University of Michigan Retirement Research Center

Working Paper

WP 2004-091

Using a Structural Retirement Model to Simulate the Effect of Changes to the OASDI and Medicare Programs

John Bound, Todd Stinebrickner, and Timothy Waidman

MRRC Project #: UM04-06

"Using a Structural Retirement Model to Simulate the Effect of Changes to the OASDI and Medicare Programs"

John Bound
University of Michigan and NBER

Todd Stinebrickner
University of Western Ontario

Todd Stinebrickner
Urban Institute

October 2004

Michigan Retirement Research Center
University of Michigan
P.O. Box 1248
Ann Arbor, MI 48104

Acknowledgements

This work was supported by a grant from the Social Security Administration through the Michigan Retirement Research Center (Grant # 10-P-98358-5). The opinions and conclusions are solely those of the authors and should not be considered as representing the opinions or policy of the Social Security Administration or any agency of the Federal Government.

Using a Structural Retirement Model to Simulate the Effect of Changes to the OASDI and Medicare Programs

John Bound, Todd Stinebrickner, and Timothy Waidman

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. The model is estimated using data from the Health and Retirement Study. We use the model to simulate the impact on behavior of raising the normal retirement age, eliminating early retirement altogether and introducing universal health insurance.

Authors' Acknowledgements

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. 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 progam. 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 Chris Swann, Steve Haider, and Paul Contoyannis.

Section 1. Introduction

With an aging population on the one hand and difficulties in financing public and private pensions on the other, understanding the determinants of individuals' retirement behavior is of considerable research and policy importance. Much of the research on the labor force behavior of older, working aged adults has focused on the effects of financial incentives such as Social Security and private pensions, generally showing that these incentives have powerful behavioral effects. At the same time, econometric studies of retirement behavior have provided strong evidence for the importance of health factors

Unfortunately, efforts to understand the effects of financial incentives on retirement behavior and understand the effects of health on retirement behavior have remained largely separate avenues of research. There exist strong empirical correlations among health, socioeconomic status, and the rewards from continuing to work. Behaviorally, health status and economic factors interact in their effects on labor force behavior. For example, people in poor health will presumably continue to work unless they have the resources to permit them to stop, while those are who not able to stop working may adapt to their poor health in other ways, for instance by changing jobs. As a result, economic factors confound the effects of health on retirement, and vice versa.

In this paper we hope to improve understanding of the labor force behavior of older Americans by specifying and estimating a model of labor force behavior that builds on the strengths and addresses some of the weaknesses of the two largely separate literatures. The benefits of drawing on valuable lessons learned in each are readily apparent. On one hand, research examining the relationship between financial resources and retirement decisions has established the importance of viewing retirement as a dynamic process which can best be viewed longitudinally. Acknowledging this basic idea seems necessary if one wishes to fully understand the relationship between health and retirement. Although variation in mental health, cognitive functioning, and physical health exists at all ages and affects early educational and occupational attainment, it is the decline in physical and mental health starting in late middle age – often in combination with a changing occupational environment – that is likely to create a mismatch between an individual's capabilities and the requirements of his job. Whether and how workers respond to declines in health depends on various factors, including the nature of the declines, their expected persistence, the age at which they occur, and workers' human capital, economic situation, and preferences for leisure and consumption. Research on the effect of health on retirement has virtually ignored these dynamic issues.

On the other hand, research examining the relationship between health and retirement has stressed the importance of carefully dealing with issues related to the measurement of health with a specific focus on the potential problems associated with using global survey measures such as self-rated work limitations and global self-rated health.   There are a number of potential problems with such survey measures, particularly relating to measurement error and endogeneity. First, respondents are asked for subjective judgements which may not be entirely comparable across individuals. Second, responses may not be independent of the very labor market outcomes investigators hope to explain. Third, since health may represent one of the few legitimate reasons for working-age adults to be out of work, respondents out of the labor force may mention health problems to rationalize their behavior. Fourth, since early retirement benefits are often available only for those deemed incapable of work, respondents may have a financial incentive to identify themselves as disabled, an incentive that will be particularly high for those for whom the relative rewards from continued work are low. It is important to note that each of these problems will lead to different kinds of biases (Bound, 1991) and biased estimates of health's impact on outcomes will also lead to biased coefficients on any variable correlated with health.[1] Unfortunately, these issues have largely been ignored in most past longitudinal retirement research which has suffered from the availability of only very limited measures of health.[2]

The importance of dynamic considerations motivates our use of a dynamic, discrete choice model. This framework has been found to be useful in previous longitudinal studies of retirement such as Berkovec and Stern (1991) and Rust and Phelan (1997). However, within this general framework we embed a statistical model of health that has been used in the health literature and allows us to deal explicitly with the problems of endogeneity and measurement error that may arise from the use of global measures of self-reported health.

