they did not work past 1992 because they lacked the required number of quarters of covered employment. Finally we excluded 15 respondents who had missing data for any of the variables used in our models. These exclusions left 328 respondents who were included in the initial conditions sample. Of these, 132 were not employed (or were self employed) at the date of their wave 1 (1992/93) interview. The remaining 196 respondents make up the "behavioral sample."

Table A.1 Sample Definition

HRS Age-Eligible Single Men	733
...with a social security earnings history available	527
...with valid pension plan data	496
...who were covered for Social Security retirement and disability	343
...with no missing data (*Initial Conditions Sample*)	328
...employed as of wave 1(Behavioral Sample)	196

A.2 Timing Issues Related to Income Sources

Given our assumption that an individual consumes all of the income that he receives in the same year, characterizing consumption in a particular year requires that we describe our assumptions related to the timing of the income from a variety of sources. For some sources of income, institutional details imply that the timing of receipt is obvious given the labor supply decisions that a person makes. For example, that defined benefit (DB) pension plans do not typically include an actuarial adjustment for delayed receipt implies that individuals will begin receiving DB pension payments from a particular job as soon as they become eligible for benefits and are no longer working in a job. We assume that DB pension wealth is not accrued in bridge jobs. While this assumption is made largely for computational reasons, the data suggest that it is a reasonable simplifying assumption. Individuals also have little discretion about the timing of income from food stamps, Supplemental Security Income, and other exogenous sources of income. In addition, the receipt of Disability insurance payments starts automatically after one applies and is approved for benefits. We treat defined contribution pension plans analogously to defined benefit pension plans by assuming that benefits from these plans are only accrued on career jobs and that an individual starts receiving benefits as soon as he becomes eligible and is no longer working at his career job.

For Social Security, some individual discretion about the timing of benefits remains given a person's labor supply choices, and we try to make reasonable assumptions about the timing of benefits. Specifically, we assume that an individual begins receiving Social Security benefits in the first year he is both eligible (i.e, he is 62 years old or older) and not-working. As noted, the choice-set allows individuals who are younger than 70 years of age the flexibility of returning to work after leaving the workforce. If an individual returns to work after beginning to collect benefits, earnings are taxed away and actuarial adjustments are made to future earnings in accordance with the rules of the Social Security system.

Perhaps the most difficult timing issue relates to the manner in which a person spends the wealth that he has accumulated by the beginning of the sample period. Our approach for dealing with wealth takes into account the intuitively appealing notion that individuals who are not working are more likely to use portions of their wealth for consumption. Specifically, we compute the yearly value of a hypothetical annuity that a person could buy at time t=0 given his wealth. In any year for the remainder of his life that the person does not work, he is assumed to consume the annuitized value of his wealth. In any year that the person does work, we assume that the person saves the annuitized amount.[44] This additional savings is used to increase consumption when he retires permanently (i.e., takes a year off at age 70 or older).[45] Finally, we assume that health expenditures are paid out of income in the year that they are incurred. However, consistent with the notion that there exists government assistance, we assume that a person's consumption in a period cannot fall below a minimum level of subsistence.[46]

A.3. More detailed information related to various sources of income

A.3.1 The health expenditure equation

Following French and Jones (2004), we assume that a person's health expenditures at time t depends on the person's health insurance status k, age, and health:

(A.1) $\text{Ln Expenditures}(t) = \Lambda_1^k + \Lambda_2 (\text{Age-53}) + \Lambda_3 \eta_t + e^k$

where we differentiate between four possible health insurance categories - $k \in$ {ESI (employer sponsored health insurance) only, ESI and Medicare, Medicare only, no coverage}. Λ_1 is a constant which depends on a person's health insurance status k. Λ_2 and Λ_3 measure the impact of age and health respectively. e is a a random component which is assumed to be normally distributed with a mean of zero and a variance which depends on a person's health insurance status.

To estimate the model of out-of-pocket health care expenditures, we used the 1996 and 1998 waves of the HRS. The sample we use (n= 4,619) consists of age-eligible men, who had not applied for Disability Insurance, and who did not have Medicaid between 1995 and 1998. Respondents fall into one of four insurance groups: those with employer sponsored health insurance (ESI) and Medicare coverage (n=261), those with ESI but no Medicare (n=2,419), those with Medicare but no ESI (n=527), and those with neither ESI nor Medicare (n=1,412).

The expenditure variable we use includes out-of-pocket payments for medical care provided during the two years between interviews plus any insurance premiums paid by the person. As a result, during estimation we divide the amounts from the expenditure equation by two. Payments included are all those for nursing home and hospital stays, doctor visits, outpatient surgeries, dental visits, prescriptions drugs, and in-home medical care. The expenditure model is estimated using a Tobit method with expenditure censored on the left at the

---

[44] The formulation precludes any active savings. As mentioned in footnote 2, the population we are specifically interested in does not appear to have much in the way of savings at baseline. For this reason, we do not believe our rather ad-hoc way of dealing with savings and wealth seriously distorts our estimates of the impact of health on labor market behavior.

[45] At the time the person retires permanently we compute the yearly value of a second hypothetical annuity that could be purchased with the savings he has accumulated by not consuming the first annuity (i.e., by working in some years between t=0 and the time he retires permanently). From the date of permanent retirement until the end of his life, the person receives both the yearly value of the first annuity and the yearly value of the second annuity.

[46] We assume that this consumption floor corresponds to the maximum payment that a person could receive from the food stamps program. Thus, no individuals in our model receive consumption lower than the maximum food stamp payment.

median level of expenditures and on the right at the 99th percentile. The censoring is done, first, for the entire sample and, then, separately by insurance group. Residual variances are allowed to vary across insurance groups but are fixed within them. The estimates of the expenditure model are shown in Table A.2.

Table A.2 Estimates of parameters in health expenditure equation

	Common Truncation		Separate Truncation	
Variable	**Coefficient**	**Std Error**	**Coefficient**	**Std Error**
Age-55	0.015	0.008	0.014	0.008
Health	0.335	0.032	0.339	0.032
ESI, Medicare	0.407	0.119	0.382	0.129
ESI, no Medicare	0.214	0.070	0.218	0.068
no ESI, Medicare	0.344	0.099	0.413	0.099
Constant	6.926	0.102	6.926	0.101
Insurance Group	**Variance**		**Variance**	
ESI, Medicare	1.893		2.042	
ESI, no Medicare	1.659		1.682	
no ESI, Medicare	1.840		1.632	
no ESI, no Medicare	2.392		2.421	

To see the effect of insurance group on the distribution of expenditures, we produced hypothetical distributions of expenditure using the Λ_1^k and variance estimates while holding age and health constant. Our predictions imply that a ESI reduces out of pocket expenditures in the upper tail of the distribution, but not at the median. This finding is similar to that of French and Jones (2004), and is consistent with findings from the Medical Expenditure Panel Survey (MEPS) (Olin & Machlin, 2004) that persons with more comprehensive insurance coverage receive substantially more care, and as a result may have significant coinsurance expenses.

	25th Percentile	**Median**	**75th Percentile**
Common Truncation:			
ESI, Medicare	$ 284.86	$ 715.48	$ 1,789.05
ESI, no Medicare	$ 483.76	$ 602.89	$ 749.57
no ESI, Medicare	$ 271.26	$ 672.50	$ 1,659.87
no ESI, no Medicare	$ 168.88	$ 475.41	$ 1,331.61
Separate Truncation:			
ESI, Medicare	$ 269.60	$ 685.78	$ 1,809.54
ESI, no Medicare	$ 480.64	$ 600.81	$ 746.96
no ESI, Medicare	$ 307.93	$ 709.37	$ 1,688.58
no ESI, no Medicare	$ 168.94	$ 466.89	$ 1,342.89

A.3.2. The earnings equations

We assume that the log of a person's real earnings in his career job at time t evolves according to the fixed-effect specification

$$\text{(A.2)} \quad \mathrm{Ln}(W^C_t) = \alpha_1^i + \alpha_2 t + \psi_t^C \quad \Rightarrow \quad W^C_t = e^{\alpha_1^i + \alpha_2 t} e^{\psi_t^C}$$

where α_1^i is a person-specific fixed-effect that can be estimated given multiple career-job earnings observations in the Health and Retirement Study and ψ^C_t is the stochastic component of career job earnings. The log of a person's real earnings in his bridge job option is given by

$$\text{(A.3) } Ln(W^B_t) = \alpha_3 EXITAGE + \alpha_4 ln(W_o^C) + \psi_t^B \qquad \Rightarrow W^B_t = e^{\alpha_3 EXITAGE + \alpha_4 ln(W_o^C)} e^{\psi_t^B}$$

where ψ^B_t is the stochastic component of bridge job earnings, and EXITAGE is the age at which a person left his career job. Our specification allows for the possibility that the earnings a person received in his career job at t=0 may contain information more generally about his earnings potential.

We do not make functional form assumptions about the distributions of ψ_t^C and ψ_t^B since our simulation approach described in Section 3 allows us to use the empirical distributions of these random variables.

Earnings in the career job are estimated using Internal Revenue Service (Form W-2) data linked to HRS records. The data cover reported wage and salary income and self-employment income from the years 1980-1991. We estimated the rate of growth in real earnings for the persons in our sample over the period to be 1.93% (s.e.=0.23%) per year. Future real earnings were thus assumed to grow at this rate from the base reported in 1992. Earnings in bridge jobs were estimated using self-reported data from survey year 1994. Sample individuals who reported working in a job different from the one reported in the 1992 survey were included in the estimation sample.

Estimates of parameters in earnings equations

	Coefficient	***Standard Error***
Career Job		
Age	0.019	*0.002*
Constant	8.954	*0.108*
σ_u (dispersion of fixed effects)	0.994	
σ_e (dispersion of residuals)	0.396	
ρ (fraction of variance due to u)	0.863	
n	4923	
F test that all u_i=0:	F(2142, 2779)=	*9.81*
Bridge Job		
Age left career job	-0.061	*0.021*
1992 log wage	0.381	*0.082*
Constant	9.429	*1.358*
n	230	
R^2	0.107	

A.3.3. Issues related to Social Security Benefits

Recall that the choice set in Section 2.1 allows individuals to leave the workforce and return at a later time. With respect to bridge job information, in reality the actual SS benefit formula implies that a person's SS benefits would depend on not only the number of years that a person works in a bridge job before t and the wages in those years but also to some extent on the specific ages that the bridge job wages were received. This is the case because earnings that enter the Average Indexed Monthly Earnings (AIME) which determines SS benefit amounts are indexed by year-specific factors (that are different than standard rates of inflation) at ages less than 60 and are not indexed at ages greater than 60. We choose not to use this age information when computing SS benefits because including it in the model is not computationally feasible since it would require that the model contain a set of state variables that characterize the specific ages at which the person worked in bridge jobs (or alternatively state variables that keep track of the years the person was out of the workforce before time t). This assumption does not seem particularly problematic because the age information tends to have a relatively small effect on benefits. The SS benefits that we use in our model are those obtained with the person's bridge years taking place immediately after a person's career years.

The SS rules imply that, when recomputing the actuarial adjustment factor, one month is added to the Social Security starting age for each month before the normal retirement age that a person incurs a benefit reduction (including partial reductions) due to work. Thus, the actuarial adjustment factor typically changes when a person returns to work before the normal retirement age (after beginning to collect SS benefits) because a partial benefit reduction occurs in a month for anyone who earns more than the monthly equivalent of approximately $10,000 a year. However, working after the normal retirement age (after beginning to collect SS benefits) only leads to a change in the actuarial adjustment factor if the person incurs a full benefit reduction due to work in a particular month. Thus, the actuarial adjustment factor does not typically change for those working after the normal retirement age (after beginning to collect SS benefits) because, for example, a person with $20,000 in yearly SS benefits would only incur a full benefit reduction in a month if he earned more than the monthly equivalent of approximately $70,000 a year in his bridge job. This implies that the state variable SSEX in Section 2.3.2 keeps track of the number of years that the person is out of the workforce between the age of 62 and his normal retirement age, and, if the person is not out of the workforce in any of these years, SSEX keeps track of the first subsequent age that the person is out of the workforce.

Social security benefits are calculated using SSA data on covered earnings that have been attached to HRS records and self reported data from 1992 using program rules in effect as of 1992. Earnings histories are available for persons in covered employment beginning in 1950. Future earnings are assumed to evolve according to the estimates described above. Individuals who do not give HRS permission to access earnings histories are excluded from the analysis.

A.3.4 Issues related to Defined Benefit Pensions

In order to be eligible for a defined benefit payment from a particular job at time $t \geq 1$, a person must have left this job at this point. In this case, whether a payment is received and the amount of the payment depends on a person's earnings history in that job as of time t and the details of the employer's pension plan. Much heterogeneity exists in both eligibility ages and payment structures across plans.