The estimation of the specified model requires a rich data source. We take advantage of the Health and Retirement Study (HRS) which, in an effort to allow researchers to address the types of research questions posed here, biennially surveys a sample of Americans who were aged 50 to 61 when the initial survey wave was collected in 1992. The HRS was designed with careful attention to the state of the art in measuring health status

---

[1]Evidence to date on the nature of the biases involved in using global self reported health measures (e.g.. responses to questions such as "Do you have any impairment or health problem that limits the kind or amount of paid work you can do?" or "Would you say your health is excellent, very good, good, fair or poor?" as are asked in the HRS) in behavioral models is mixed. Stern (1989) and Bound (1991), using instrumental variable procedures to deal with the potential endogeneity of global self reported health measures, found evidence that global measures neither dramatically over or under-estimated the effect of health on labor force participation in reduced form cross sectional models of labor force participation. Bound, Schoenbaum, Stinebrickner and Waidmann (1998) obtain similar results using longitudinal data from the Health and Retirement Study. Bound (1991) and Bound et al. (1998) interpret this result as suggesting that the endogeneity and errors-in-variables bias's approximately cancel each other out. However, even if this is the case, biases may remain in the estimated effects of any variables that are correlated with health. Kreider (1999) produces evidence suggesting that the use of global self-reported measures will tend to yield results that underestimate the impact of financial incentives on labor force behavior. The modeling strategy we follow in this paper allows for both random and systematic components to any measurement error in self reported health.

[2]In a recent paper Benitez-Silva, Buchinsky, Chan, Cheidvasser and Rust (2000) argue that, at least when studying the decision to apply for Social Security Disability Insurance, self reported measures of work incapacity (i.e. responses to the questions: "Do you have any impairment or health problem that limits the kind or amount of paid work you can do? If so, does this limitation prevent you from working altogether?") provide unbiased measures of work capacity. This finding is at odds with the common presumption in the literature, at odds with the evidence that, both across countries and across time within the U.S., the fraction of individuals reporting themselves unable to work mirrors the fraction of individuals receiving disability insurance (Waidmann, Bound and Schoenbaum, 1995; Bound and Waidmann, 2002; Burkhauser, Dwyer, Lindeboom, Theeuwes, Woittiez, 1999), and at odds with results reported in Bound et al. (1998). In particular, in this latter paper, strategies similar to some of those used by Benitez-Silva et al. indicate that, while a dichotomous measure based on the question "Do you have any impairment or health problem that limits the kind or amount of paid work you can do?" produces estimates that neither dramatically over or underestimate the impact of work limitations on retirement, a dichotomous measure that is coded 1 if the person identifies himself as unable to work and 0 otherwise (this is the same kind of measure as is used by Benitez-Silva et al.) tends to dramatically exaggerate the impact of work limitations on retirement. We suspect that the most important reason for the differences between the Bound et al. and the Benitez-Silva et al. findings has to do with differences between the exclusion restrictions used by the two sets of authors. In particular, Benitez-Silva et al. use a much sparser list of exclusion restrictions than do Bound et al. As a result, one concern is that the Benitez-Silva et al. tests may have low power.

in self-reported surveys. As such, it includes the type of detailed health information that is required to estimate the health portion of our model but has not been previously available in labor force surveys. Similarly detailed is information about the economic resources of survey respondents. Accurate information on all potential financial sources that could influence the retirement decision is important since, for reasons alluded to earlier, incomplete information about economic resources will tend to lead to biased estimates of the effects of both other financial variables and health on labor force participation. Finally, the HRS contains detailed information about a person's activity state in a particular year. While some individuals who encounter poor health may have the economic resources to simply leave the workforce, other individuals with insufficient economic resources may find that qualifying for Disability Insurance benefits represents the only plausible way to leave the workforce. Still others may try to adapt to their health status by finding a new job. While fully understanding the relationship between financial resources, health, and labor force participation would seem to require that one consider the variety of ways individuals adapt to poor health, to date very little research has attempted to model the effect of health on labor force transitions other than retirement.

We restrict our attention to single men. We restrict our attention to singletons to avoid complications associated with joint decision making. We restrict our attention to men since single women's Social Security benefits often depend on the earnings of their ex-husbands, and this is information the is unavailable to us.

The estimation of our specified model also requires state of the art econometric techniques. The health methods we embed in our dynamic programming models use a latent variable model to construct a continuous index of health. Past research suggests that it is generally reasonable to assume that this continuous health measure follows an autoregressive (1) process (Bound et al. 1999). However, this raises issues from both the standpoint of solving the value functions that are generated by the model and from the standpoint of estimating the model given the solved value functions. With respect the former, the use of serially correlated, continuous state variables has traditionally represented one of the big challenges for researchers working with these types of models (Rust 1997, Keane and Wolpin 1994, Stinebrickner 2000, Brien et al. 2001). With respect to the latter, the correlation of health across decision periods implies correlations between the composite random variables which enter the large number of conditions that determine the likelihood contribution of a particular person.

These issues, when combined with our desire to include unobserved heterogeneity, our use of up to six years of observed choices and an initial condition equation for each person, 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. Estimation of our model is made feasible by specifying the model in a way that allows for the use of derivative-based updating algorithms, our willingness to undertake the painstaking task of computing analytical derivatives of the likelihood function with respect to each of the parameters to be estimated, and by taking advantage of developments in parallel processing techniques and the exclusive access we were given during estimation to one of the fastest academic parallel processing supercomputers in North America.[3]

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, it is worth stressing that the specification of the choice set was motivated by the reality that there is also a potentially significant cost to expanding the choice set to allow for more endogenous choices. As discussed in Section 4, even with the efforts described above to make estimation feasible, we are currently at the computational edge with a single likelihood function iteration taking approximately nine hours. 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 changes 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

[3]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.

evidence exists in the literature about this issue more generally.[4] It was the virtue of avoiding uncertainty about approximation quality that led to the specification of a parsimonious choice set.[5] Nonetheless, we stress that we have no evidence (or particularly strong feelings) about the benefits of our decision relative to the alternative.

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 in detail the estimation methods we use. In Section 4 we present parameter estimates and simulations which highlight important aspects and implications of our model. In Section 5 we conclude.