For the discussion here it is worthwhile to group the set of jobs that a person holds during his working lifetime as jobs that the person left before t=0 (referred to hereafter as previously held jobs), the job the worker held at t=0 (defined above as a career job), and jobs that the worker begins after t=0 (defined above as bridge jobs). With respect to bridge jobs, for reasons of model tractability we assume that defined benefit pensions are not accumulated in these jobs.[47] Descriptive work shows that this is a reasonable simplifying assumption. With respect to any previously held job, because the worker has left this job as of t=0, θ contains all details of the pension plan and information about the worker's earnings history at the job that is necessary for the person at t=1 to compute the defined benefit payment he will receive at any $t \geq 1$. With respect to the career job, because the worker has not left this job at t=0, in order he must know not only the details of the employer's pension plan and information about his earnings history at the job that are contained in θ but also the endogenously determined number of years that he will remain in his career job and the wages that he will earn in each of these years. Letting EXC(t) denote the years of experience that a person has accumulated in his career job as of time t, the number of years that the person stays in his career job before leaving is EXC(T) and the set of wages relevant to the DB calculation is given by $W^C_1, W^C_2, \ldots, W^C_{EXC(T)}$.[48] The assumption that an individual considers expected future earnings when thinking about future DB benefits implies that the information in θ and EXC(T) is sufficient for the person at t=1 to compute the DB that he will receive in some year in the future.

Values for defined benefit pension income are computed by the HRS Pension Benefit Calculator. In 1992 the HRS asked respondents who were working for the name and address of their current employer as well as their previous employer. Using this information, the HRS contacted the employers to obtain summary plan descriptions. These descriptions, supplemented with data on pension plans maintained by the U. S. Department of Labor were used to produce formulas for the calculation of individual pension benefits based on the individual's wage history for any given date of retirement. See Curtin et al. (1998) for details on the pension

[47] Allowing DB pensions to be accumulated in jobs that are started after time t=0 is difficult because it requires a person to think about the types of pension plans that could arrive with all possible future job offers. In this case, variables that are capable of describing all pension plans that arise in the future would have to be included in the model as state variables.

[48] EXC(T), the number of years of career job experience as of the end of his decision horizon, is the relevant number because the person cannot return to his career job after leaving.

calculator program. For each sample member with a pension plan recorded in the plan database we generated annual pension benefit values for every future age of retirement from the current job beginning in the current year through the year in which the respondent reaches age 70.

A.3.5 Issues related to Defined Contribution Pensions

Defined contribution pensions are treated analogously to defined benefit pensions. To be consistent with our DB assumption regarding bridge jobs, we assume that DC benefits are not accumulated in jobs that begin after t=0. The defined contribution benefits at future time t associated with jobs that ended before t=0 are entirely characterized by information about the total amount of contributions that have been made to the plan and the plan's age of eligibility that is contained in our baseline financial characteristics β.[49] With respect to the career job, because the worker has not left this job at t=0, in order for the person at t=0 to compute the DC that he will receive in some year in the future he must know not only information about previous contributions and the plan's age of eligibility that is contained in β but also the endogenously determined number years he will remain in his career job and how much he will contribute to his DC plan in each of these years. We abstract from the endogeneity of the contribution decision by assuming that each individual continues to contribute the same percentage of his income to the DC plan at his career job in the future as he has in the past. In this case, given the past rate of contribution which is information contained in β, the information needed to compute future DC benefits is very similar to that of the DB case. In particular, in addition to knowing the information in β, the person must know how many years EXC(T) that he will remain in his career job and the wages $W^C_1, W^C_2, \ldots, W^C_{EXC(T)}$ that he will earn in each of these years. The assumption that an individual considers expected future earnings when thinking about future DC benefits implies that the information in β and EXC(T) is sufficient for the person at t=1 to compute the DC that he will receive in some year in the future.

Values for defined contribution pension income are derived as an actuarially fair annuity based on the accumulated value in the DC pension account at the time of retirement from the career job, according to the formula

$$W(a) = \sum_{x=a}^{119} \frac{l_i(x)}{l_i(a)} \frac{A(a)}{(1+r)^{x-a}}$$

where $W(a)$ is the accumulated value of the pension fund at age a, $l_i(x)$ is the adjusted (see above) life table probability of survival to age x, $A(a)$ is the annuity payment for an individual who begins receiving it at age a, and r is the (real) interest rate, assumed here to be 3.29%. Pension wealth is self-reported at baseline, and it is assumed to accumulate with annual contributions from either or both employers and employees calculated as percentages of wage and salary income from the career job until retirement. Interest is accrued on these pension assets until they are converted to an annuity.

A.3.6. Non-pension wealth

Our general approach for dealing with wealth is described in Appendix A.2. The HRS respondents are asked to value a wide variety of assets and debts (see Smith, 1995, and Hill, 1993, for details). We calculate the total net value of non-pension sources of wealth available at baseline (1992).

A.4 Survival Probabilities

We calculate health-adjusted survival probabilities at each future age for the purposes of discounting future utility and converting stocks of wealth into annuity income estimates. The relationship between current health and survival is estimated using a discrete time hazard model of mortality and the longitudinal data from the HRS. In particular, for persons who are interviewed (taken as an objective measure of survival) in successive waves, we record the date of each interview, and include in the likelihood function an expression for the probability that the individual survives to that date. For those who die between waves, we include in the likelihood function the cumulative probability that the individual dies sometime in the two-year period following their last interview. The survival function is assumed to be exponential. Thus , the probability of survival from year t-1 to year t is next is

$$S(t) = \exp[-\exp(X_t'\beta)],$$

[49] We assume an anticipated real growth rate of 3.29% on all investments.

where single year of age dummies and a health index (the predicted latent variable from an ordered probit estimate of self-assessed health status) are included in *X*. The probability of surviving from one survey (t-2) to the next (t) is then $S(t\text{-}1)S(\text{t})$ and the probability of dying during the interval is just $1 - S(t\text{-}1)S(\text{t})$. Health as of *t-2* and age as of *t-1* and t are included in the model. The coefficient on the health index was 0.623 (0.071). Our estimate of the coefficient on health in the survival model implies that a one standard deviation increase in the latent health index (worse health) increases the mortality hazard by a factor of 1.86. Using this coefficient estimate and individual values of the latent health index, individual survival probabilities to future ages are increased or decreased relative to the national probabilities obtained from Social Security life tables (SSA, 1995).

A.5 Characterizing a person's decisions

An individual in the behavioral sample is classified as having applied for disability benefits at time t if he reported applying for Social Security Disability or Supplemental Security Income sometime between the target dates t-1 and t, regardless of his work status as of time t. Among those not classified as disability applicants, those who report that they are doing no work for pay as of time t are classified as "non-workers." Note that this classification can include respondents who had applied for disability benefits but did so more than a year before time t-1. At each survey, workers are asked if they work for the same employer, named by the interviewer, they reported in the previous wave. If not (or if they reported being self-employed), they are asked the dates at which the previous job ended and the current job began. This information is used to classify individuals as working in a career job or a bridge job. Finally, respondents who either denied working for the employer given in a previous interview or didn't know if they worked for the same employer were excluded from the behavioral sample. Persons who were working but self-employed at wave 1 were also excluded from the behavioral sample.