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 we assume corresponds to the first wave (1992) wave 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, the group of individuals who are working at our

[4]While other work also relates to approximation quality (Rust, 2000, 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 the 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.

[5]Given our specification 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..

baseline time period is 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 the unobserved characteristics and preferences 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 attempt to address the former concern by adding a reduced form initial conditions equation that describes whether a person is working at our time period. We address the latter concern in a manner proposed by Bound (1991) 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.

In this section we describe the behavioral portion of our model and the methods we use to solve value functions given the presence of the serially correlated health variable. This discussion implicitly assumes that the true health (η_t) of each individual is known at each time t 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 deal with 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 t=1 (1993) and ending at year t=T.[6] 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 time t=0), B is the option of accepting a bridge job (defined to be a job other than the job held at baseline t=0), 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, 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 three 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 respect to this assumption, we allow for both the possibility that a person could leave his career job by choice and for 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 leaves the workforce at or after the age of 70, he will not return to work. Finally, 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). These assumptions imply that sufficient to characterize 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 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

[6]In practice, we assume that T is the year that the person turns 75 years old. After year T, individuals are assumed to remain out of the workforce for the remainder of their lives.

person consumes all of his "income" in year t. [7] As discussed in Section 2.2.1, we view pension wealth, non-pension wealth, Social Security earnings, Disability Insurance payments, and other 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 individuals applying for Social Security Disability benefits will incur financial costs. For the year of application they will forgo all earnings. Further, if their applications are rejected, they will not be able to return to their 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 earnings 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 is given by

(1) $\sum_{j \in D_t} R^j(t) d^j(t)$

where $R^j(t)$ is the current period reward associated with the jth alternative at time t. The reward $R^j(t)$ 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) $R^j(t) = U^j_{cons}(t) + U^j_{np}(t)$ j=C,B,N,A.

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

---

[7] In recent work researchers have begun to introduce savings into dynamic programing models of retirement (e.g . French, 2000a,b; van der Klaauw and Wolpin, 2002; Rust, Buchinsky and Benitez-Silva, 2001). 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 render our already complicated model intractable. Given our interest in the the interplay among health, financial resources, and the labor market behavior , we believe our choice was a reasonable one.

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

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

where θ determines the level of risk aversion and τ (along with parameters in the non-pecuniary utility equation (6) 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 wages, Social Security entitlements, defined benefit and defined contribution pension plans, Disability Insurance payments, non-pension wealth, food stamps, Supplemental Security Income, and other exogenous sources of income (such as veteran benefits) minus expenditures on health care. This approach to modeling the effect of the availability health insurance is now common in the literature (Rust and Phelan, 1997; Blau and Gilleskie, 2001; French and Jones, 2000).

Given our assumption that an individual consumes all of the income that he "receives" in a particular year, characterizing consumption in a particular year requires that we describe our assumptions related to the timing of the income from the sources above and describe any stochastic processes that influence these sources. These issues are discussed in the next two subsections.

Timing Issues Related to Income 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, the reality that defined benefit pension plans do not typically include an actuarial adjustment for delayed receipt implies that individuals will begin receiving defined benefit 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 defined benefits are 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 foodstamps, 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. This additional savings is used to increase consumption when he retires permanently (i.e., takes a year off at age 70 or older).[8]

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.[9]

Specification Issues Related to Processes that Influence Net Income

[8]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.

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

Following French and Jones (2000), we assume that a person's health expenditures at time t depends on the person's health insurance status k, the person's age, and the person's health:

(4) Ln Expenditures(t) = $\Lambda_1^k + \Lambda_2$ (Age-53) + $\Lambda_3\, \eta_t + e^k$

where we differentiate between four possible health insurance categories - k$\in${EHI (employer provided health insurance) only, EHI 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.

As a final specification decision related to current period income, we assume that the log of a person's real wage in his career job at time t evolves according to the fixed-effect specification

(5a) $\mathrm{Ln}(W^C_t) = \alpha_1^i + \alpha_2 t + \psi_t^C \quad \Rightarrow \quad W^C_t = e^{\alpha_1^i + \alpha_2 t} e^{\psi_t^C}$

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

(5b) $\mathrm{Ln}(W^B_t) = \alpha_3 \mathrm{EXITAGE} + \alpha_4 \ln(W_o^C) + \psi_t^B \quad \Rightarrow W^B_t = e^{\alpha_3 \mathrm{EXITAGE} + \alpha_4 \ln(W_o^C)} e^{\psi_t^B}$

where ψ^B_t is the stochastic component of bridge wages, and EXITAGE is the age at which a person left his career job. Our specification allows for the possibility that the wage 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 assumption 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.

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. We allow health insurance status to have a direct effect on a persons utility rather than simply allowing health insurance status to effect disposable income because there is ample evidence that health insurance effect health care utilization. Presumably this implies that health insurances effect on out of pocket medical expenses understates its effect on utility. 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.

$$(6) \qquad U_{np}^{j}(t) = \lambda_X^j X(t) + \lambda_{HI} HI^{j}(t) + \kappa\, 1(j = C \text{ or } B) + \lambda_\eta^j \eta_t + \varepsilon_t^j \quad \text{for } j = C,B,A,N.$$

where 1(•) is an indicator function that takes on a value of unity if the condition in parentheses is true.

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.