Appendix B. Computing Value Functions

For illustrative purposes, consider the calculation of the expected future utility associated with the option of staying in one's career job. For clarity below, it will be useful to separate the stochastic components in S(t+1) from the components in S(t+1) that are deterministic given S(t) and the decision to choose j at time t. Denoting the vector of non-stochastic components $S^{NS}(t+1)$, $S(t+1)=S^{NS}(t+1)\cup\{DI(t+1), \eta_{t+1}, \epsilon(t+1), \psi(t+1), HI^{B}(t+1), L(t+1)\}$,

$$
\begin{aligned}
\text{(B.1)}\quad & E[V(t+1, S(t+1)|S(t),d^{C}(t)=1] \\
&= PR(L(t+1)=1)\cdot PR(HI^{B}(t+1)=1)\cdot\int\int\int [V(t+1,S^{NS}(t+1),L(t+1)=1,\psi(t+1),\epsilon(t+1),\eta_{t+1},HI^{B}(t+1)=1] \\
&\qquad dG(\epsilon(t+1))\, dF(\psi(t+1))\, d\,H(\eta_{t+1}|\eta_{t}) \\
&+PR(L(t+1)=0)\cdot PR(HI^{B}(t+1)=1)\cdot\int\int\int [V(t+1,S^{NS}(t+1),L(t+1)=0,\psi(t+1),\epsilon(t+1),\eta_{t+1}, HI^{B}(t=1)=1] \\
&\qquad dG(\epsilon(t+1))\, dF(\psi(t+1))\, d\,H(\eta_{t+1}|\eta_{t}) \\
&+PR(L(t+1)=0)\cdot PR(HI^{B}(t+1)=2)\cdot\int\int\int [V(t+1,S^{NS}(t+1),L(t+1)=0,\psi(t+1),\epsilon(t+1),\eta_{t+1}, HI^{B}(t=1)=2] \\
&\qquad dG(\epsilon(t+1))\, dF(\psi(t+1))\, d\,H(\eta_{t+1}|\eta_{t}) \\
&+PR(L(t+1)=1)\cdot PR(HI^{B}(t+1)=2)\cdot\int\int\int [V(t+1,S^{NS}(t+1),L(t+1)=1,\psi(t+1),\epsilon(t+1),\eta_{t+1}, HI^{B}(t=1)=2] \\
&\qquad dG(\epsilon(t+1))\, dF(\psi(t+1))\, d\,H(\eta_{t+1}|\eta_{t})
\end{aligned}
$$

where H, G, and F are the distribution functions of η_{t+1} given η_t, ϵ, and ψ respectively.[50]

The normality assumption for ϵ described in Section 2.3.2 is made primarily for practical reasons related to the necessity of allowing certain correlations in our model that are discussed in detail in Section 3. Unfortunately, this distributional assumption, along with the equation (8) assumption about η_{t+1} given η_t and our use of the empirical distribution of ψ, implies that equation (B.1) does not have a closed form solution. Simulation approaches have been found to be useful in such contexts. For illustration, considering the first of the three-dimensional integrals in equation (B.1), our approximation approach has its foundations in the "naive" simulator given by

$$
\text{(B.2)}\quad \frac{1}{D}\sum_{d=1}^{D} [V(t+1,S^{NS}(t+1),L(t+1)=1,\psi^{d}(t+1),\epsilon^{d}(t+1),\eta^{d}_{t+1,}HI^{B}(t+1)=1]
$$

where η^{d}_{t+1}, $\epsilon^{d}(t+1)$, and $\psi^{d}(t+1)$ and represent the dth of D draws from the distributions of η_{t+1} given η_t, $\epsilon(t+1)$, and the empirical distribution of $\psi(t+1)$ respectively.

We deviate from this naive simulator in two ways. The first deviation is that, motivated by findings in Stinebrickner (2000), we choose to approximate the outer integral in each of the three-dimensional integrals in equation (B.1) by Hermite Gaussian Quadrature rather than by simulation[51]. The second deviation is motivated by the reality that the naive simulator in equation (B.2) does not have continuous derivatives with respect to the parameters of our model.[52] This causes problems with the derivative-based optimization

[50] L(t+1) is relevant for only the choice C and DI(t+1) is not relevant for the choice C or B since we assume that individuals do not imagine returning to work if they are approved for DI.. Uncertainty about $HI^{B}(t+1)$ exists in all cases except when a person is working in a bridge job with health insurance at time t. In this case, the individual assumes that he will have health insurance if he remains in a bridge job in time t+1. The future utility associated with applying for DI depends on whether the person is approved or not. If not approved the person has multiple options in the next period. If approved, the individual receives the utility from remaining out of the workforce and collecting DI benefits for the remainder of his life. Thus, the expected value is a weighted average of the utility from these sources with the weights coming from a person's belief about the probability that his DI application will be approved.

[51] Stinebrickner(2000) found that Gaussian Quadrature methods performed well relative to simulation methods in a similar context with a serially correlated variable. We have no strong evidence (or strong beliefs) that the quadrature method is a better choice than simulation in this specific context, especially since we are also simulating other dimensions of the integral.

[52] A slight change in a particular parameter can potentially change the option that is the maximum, and, therefore, lead to a discrete change in the derivative of V(·) with respect to the parameter.

algorithms that are a necessity given the computational demands of our model. We address this issue in a manner similar to Keane and Moffitt (1998) by adding an additional extreme value smoothing random variable to each of the current period utility equations (4). The additional parameter produces a smooth simulator which approaches the value in equation (B.1) as the variance of the extreme value smoothing random variables approaches zero. In practice, we set the variance to be a small number.

Specifically, these two deviations from the naive simulator lead to an analog of equation (B.2) that is given by

$$\text{(B.3)}\ \tfrac{1}{\sqrt{\pi}}\sum_{q=1}^{Q} w_q \tfrac{1}{D}\sum_{d=1}^{D} E[V(t+1), S^{NS}(t+1), L(t+1)=1, \psi^d(t+1), \varepsilon^d(t+1), \eta_{t+1} = \rho(\eta_t - \pi X_t) + \pi X_{t+1} + \sqrt{2}\sigma_{\xi} m_q, HI^B(t+1)=1]$$

where, as described in detail in Stinebrickner (2000), m_q and w_q, q=1,...,Q are the Q quadrature points and weights respectively from the Hermite quadrature method and the term $\frac{1}{\sqrt{\pi}}$ arises from a change of variables that is necessary to use the Hermite Quadrature method. The expected value is over the extreme value smoothing random variables that have been added to each current period utility equation. The benefit of adding the smoothing variable is that E[•] has a closed form solution that has continuous derivatives with respect to all of the parameters that are being estimated (Keane and Moffitt, 1998).

The expectations of the type in B.1 are calculated as part of the backwards recursion solution process. In order to solve all necessary value functions at each time t of the backwards recursion process, it is necessary to know value functions at time t+1 for each combination of the state variables in S(t+1)={6, X(t+1), EXC(t+1), EXB(t+1), SSEX(t+1), DI(t+1), ϵ(t+1), ψ(t+1), L(t+1), κ, η_t , HI^C, and HI^B(t)} that could arise at time t+1.

The variables 6, X(t+1), HI^C, ϵ(t+1), and ψ(t+1) are not computationally burdensome from the standpoint of solving value functions. The first three of these variables are not computationally burdensome because they are assumed to be exogenous and predetermined. This implies that value functions at time t+1 need to be solved only for the known value of these variables. The last two of these variables are not computationally burdensome because they are assumed to be serially independent, and, as a result, influence t+1 value functions only through the current period utility.[53]

The computational burden of the DP model arises primarily from the variables EXC(t+1), EXB(t+1), SSEX(t+1), DI(t+1), HI^B(t+1), η_{t+1}, and κ. For each of these variables, the computational burden arises because 1) there are multiple values for which value functions are needed at time t+1 and 2) the current period value of the variable provides information about both current and future utility.[54] It is worth noting that the reasons that multiple values need to be solved varies somewhat across the variables. For the first six variables, the agent needs to solve for the multiple values as part of his decision process. However, this is not the case for κ since each person is assumed to know his person-specific value. Instead, the necessity of solving value functions for multiple values arises as an estimation issue because the econometrician does not observe the person-specific value. The method for dealing with this "missing data" problem is an estimation which, as described in Section 3, implies that the econometrician needs to solve value functions for different values of κ.

---

[53]Thus, the implication that the time t+1 value functions need to be known for many values of ϵ(t+1) and ψ(t+1) is not problematic since, for any j, knowledge of V_j(t+1,S(t+1))= R^j(S(t+1))+β(S(t+1))·E[V(t+2,S(t+2))|S(t+1),d^j(t+1)=1] for any values of ϵ^j(t+1) and ψ^j(t+1) implies that V_j(S(t+1),t+1) can be found for any other values of ϵ^j(t+1) and ψ^j(t+1) by simply recalculating the current period reward R_j(S(t+1)). This calculation is not computationally demanding.

[54]The latter characteristic of these variables implies that, for any option j, computing a time t+1 value function for any particular combination of these variables requires that one recompute the computationally demanding βE[V(t+2,S(t+2))|S(t+1),d_j(t+1)=1].

The first five of the variables in the previous paragraph are discrete variables that are determined endogenously by individual decisions, and the possible combinations of these variables at any time t+1 can be easily characterized.[55] The simulation approach for dealing with unobserved heterogeneity which is described in Section 3 determines a finite number of values for κ for which value functions need to be solved. However, η_{t+1} is a serially correlated continuous variable and this causes well-known difficulties for the backwards recursion solution method. As discussed in detail in Keane and Wolpin (1994), Rust (1997), and Stinebrickner (2000), quadrature or simulation methods such as those mentioned earlier in this section are a useful tool for addressing the difficulties of serially correlated, continuous variables because, in effect, they serve to discretize the state space - an obvious necessity given that the backwards recursion process requires that value functions be solved for all combinations of state variables. Unfortunately, while finite, the number of possible values of η_{t+1} for which value functions need to be solved at time t+1 is very large for all but the smallest values of t+1.[56] As a result, for all but the smallest values of t+1 it is infeasible to solve value functions using standard methods for all possible combinations of EXC(t+1), EXB(t+1), SSEX(t+1), DI, HI^B(t+1), κ and η_{t+1} that could arise. We address this issue by implementing a modified version of the standard backwards solution process. The first step in the modified backwards recursion approach, which takes place before the backwards recursion process begins, involves determining the range of possible values that η_t could have in t=1,2,...,T. The set of possible values of η_1 are determined by the simulation process (described in Section 3) that is needed to compute the likelihood contribution of the person given the reality that true health in t=1, η_1, is not observed. Given the range of possible values of η_1, equation (B.3) can be used to determine the possible values for η_2 that are needed to compute the future component of time value t=1 functions associated with these values. Additional values of η_2 are generated by the simulation process associated with the likelihood contribution which takes into account that η_1 is not observed. The range of possible values of η_2 can be constructed from the set of all possible values of η_2, and this process can be repeated one period at a time to determine the possible range of values of $\eta_3,...,\eta_T$.

Once the range of values for $\eta_1,...,\eta_T$ have been determined, the modified backwards recursion process can take place. At each time t in the backwards recursion process, rather than solving value functions for all possible values of η_t, value functions are solved for the largest possible value of η_t, the smallest possible value of η_t, and some subset of the possible values in between. We refer to these values of η_t for which value functions are solved at time t as the time t grid points and denote them $\eta_t^{*1},...,\eta_t^{*N^\eta}$ where N^η is the total number of grid points at time t.[57] Equation (B.3) indicates that solving the value functions associated with the grid points at time t requires knowledge of value functions at time t+1 for various values of η_{t+1}. The reality that these values of η_{t+1} will not correspond to the time t+1 grid points (for which value functions were

[55] If a person leaves the workforce at an age of 70 or older or if a person successfully applies for Disability Insurance he enters a terminal state in which he remains out of the workforce for the remainder of his lifetime. These terminal state assumptions reduce computational burden. For example, the latter implies that value functions C, B, and A do not need to be solved at time t unless DI(t)=0. The former implies that value functions C, B, and A do not need to be solved for any combinations of state variables that indicate that a person left the workforce at any year when he was age 70 or older.

[56] To see this note that Equation (B.3) in appendix B indicates that solving value functions for a particular value of η_1 would require that value functions be computed at time t=2 for the Q values of η_2 given by $\rho(\eta_1 - \pi X_1) + \pi X_2 + \sqrt{2}\sigma_\xi m_q$ q=1,...,Q. Similarly, solving each of these time t=2 value functions would require that time t=3 value functions be computed for Q values of η_3. Thus, solving value functions for a particular value of η_1 would require that we compute value functions at any time t+1 for Q^t values of η_{t+1}. Further, as described below, our methods for dealing with the reality that true health in period 1, η_1, is unobserved requires that we solve value functions at t=1 for multiple values of η_1.