2.2.3 Summary of $R^j(t)$

To summarize,

$$(7) \qquad R^{j}(t) = U_{cons}^{j}(t) + U_{np}^{j}(t) = \tau \frac{Y^{j}(t)^{1-\theta}}{1-\theta} + \lambda_X^j X_t + \lambda_{HI} HI^{j}(t) + \kappa\, 1(j = C \text{ or } B) + \lambda_\eta^j \eta_t + \varepsilon_t^j, \quad j = C,B,N,A$$

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

2.3 *Discounted Expected Utility - Value Functions*

2.3.1 Specification of Value Functions

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):

$$(8) \qquad V_j(t, S(t)) = R^j(S(t)) + \beta(S(t)) \cdot E[V(t+1, S(t+1)) | S(t), d^j(t) = 1]$$

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:

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

As we discuss in Section 4, we examine the robustness of our results to different values of β^{Common}. 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.[10]

2.3.2 State Variables

The set of state variables at time t, S(t), includes all variables that provide information about the 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 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 (6), indicates that non-pecuniary utility at time t is determined by X(t), $\varepsilon(t)$, κ, η_t , and HI(t) where $\varepsilon(t)\subset\{\varepsilon^C_t,\varepsilon^B_t,\varepsilon^N_t,\varepsilon^A_t\}$ 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,

[10]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 Bound et. al (??) for details.

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. ε(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 $\varepsilon_t^j \sim N(0,1)$, j=C,B,A,N and that $E(\varepsilon_t^j,\varepsilon_r^k)=0$ if $j \neq k$ or $t \neq r$.[11]

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:

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

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 (10) 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.

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

[11]Normalizing the variance implies that a constant can be estimated as part of λ_X^j in equation (8).

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 $HI^B(t)=2$; or 4). $Age(t+1)\geq 65$. 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 final condition is present because everyone who has turned 65 years of age receives medicare. 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 some probability that the bridge offer he receives in time t+1 will include health insurance. [12] 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.[13]

Thus, at time t, the state variables that influence non-pecuniary utility are {$X(t)$, $\varepsilon(t)$, 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. Specifically, as discussed at the end of Section 2.2.1, $X(t)$, η_t , HI^C, and $HI^B(t)$ 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, θ, 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

---

[12] In practice, 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. ??John or Tim, include a discussion of where these numbers came from

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

information about a person's wealth at time t=1. In addition, because θ 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: Wage earnings, the SS and DI systems, and DB and DC pension plans. Below we describe the state variables that are needed (in addition to θ) to characterize what a person knows about the income from each of these sources.

The wage earnings equations are given in Section 2.2.1. Earnings in career jobs depend on a fixed effect (which can be viewed as an element of θ) and ψ^C_t. Earnings in bridge jobs depend on W^C_O (which is contained in θ), ψ^B_t, and the age at which a person left his career job. Sufficient for the latter is the person's age at baseline (contained in X) and the number of years of experience that he person worked in his career job as of time t (which we refer to as EXC(t)).[14]

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.[15] In this case, sufficient for computing the SS

[14]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..

[15]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 wages (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 or bridge wages (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 equation (5a) is relatively small. Second, over several years, positive shocks to wages in some years tend to be offset by negative shocks to wages 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.

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*). These three state variables are endogenously determined within the model.[16]

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, EXC(t*).[17] 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*), and we assume that an applicant at time t* is approved for DI benefits if

(11) $$A_1^{DI} + A_2^{DI}\eta_{t*} + e^{DI} > 0$$

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.

---

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

[17]The year that an individual left his career job is given by the amount of experience that a person had at the end of his decision making horizon, EXC(T), since individuals cannot return to career jobs after leaving these jobs.

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

2.3.3 Solving value functions

The expected value in equation (8) 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), $\varepsilon(t+1)$, $\psi(t+1)$, 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.[18]

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., $\varepsilon(t)$ in our case) to be iid extreme value, the expected value in equation (8) has a closed form solution conditional on the values of the other state variables.[19] However, in this application, the Section 2.3.2 normality assumption for $\varepsilon(t)$ 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 (10) assumption about the distribution of η_{t+1} given η_t and our use of the empirical distribution of ψ, implies that the expected value in equation (8) 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 feasibility, the derivatives of the expected value "approximator" are continuous with respect to model parameters.

[18]If the person has chosen a time t option other than CAREER, L(t+1) is not relevant and the dimensions of $\varepsilon(t+1)$ and $\psi(t+1)$ are reduced by one. In addition, if a person is working in a bridge job with health insurance in time, 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.

[19]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.

2.3.3 Backwards Recursion Solution Method and Adjustments to Deal with Serially Correlated Health

The recursive formulation of value functions in equations (8) motivates the backwards recursion solution process that is standard in finite horizon, dynamic, discrete choice models. The most basic property of this 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)={β, 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 β, 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.[20]

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 tthe variable provides information about both current and future utility.[21] 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

---

[20]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 $V_j(t+1,S(t+1))= R^j(S(t+1))+\beta(S(t+1))\cdot E[V(t+2,S(t+2))|S(t+1),d^j(t+1)=1]$ for any values of $\varepsilon^j(t+1)$ and $\psi^j(t+1)$ implies that $V_j(S(t+1),t+1)$ can be found for any other values of $\varepsilon^j(t+1)$ and $\psi^j(t+1)$ by simply recalculating the current period reward $R_j(S(t+1))$. This calculation is not computationally demanding.

[21]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 $\beta E[V(t+2,S(t+2))|S(t+1),d_j(t+1)=1]$.

value. The method for dealing with this "missing data" problem is an estimation which, as described in Section 3, requires that the econometrician needs to solve value functions for different values of κ.

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.[22] 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 and detailed in Bound et al. (??) 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 value values of t+1.[23] 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. At each time t+1 of the backwards recursion process, we solve and store value functions for a computationally feasible set of η_{t+1} values that we call "grid points." At each time t of the backwards recursion process, the solved value functions associated with these grid points serve as inputs into a

---

[22]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.

[23]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.

non-parametric value function approximation which allows us to approximate the time t+1 value functions associated with all values of η_{t+1} that arise when computing the expected future utility component of the time t value functions using equations such as equation (8). This approach is described in detail in Bound et. al (??).[24]

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 describe our simulated maximum likelihood estimation approach.

3.1 Health

The strategy we have chosen to deal with the reality that true health is not observed is to 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., specific health conditions), Z_t ; and other unobserved factors ν_t.

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

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

(13) $$h_t^*=\eta_t+\mu_t.$$

[24]The value functions for some value of time t+1 health are interpolated using weighted average of the value functions associated with health grid points that are nearby.

We assume that μ_t and η_t are uncorrelated. Substituting equation (12) into equation (13), we get

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

If $\nu_t+\mu_t$ is assumed to be normally distributed with a variance that is normalized to be one, equation (14) represents a probit model in which h_t* is greater than zero if the person reports that he is work limited. The relative size of var(ν_t) and var(μ_t) is not important for the estimation of π and γ in equation (14) 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 (14), $\nu_t+\mu_t$, 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 (7) 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(μ_t, ε_t^N)=0) and estimate the three covariance parameter COV(μ_t, ε_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*. This 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 model to 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 in our behavioral questions. First, the measures reflect only a component of health. Second, our latent variable model substantially reduces the number of parameters we need to estimate.

Substituting equation (12) and equation (7) into equation (8) shows that the value functions at time t can be rewritten as

(15)

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

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. In an attempt to account for this fact, 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

(16) $I^* = \Pi_1 X_0 + \Pi_2 Z_0 + \varepsilon^I$.

We assume that $\varepsilon^I \sim N(0,1)$ in which case equation (16) 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 specific health conditions 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 (10) implies that for t>1, $COV(\varepsilon^I, \nu_t)$ is a function of $COV(\varepsilon^I, \nu_1)$ and ρ, and we estimate $COV(\varepsilon^I,\nu_1)$. 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 effects 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.

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 (14), (15), and (16) that define our model and contain the parameters to be estimated.

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 (14) must lie. While individuals make decisions yearly 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 we discuss the timing of survey dates in more detail. For concreteness at this point, assume that the person's work status is reported three times, 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 (14) implies the three conditions that must be satisfied:

(17) $h_2* = \pi X_2 + \gamma Z_2 + [\nu_2 + \mu_2] < 0 \Rightarrow \nu_2 + \mu_2 < -\pi X_2 - \gamma Z_2$, $h_3* = \pi X_3 + \gamma Z_3 + [\nu_3 + \mu_3] < 0 \Rightarrow \nu_3 + \mu_3 < -\pi X_3 - \gamma Z_3$,

$h_5* = \pi X_5 + \gamma Z_5 + [\nu_5 + \mu_5] > 0 \Rightarrow \nu_5 + \mu_5 > -\pi X_5 - \gamma Z_5$

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 (16) implies that

(18) $\varepsilon^I > -\Pi_1 X_0 - \Pi_2 Z_0$.

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 at 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.[25]

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,

(19) $V_C(1,S(1)) > V_B(1,S(1))$, $V_C(1,S(1)) > V_A(1,S(1))$, and $V_C(1,S(1)) > V_N(1,S(1))$.

Define $V^*_j(t,S(t),\psi_t,\eta_t) = V_j(t,S(t),\psi_t,\eta_t) - \kappa \cdot 1(j=C \text{ or } B) - \lambda^j_\eta \nu_t - \varepsilon^j_t$ 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 (19) can be rewritten as

(20) $\kappa + \lambda_\eta^C \nu_1 + \varepsilon_1^C - \kappa \cdot 1(j=B) - \lambda_\eta^j \nu_1 - \varepsilon_1^j > V^*_j(1,S(1),\psi_1,\eta_1) - V^*_C(1,S(1),\psi_1,\eta_1)$, 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.

[25]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..

At t=4, the person chooses the bridge option and the three conditions analogous to those in equation (20) indicate that the value of the the bridge option at t=4 is greater than the value of the three other options at t=4:

(21) $\kappa+\ \lambda_\eta^B\nu_4+\varepsilon_4^B\ -\kappa\cdot 1(j{=}C)-\ \lambda_\eta^j\nu_4-\varepsilon_4^j > V^*_j(4,S(4),\psi_4,\eta_4)-V^*_B(4,S(4),\psi_4,\eta_4)$ 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 (20) 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:

(22) $\lambda_\eta^N\nu_5+\varepsilon_5^N\ -\kappa\cdot 1(j{=}B)-\ \lambda_\eta^j\nu_5-\varepsilon_5^j > V^*_j(5,S(5),\psi_5,\eta_5)-V^*_N(5,S(5),\psi_5,\eta_5)$ 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 (17), (18), (20), (21), and (22) are jointly true. The ψ_t's are independent of all other stochastic elements that enter these equations so that that likelihood contribution of person i can be written as