[57] The number of grid points can change with t. In practice, to reduce computational expense slightly we allowed the spacing between grid point to increase as t increases and for values of η that are a long way from the mean of η for the sample. The spacing we choose between grid points is .35 for the years in which choices are observed which is approximately .30 of a health standard deviation. Stinebrickner (2000) suggests that the methods perform well even when grid points are spaced substantially further apart than this. Consistent with this, we find that reducing spacing further had virtually no effect on parameter estimates.

solved at time t+1) necessitates a value function approximation. Specifically, we interpolate the t+1 value function associated with a particular value of η_{t+1} as the weighted average of the value function associated with the smallest grid point at time t+1 which is larger than η_{t+1} and the value function associated with the largest t+1 grid point which is smaller than η_{t+1}. This nonparametric linear interpolation approach based on "surrounding" grid points has the virtue that the interpolated value function for η_{t+1} converges to the true value function as the number of grid points is increases (i.e., as the grid points used in the weighted average become close to the value of η_{t+1} for which value functions are being approximated).[58]

[58]Keane and Wolpin (1994) describe a value function approximation technique based in Ordinary Least Squares and demonstrate the usefulness of this in several applications including Keane and Wolpin (1997). An alternative for solving value functions in the presence of serially correlated state variables that does not require value function approximation is the self-interpolating approach suggested by Tauchen and Hussey (1991) and Rust (1997). Stinebrickner (2000) provides some intuition about cases in which one might expect the interpolating method to perform better and cases in which one might expect the self-interpolating method to perform better. Stinebrickner (2000) suggests that the interpolating method has an advantage in cases such as this where the degree of serial correlation is high.

Appendix C. The Likelihood function

In either the case of a person who is in our behavioral sample or a person who is not in our behavioral sample, the first conditions that must be satisfied come from the fact that, in each year that a person reports his work limitation status, this report dictates an interval in which the latent health index h_t^* from equation (12) must lie. While individuals make decisions annually in our behavioral model, health is only reported at survey dates. Thus, given that we are using three waves of HRS data after the baseline wave, we observe up to three health observations per person after the baseline period. In Section 4.2 we discuss the timing of survey dates in more detail. For concreteness at this point, assume that the person's health status is reported three times at t=2, t=3, and t=5, and that the reports indicate that the person's work is not limited at t=2 and t=3 but is limited at t=5. Then, health equation (12) implies the three conditions that must be satisfied:

$$\text{(C.1)}\quad \begin{aligned} h_2^* &= \pi X_2 + \gamma Z_2 + [\nu_2 + \mu_2] < 0 \Rightarrow \nu_2 + \mu_2 < -\pi X_2 - \gamma Z_2, \\ h_3^* &= \pi X_3 + \gamma Z_3 + [\nu_3 + \mu_3] < 0 \Rightarrow \nu_3 + \mu_3 < -\pi X_3 - \gamma Z_3, \\ h_5^* &= \pi X_5 + \gamma Z_5 + [\nu_5 + \mu_5] < 0 \Rightarrow \nu_5 + \mu_5 < -\pi X_5 - \gamma Z_5 \end{aligned}$$

For each person, an additional condition that must be satisfied comes from the initial conditions equation. For a person who is working at baseline, equation (14) implies that

$$\text{(C.2)}\quad \epsilon^I > -\Pi_1 X_0 - \Pi_2 Z_0.$$

The direction of the inequality is reversed for individuals who are not working at baseline.

The likelihood contribution for an individual who is not in the behavioral sample, comes exclusively from the health and initial conditions observations. For an individual who is in our behavioral sample, additional conditions come from the expression for the behavioral choice that a person makes in each year a choice is observed. For concreteness assume that the person is younger than 61 years of age at t=1, five behavioral decisions are observed and the decisions are to work in his career job in years t=1,2,3, to work in a bridge job in year t=4 and to be out of the workforce in year t=5.[59] In Section 4.2 we describe how we map the continuous work histories that can be constructed from the data to the sequence of yearly decision periods that are each characterized by a single activity status.

Starting with t=1, the fact that the person chooses the Career option implies that the value of this option is greater than the value of each of the other options for this individual at time t=1,

$$\text{(C.3)}\quad V_C(1,S(1))>V_B(1,S(1)),\ \ V_C(1,S(1))>V_A(1,S(1)), \text{ and } V_C(1,S(1))>V_N(1,S(1)).$$

Define $V^*_j(t,S(t),\psi_t,\eta_t)=V_j(t,S(t),\psi_t,\eta_t)-\kappa^j- \lambda_\eta^j \nu_t -\epsilon_t^j$ for j=C,B,A,N, where for expositional reasons we have included ψ_t and η_t explicitly in V* even though they are both elements of S(t). Then, the three conditions in equation (C.3) can be rewritten as

$$\text{(C.4)}\quad \kappa +\lambda_\eta^C\nu_1+\epsilon_1^C - \kappa^j- \lambda_\eta^j\nu_1-\epsilon_1^j >V^*_j(1,S(1),\psi_1,\eta_1)-V^*_C(1,S(1),\psi_1,\eta_1), \qquad j=B, A, N.$$

For t=2 and for t=3, three conditions of identical form exist with the exception that all variables would be indexed by time t=2 and t=3 respectively rather than t=1.

At t=4, the person chooses the bridge option and the three conditions analogous to those in equation (C.4) indicate that the value of the bridge option at t=4 is greater than the value of the three other options at t=4:

---

[59]The reason to specify the age of the example person is that a person's age has an effect on his choice set. Given the specified age, the example person can apply for DI (option A) in all periods t=1,2,3,4,5..

(C.5) $\kappa+\ \lambda_\eta^B \nu_4+\epsilon_4^B\ -\kappa^j-\ \lambda_\eta^j \nu_4-\epsilon_4^j > V^*_j(4,S(4),\psi_4,\eta_4)-V^*_B(4,S(4),\psi_4,\eta_4)$ j=C,A,N.

Finally, at t=5, the person no longer has the option to work in his career job because he left his career job at time t=4. The two conditions analogous to those in equation (C.4) indicate that the value of not working (N) at t=5 is greater than the value of the two other available options at t=4:

(C.6) $\lambda_\eta^N \nu_5+\epsilon_5^N\ -\kappa^j-\ \lambda_\eta^j \nu_5-\epsilon_5^j > V^*_j(5,S(5),\psi_5,\eta_5)-V^*_N(5,S(5),\psi_5,\eta_5)$ j=B, A

Thus, the likelihood contribution L_i for our example person i is given by the joint probability that the eighteen conditions (3 health conditions, 1 initial condition, and 14 behavioral conditions) in equations (C.1), (C.2), (C.4), (C.5), and (C.6) are jointly true. The ψ_t's are independent of all other stochastic elements that enter these equations so that the likelihood contribution of person i can be written as

$$
\begin{aligned}
\text{(C.7) } L_i= \int PR(&\nu_2+\mu_2<-\pi X_2-\gamma Z_2 \cap \nu_3+\mu_3<-\pi X_3-\gamma Z_3 \ \cap\ \nu_5+\mu_5>-\pi X_5-\gamma Z_5\ \cap\ \epsilon^I>-\Pi_1 X_0-\Pi_2 Z_0 \ \cap \\
&\kappa+\lambda_\eta^C \nu_1+\epsilon_1^C - \kappa - \lambda_\eta^B \nu_1-\epsilon_1^B > V^*_B(1,S(1),\psi_1,\eta_1)-V^*_C(1,S(1),\psi_1,\eta_1) \cap \ldots \cap \\
&\lambda_\eta^N \nu_5+\epsilon_5^N - \lambda_\eta^A \nu_5-\epsilon_5^A > V^*_A(5,S(5),\psi_5,\eta_5)-V^*_N(5,S(5),\psi_5,\eta_5) \quad dF(\psi_1,\ \psi_2,\ \psi_3\ \psi_{4\ ,}\psi_5).
\end{aligned}
$$