(23) $$L_i= \int PR(\nu_1+\mu_1<-\pi X_1-\gamma Z_1 \cap \nu_2+\mu_2<-\pi X_2-\gamma Z_2 \cap \nu_3+\mu_3>-\pi X_3-\gamma Z_3 \cap \varepsilon^I>-\Pi_1 X_0-\Pi_2 Z_0 \cap$$
$$\kappa+\lambda_\eta^C\nu_1+\varepsilon_1^C - \kappa - \lambda_\eta^B\nu_1-\varepsilon_1^B > V^*_B(1,S(1),\psi_1,\eta_1)-V^*_C(1,S(1),\psi_1,\eta_1) \cap \ldots \cap$$
$$\lambda_\eta^N\nu_5+\varepsilon_5^N - \lambda_\eta^A\nu_5-\varepsilon_5^A > V^*_A(5,S(5),\psi_5,\eta_5)-V^*_N(5,S(5),\psi_5,\eta_5) \quad dF(\psi_1, \psi_2, \psi_3\, \psi_4, \psi_5).$$

The integral in equation (23) can be simulated in a straightforward manner given the empirical distribution of $\psi_1,\ldots,\psi_5$. What remains is to show how to compute the probability expression in (23) given values of the ψ's. The joint probability is the area under the joint density of $\Psi=\{\nu_1+\mu_1, \nu_2+\mu_2, \nu_3+\mu_3, \varepsilon^I, \kappa+\lambda_\eta^C\nu_1+\varepsilon_1^C - \kappa - \lambda_\eta^B\nu_1-\varepsilon_1^B,$ $\kappa+\lambda_\eta^C\nu_1+\varepsilon_1^C - \lambda_\eta^A\nu_1-\varepsilon_1^A, \ldots, \lambda_\eta^N\nu_5+\varepsilon_5^N - \kappa - \lambda_\eta^B\nu_5-\varepsilon_5^B, \lambda_\eta^N\nu_5+\varepsilon_5^N - \lambda_\eta^A\nu_5-\varepsilon_5^A\}$ where the conditions that appear in the probability expression (i.e., the conditions in equations 17, 18, 20, 21, and 22) 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 $var(\nu_t+\mu_t) = var(\varepsilon^I) = var(\varepsilon_t^C) = var(\varepsilon_t^B) = var(\varepsilon_t^N) = var(\varepsilon_t^A) = 1$, by σ_κ and $var(\nu_t)$, and by the coefficients $\lambda_\eta^C, \lambda_\eta^B, \lambda_\eta^N, \lambda_\eta^A$. The non-diagonal elements of the covariance matrix are determined by the covariance terms $COV(\varepsilon^I, \nu_1)$, $COV(\varepsilon^I, \kappa)$, $\{COV(\mu_t, \varepsilon_t^j), j=C,B,A\}$, by the normalization of $COV(\mu_t, \varepsilon_t^N)$ 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 (23). Roughly speaking, this approach centers around the rewriting of the joint probability in equation (23) 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. This procedure is described in detail in Appendix C. We note here that the presence of the serially correlated unobservable η_t implies that value function approximation must take place during the GHK simulation of equation (23) because an implication of the backwards recursion modification discussed in Section 2 is that the V*'s will be solved for only a set of η_t grid points, and this set will not include all values of η_t that arise as part of the GHK simulation of equation (23). We use a non-parametric value function approach of the type described in Section 2.3.3.

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 Stern and ??). 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.

Section 4. Results

4.1 Data

Data for this research come from the Health and Retirement Survey, a multi-purpose social science survey conducted by the University of Michigan and funded by the National Institute on Aging. 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 estimation of our model in equations (24)-(26) 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. The HRS is described in additional detail in Juster and Suzman (1995).

The HRS covers a representative national sample of non-institutionalized men and women born between 1931 and 1941 (inclusive), so that respondents in the sample frame were aged 50-62 at the time of the first wave. The HRS 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; the second wave re-interviewed 11,642 respondents, representing 92% of the original sample.

The HRS includes the spouses/partners of the survey population even if they are not themselves in the age range of the sample frame; since respondents out of the sample frame do not constitute a representative sample, they are excluded here. The age-eligible first wave sample consists of 9,824 respondents, of whom 4,522 are men, of whom 733 are not married/partnered. Table 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.”

4.2 Descriptive Statistics

We are interested in understanding how the availability of economic resources and health affect economic behavior. Table 2 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 for those reporting work limitations are quite different from those not reporting them. 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, the Supplemental Security Income (SSI) and the Social Security Disability Insurance (DI) programs.[26] Crudely put, Table 2 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 3 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 4 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

[26]In Table 2, DI payments and regular SS payments are grouped under the category Social Security. Although not shown explicitly in Table 2, approximately 44% of those not working and reporting work limitations are receiving income from the DI program.

second column of Table 4 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 4 report descriptive statistics on our behavioral sample broken down by whether they chose C, B, A, or N in the final survey.[27] 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 that 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.

The average person in our behavioral sample would receive \$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 wages and bridge wages at the final survey based on estimates of equation (5) and equation (6) are \$32,854 and \$13,098 respectively. The final row of Table 4 show that the proportion of individuals in the sample who choose the options C, B, N, and A is .430, .204, .279, and .087 respectively.

[27]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.”

4.3 Model Estimates and Simulations

The parameters that enter our model are:

1). the parameters of the DI approval equation (11).