The integral in equation (C.7) can be simulated in a straightforward manner given the empirical distribution of $\psi_1,...,\psi_5$. What remains is to show how to compute the probability expression in (C.7) given values of the ψ's. The joint probability is the area under the joint density of $\Psi=\{\ \nu_2+\mu_2,\ \nu_3+\mu_3,\ \nu_5+\mu_5,\ \epsilon^I,\ \kappa+\lambda_\eta^C \nu_1+\epsilon_1^C -\kappa - \lambda_\eta^B \nu_1-\epsilon_1^B, \kappa+\ \lambda_\eta^C \nu_1+\epsilon_1^C - \lambda_\eta^A \nu_1-\epsilon_1^A, \ldots,\ \lambda_\eta^N \nu_5+\epsilon_5^N -\kappa-\ \lambda_\eta^B \nu_5-\epsilon_5^B,\ \lambda_\eta^N \nu_5+\epsilon_5^N -\lambda_\eta^A \nu_5-\epsilon_5^A\}$ where the conditions that appear in the probability expression (i.e., the conditions in equations C.1, C.2, C.4, C.5, and C.6) are all true.

The eighteen elements of Ψ have a joint normal distribution. The diagonal elements of the covariance matrix are determined by the previously discussed normalizations $var(\nu_t+\mu_t) = var(\epsilon^I) = var(\epsilon_t^C) = var(\epsilon_t^B) = var(\epsilon_t^N) = var(\ \epsilon_t^A) = 1$, by σ_κ and $var(\nu_t)$, and by the coefficients λ_η^C, λ_η^B, λ_η^N, λ_η^A. The non-diagonal elements of the covariance matrix are determined by ρ, by the covariance terms $COV(\epsilon^I, \nu_1)$, $COV(\epsilon^I, \kappa)$, $\{COV(\mu_t, \epsilon_t^j)$, j=C,B,A$\}$, by the normalization of $COV(\mu_t, \epsilon_t^N)$ to zero, and by the fact that ν_t, t=1,2,3,4,5 enters multiple elements of Ψ.

Given the covariance matrix that can be computed given the model parameters at a given iteration of our updating algorithm, we use the GHK simulator of Geweke (1991), Hajivassiliou (1990), and Keane (1994) to approximate the probability expression in equation (C.7). Roughly speaking, this approach centers around the rewriting of the joint probability in equation (C.7) as a product of conditional probabilities and approximating this product by repeatedly computing marginal probabilities, drawing values of random variables consistent with conditions that must be satisfied, and updating conditional distributions. It is worth noting that the GHK simulation approach provides a very natural way to deal formally with important missing data issues that arise in this context (see Stinebrickner, 1999, and Lavy et al., 1998). Consider the likelihood contribution for our example person. The person's true health η_t in t=1,2,3,4,5 enters the likelihood contribution. However, η_t is never observed since, even when a self-report of health occurs at time t, the self-report represents only a noisy measure of true health. In essence, our approach allows us to deal with the missing data problem by "integrating out" over the joint distribution of η_1, η_2, η_3, η_4, ..., η_5 that is appropriate given a person's self-reported health in t=2, t=3, and t=5.

Specifically, he likelihood in equation (C.7) can be approximated as

$$
\text{(C.8)} \quad \frac{1}{D}\sum_{d=1}^{D}\{
$$

$$
\begin{aligned}
&PR(\nu_2+\mu_2<-\pi X_2-\gamma Z_2)\bullet\\
&PR(\nu_3+\mu_3<-\pi X_3-\gamma Z_3\,|(\nu_2+\mu_2)^d)\bullet\\
&PR(\nu_5+\mu_5<-\pi X_5-\gamma Z_5\,|(\nu_2+\mu_2)^d,(\nu_3+\mu_3)^d)\bullet\\
&PR(\epsilon^I>-\Pi_1X_0-\Pi_2Z_0|(\nu_2+\mu_2)^d,(\nu_3+\mu_3)^d,(\nu_5+\mu_5)^d,\nu_1{}^d,\nu_2{}^d,\nu_3{}^d,\nu_4{}^d,\nu_5{}^d,\kappa^d)\bullet\\
&PR(\kappa^d+\lambda_\eta{}^C\nu_1{}^d+\epsilon_1{}^C-\kappa^d-\lambda_\eta{}^B\nu_1{}^d-\epsilon_1{}^B>V^*{}_B(1,S(1),\psi_1{}^d,\eta_1{}^d)-V^*{}_C(1,S(1),\psi_1{}^d,\eta_1{}^d|\\
&\qquad(\nu_2+\mu_2)^d,(\nu_3+\mu_3)^d,(\nu_5+\mu_5)^d,\nu_1{}^d,\nu_2{}^d,\nu_3{}^d,\nu_4{}^d,\nu_5{}^d,\kappa^d,\epsilon^{Id})\\
&\bullet\\
&\bullet\\
&\bullet\\
&PR(\lambda_\eta{}^N\nu_5{}^d+\epsilon_5{}^N-\lambda_\eta{}^A\nu_5{}^d-\epsilon_5{}^A>V^*{}_A(5,S(5),\psi_5{}^d,\eta_5{}^d)-V^*{}_N(5,S(5),\psi_5{}^d,\eta_5{}^d|\bullet)\;\}
\end{aligned}
$$

where the superscript d denotes the dth of D total draws of a particular variable. $\psi_t{}^d$ t=1,2,3,4,5 is a draw of ψ_t from its empirical distribution as part of the simulation of the integral in equation (C.7), $(\nu_2+\mu_2)^d<0$ is drawn from the marginal distribution of $\nu_2+\mu_2$ given the joint density of the eighteen elements in Ψ. $(\nu_3+\mu_3)^d<0$ is drawn from the marginal distribution of $\nu_3+\mu_3$ given $(\nu_2+\mu_2)^d$. $(\nu_5+\mu_5)^d>0$ is drawn from the marginal distribution of $\nu_5+\mu_5$ given $(\nu_2+\mu_2)^d$ and $(\nu_3+\mu_3)^d$. $\nu_1{}^d,\ldots,\nu_1{}^5$ are drawn unconditionally from the marginal distribution of $(\nu_1{}^d,\ldots,\nu_1{}^5)$ given $(\nu_2+\mu_2)^d$, $(\nu_3+\mu_3)^d$, and $(\nu_5+\mu_5)^d$. Given $\nu_1{}^d,\ldots,\nu_1{}^5$, the values of $\eta_1{}^d,\ldots,\eta_5{}^d$ can be constructed using equation (10).[60] κ^d is drawn from its unconditional distribution (it is not correlated with any of the previously simulated values). $\epsilon^{Id}>-\Pi_1X_0-\Pi_2Z_0$ is drawn from the marginal distribution of ϵ^{Id} given $(\nu_2+\mu_2)^d,(\nu_3+\mu_3)^d,(\nu_5+\mu_5)^d,\nu_1{}^d,\nu_2{}^d,\nu_3{}^d,\nu_4{}^d,\nu_5{}^d,\kappa^d$. This process continues until all eighteen conditional probabilities that appear in equations (C.7) and (C.8) are computed. The probability expression on the last line of equation (C.8) is conditioned on all of the unobservables that have been simulated in order to compute the previous seventeen conditional probabilities.

One adjustment is needed to make the likelihood calculation in equation (C.8) feasible. Recall from Appendix B that our adjustment to the backwards recursion method implies that $V^*{}_j(\cdot,\eta_1)$ j=C,B,N,A will be solved for only a set of grid points $\eta_1^{*\,1},\ldots,\eta_1^{*\,N^\eta}$ which will not be the same as $\eta_1{}^d$, d=1,...,D. This implies that we must use an approximated value of $V^*{}_j(\cdot,\eta_1{}^d)$, j=C,B,N,A for each d=1,...,D in equation (C.8). Consistent with what is done in Appendix B we interpolate $V^*{}_j(\cdot,\eta_1{}^d)$ as the weighted average of the values of $V^*{}_j$ associated with the largest grid point less than $\eta_1{}^d$ and the smallest grid point greater than $\eta_1{}^d$.

[60]One minor complication is that the health conditions Z_t from equation (10) are not observed in the data in non-survey years (in our example, years 1 and 4). We simulate the missing γZ_t's (in our example,we would simulate γZ_1 and γZ_2).

Appendix D. Computational Time Issues

In order to make estimation feasible we paid close attention to certain properties of the likelihood function - specifically we made sure to specify our model in a way such that the derivatives of the likelihood function with respect to the parameters above are continuous and smooth. This has substantial benefits because it allows the use of a derivative-based Newton-Raphson parameter updating algorithm which requires many fewer likelihood function evaluations (and are therefore many times faster) than algorithms that do not take advantage of derivatives (e.g., simplex methods). Nonetheless, even with an updating algorithm that does not require large number of likelihood function evaluations, estimation of our specified model would not be feasible if we were to estimate the model on the fastest single windows-based CPU that is currently available using numerically computed derivatives of the likelihood function with respect to model parameters. Specifically, we found that using this approach would require between 100 and 200 days to compute a single iteration of the Newton-Raphson algorithm.

In order to make estimation feasible we took several steps. First, rather than relying on numerical derivatives (of the likelihood function with respect to model parameters) which require a likelihood function evaluation for each parameter in the model we took the painstaking approach of programming analytic derivatives for all parameters that enter the model.[61] Second, rather than relying on a single processor we took advantage of the parallel processing capabilities of one of the fastest academic supercomputers in North America which provided us with exclusive access to 48 Compaq Alpha CPU's for an extended period of time.[62] Finally, as described in Section 4.4, we estimated certain parameters separately from the structural model. These efforts reduced the time necessary to compute one iteration of the likelihood function to approximately 9 hours which made estimation time-consuming but feasible given our Newton-Raphson updating algorithm.

---

[61]The amount of time needed to compute an analytic derivative is much smaller than that needed to compute a numerical derivative (which requires at least 1 additional likelihood function evaluation for each parameter). One reason for this is that the analytic derivatives can typically be written as functions of information that has already been computed when evaluating the likelihood function.

[62]See Swann (2002) for a user-friendly description of how to use parallel processing techniques with Maximum Likelihood estimation algorithms. For some preliminary explorations we used as many as 160 CPU's. The final numbers in this paper were generated using 48 CPU's.

Appendix E. The Tradeoff Between Model "Size" and Approximation Quality

It is worth noting that, while the choice set in our model allows individuals to consider a set of work and non-work activity statuses that is more detailed than that which has typically been allowed in the retirement literature, it is somewhat parsimonious in the sense that we do not model other types of individual decisions that take place during the later part of a person's working life. As one example, as discussed in Section 2.1, we do not formally model the savings decisions of individuals.

While we recognize that there is a cost to this parsimony, the specification of the choice set was motivated by the potentially significant cost to expanding the choice set to allow for more endogenous choices. As discussed in detail in Appendix D, even after taking very substantial steps to reduce computational time as much as possible, we are currently at the computational edge. As a result, the computational increase associated with the most obvious expansions of the choice set and the change in the number of state variables that would accompany such an expansion would imply the necessity of adopting an approximation method in which value functions are solved for only a small subset of the state points in the model. While other researchers have recently chosen to take this route, in practice evaluating the quality of approximation methods that are available when value functions can be solved for only a very small subset of the state space is extremely difficult for a particular application, and very little evidence exists in the literature about this issue more generally.[63] It was the virtue of avoiding uncertainty about approximation quality that led to the specification of a parsimonious choice set.[64] Nonetheless, while we think our choices are reasonable, we stress that we have no evidence (or overly strong feelings) about the benefits of our decision relative to the alternative.

[63]While other work also relates to approximation quality (Rust, 1997, Stinebrickner, 2000), the most relevant evidence and most used methods related to the case of a very large state space is Keane and Wolpin (1994). The majority of the discussion and evidence in that paper relates to a case where the state space is not very large, but functional form assumptions imply that closed form solutions do not exist for the value function in a model. As detailed in that paper, when the state space is very large (so that it is not possible to solve or interpolate value functions for all points in the state space), it is not possible to use the specification of the interpolating function that was found to be desirable in the case where the state space is not very large. The two tested interpolating functions that are available in the case of a very large state space lead to biases of "substantial economic magnitude" in certain parameters.

[64]Given our specification which includes a continuous health measure it is not possible to avoid the issues of approximation quality altogether. However, given the nature of the approximations in our specification, we feel much more comfortable about our ability to specify a model in which little approximation error exists than we would be if we expanded our choice set and dealt with the case of a very state space. As one example, as described in Section 3 and appendix B, our approximation method for dealing with the serially correlated health variable involves interpolating in only a single dimension and we are able to use a straightforward nonparametric approach with desirable properties.