2). the parameters in the wage equations (5a) and (5b);

3). the parameters of the health expenditure equation (4);

4). the parameters π, γ, and σ^2_v from the health equation (12) and the parameters ρ and σ^2_ξ from equation (10);

5). τ, λ_{HI}, and $\{\lambda^j_X, \lambda^j_\eta$:j=C,B,A$\}$ from the current period utility equation (7);

6). the parameters Π_1 and Π_2 from the initial conditions equation (16);

7). the standard deviation of unobserved heterogeneity σ_κ;

8). The covariance parameters $\{COV(\mu_t, \varepsilon_t^j)$, j=C,B,A$\}$ discussed after equation (14),

9). the covariance parameters $\{COV(\varepsilon^I, v_1)$ and $COV(\varepsilon^I, \kappa)$, j=C,B,A$\}$ discussed after equation (16).

10). the parameter β^{Common} from the discount factor equation (9)

11).the parameter θ from the pecuniary utility function in equation (3).

We have estimate the parameters in 1), 2), and 3) outside of our behavioral model and report results in Appendix A. Ideally, it would have made sense to estimate the DI approval and wage equations jointly with the other parameters in our model. However, doing so would not have been easy.

The identification of the DI approval equation 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, A_2^{DI}=1, and Var(e^{DI}). This approach is motivated by the work of Benitez-Silva, Buchinsky and Rust (2004) who find evidence using HRS data that DI award decisions are about as equally reliable indicators of disability status as are the global self-reported measures available in the HRS.

The parameters of the wage equation are also difficult to estimate inside the structural model. One reason for this is related to the fact 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, must be calculated and stored prior to estimation. As a result, we estimate the wage equations (4a) and (4b) outside the structural model.

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. This has substantial benefits because it allows the use of a derivative-based Newton-Raphson parameter updating algorithm which require much 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 a a 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.[28] 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

[28]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.

which provided us with exclusive access to 48 Compaq Alpha CPU's for an extended period of time.[29] As a final concession to the size of the computing task we decided to reduce the number of parameters in the model by estimating the parameters the health equation (4 above) outside of the behavioral model. The estimates are shown in the first column of Table 5C.[30] 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.

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.[31] We begin by estimating a "baseline" specification in which $\beta^{Common}=0.90$ and $\theta=1$ which implies that $Y^j(t)^{1-\theta}/(1-\theta)=\ln(Y)$. We then examine the robustness of our results to changes in β^{Common} and θ. The estimates from the baseline and robustness specifications are shown in Table 5D and discussed in the next two subsections.

Baseline Specification

The first column of Table 5D shows the estimates of the behavioral equations for the baseline specification.[32] 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 5D 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

[29]See Swann (2002) for a user-friendly description of how to use parallel processing techniques with Maximum Likelihood estimation algorithms.

[30]These parameters can be estimated by using the portion of the likelihood function in Section 3.3 which corresponds to the health equations. Estimation of these parameters in this way is not time consuming because it does not require value functions to be computed.

[31]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 $COV(\mu_t, \varepsilon_t^A)=COV(\mu_t, \varepsilon_t^N)=0$.

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

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) than individuals in better health. Our estimates imply that, over and above the effect of health insurance on disposable income, health insurance has a positive effect on well being, but the estimated effect is quite small and not statistically significant. In terms of the variance/covariance terms, 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 role that economic resources and health play in determining labor decisions we begin by performing simulations using a "representative person." We construct a representative 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.[33] We first assume that the representative 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 five rows of Table 6A show simulated choice probabilities at t=1 assuming that the representative person is 55, 60, 62, 64, and 65 years of age at t=1 respectively where to simplify the discussion and make tables more readable we have combined the Career and Bridge options into a single "working" category (C+B). Since the simulated individuals would have been employed at t=0, these simulated probabilities can be thought of as one year transition rates.

The simulated choice probabilities at age 55 and age 60 are quite similar. At these ages, the representative 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 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

[33]Our representative 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.

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.[34] 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.

Table 6B shows choice probabilities at different ages for the representative person under the assumption that his health is one standard deviation below average at t=1.[35] 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 average health person 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 transitioning out of the workforce (N+A) at age 60 is 0.025 for the average health person 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 the average and below-average health individuals at age 62 indicates that the effect of economic resources depends to some extent on health. At age 62 when SS benefits become available, the total probability of transitioning out of 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.

---

[34]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.

[35]As is clear from table 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.

Table 6C shows choice probabilities at different ages for the representative person under the assumption that his health is 1.5 standard deviations below average. Comparing the results of Table 6C to those of 6B indicates that, once a person reaches poor health, an incremental worsening of health can have large effects on the decisions of individuals. When compared to the representative person in Table 6B (who has health 1 standard deviation below average), the representative person in Table 6C is at least twice as likely to apply for Disability Insurance at each age and is approximately .10 more likely to transition out of the workforce (N+A) at each age. Comparing Table 6C to Table 6A reinforces the notion that the availability of economic resources may influence individuals in poor health more than individuals in good health. At age 62 when SS benefits become available, the total probability of transitioning out of the workforce at t=1 increases by approximately 0.07 (from 0.198 to 0.269) for the person in below-average health but, as discussed before, increases by only approximately 0.02 (from 0.025 to 0.049) for the person in average health.

Robustness Checks -Changes to the Baseline Specification

We examine the robustness of the baseline results to changes in the values used for θ and β^{Common}. Estimates of these specifications are shown in Columns 2-3 of Table 5D.

Column 2 of Table 5D shows parameter estimates a first robustness check which involves increasing the level of risk aversion in the model by changing θ in equation (3) from $\theta = 1$ to $\theta = 2$.[36] The health estimates and virtually all other estimates are very similar to the baseline specification. As a result, it is not surprising that the simulations in Table 7 reveal that the choice probabilities for the representative person under this specification are very similar to the choice probabilities under the baseline specification. As a result, changing the degree of risk aversion in the model from $\theta=1$ to $\theta=2$ has little effect on conclusions about the importance of health, the importance of financial resources, and the importance of the interaction between health and financial resources.

[36]The exact specification used is $-3000 \cdot \tau \cdot Y^{j}(t)^{-1}$. The change in the utility function implies that the value of τ in column 2 of Table 5D cannot be compared directly to the value of τ from the other specifications.

Column 3 of Table 5D shows parameter estimates for the second robustness check which involves decreasing β^{Common} from 0.90 to 0.67.[37] Perhaps the most noticeable changes from this specification are the increase in the point estimate (standard error) of τ from 0.335 (0.059) in the baseline specification to 0.586 (.111) in this specification and an increase in the point estimate (standard error) of σ_κ (the standard deviation of unobserved heterogeneity) from 1.423 (.322) in the baseline specification to 1.945 (.426) in this specification. Because it is difficult to understand the qualitative differences between the models simply by examining particular point estimates, we compare simulated choices for the representative person in this robustness specification to those from the baseline specification.

The representative person simulations in Table 8 for this specification reveal that health has a somewhat smaller impact on choices than in the baseline simulation. For example, in this specification the probability of working at age 60 is .994, .962, and .917 for a person with average health, health 1 standard deviation below average, and 1.5 standard deviations below average respectively. In the baseline specification, the decrease with health is somewhat more dramatic with the three probabilities being .974, .893, and .802 respectively. The counterintuitive result that health plays a somewhat smaller role despite larger point estimates arises in part because a person's unobserved heterogeneity type plays a bigger role in this specification. The effect of consumption as revealed by the receipt of Social Security income at age 62 is similar in this specification as it was in the baseline specification.

Thus, changing θ has little impact on estimates or simulations. Our results indicate that changing the discount factor can influence estimates and simulations. However, the size of the change in the discount factor was dramatic and, even in this case, the general message from the simulations is consistent across both the baseline and robustness specification. Specifically, health and economic resources play very important roles in determining labor supply decisions of older workers and a full understanding of the effects of these variables requires that one take into account the manner in which they interact.

[37] 0.67 represents substantial discounting of the future, but is consistent with work that indicates that individuals heavily discount the future relative to the present (Angeletos, Laibson, Repetto, Tobacman, and Weinberg, 2001).

Potential Changes in Policy

The simulations involving the representative 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 four potential changes in policy. First, we examine the effect of removing the option of early SS benefits. Second, we examine the effects of a policy that has been implemented - changing the normal retirement age from 65 to 67. Third, we examine the impact of removing the Disability Insurance program entirely. Finally, we examine the impact of implementing a type of universal health insurance program.

To quantify the effects of these policy changes, we first perform a baseline simulation in which no policy change has occurred. This simulation yields the probability that each person in our behavioral sample will choose each of the options {C, B, N, A} in each of the years that a choice is observed for him in the data.[38] Aggregate measures for the sample, which are created by averaging these individual-specific choice probabilities across all sample members, are shown in the first column of Table 9a. The first entries in Column 1 represent the choice probabilities that are generated if individual choices at all ages are pooled. [39] Thus, under the baseline specification, our model indicates that individuals will choose the work option (C+B) in .859 of the periods for which decisions are observed in the data (hereafter referred to as the pooled decision periods), will choose the non-work option (N) in .125 of the pooled decision periods, and will choose to apply for Disability Insurance (A) in .016 of the pooled decision periods.[40] The remainder of the entries in Column 1 reflect the choice probabilities when choices are disaggregated by age (for select ages). Thus, for example, the simulations indicate that the probability of working (C+B) at age 55 is .863.[41]

---

[38]Thus, we take the approach of simulating choices within the sample period.

[39]Thus, for these entries we are averaging choice probabilities of 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 to the aggregate measures at a subset of these ages.

[40]In the data, the empirical probabilities are .838, .153, and .018 respectively.

[41]Given our within-sample period simulation strategy, this number is computed using all people for which a choice is observed in the data at age 55.

For each policy change, the simulation process is repeated after modifying the model appropriately to reflect the change. Column 2 of Table 9a shows the choice probabilities associated with a first policy change 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 probability of choosing the non-work option (N) is .115 under the baseline simulation and .113 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 probability of choosing the non-work option falls by approximately 18% (from .147 under the baseline simulation to .121 under the policy change). A similar effect (.172 versus .150) 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 probability of choosing N is accompanied by an increase of similar size in the probability of choosing C or B. At 65 when Social Security benefits become the same in the baseline simulation and policy change simulation, the number of people in the option N becomes more similar under the baseline and policy change (.192 vs. .184).[42]

We also construct the analog to Table 9a 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 individuals in our sample with health that is better than one standard deviation below the average for the sample (hereafter rerred to as "good" health) and show these probabilities in Table 9b and Table 9c respectively.[43] Consistent with our representative person simulations, we find 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 9C reveals that, for individuals in our sample with bad health, the policy causes the probability of choosing N at age 64 to fall by approximately .06

[42]The small difference remains because, under the policy change, individuals are less likely to have left their career jobs as of age 65, and, therefore, have an additional and often desirable work option at age 65.

[43]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.

(from .263 under the baseline to .204 under policy 1) and the probability of applying for DI to fall by approximately .03 (from .083 to .054). A comparison of the first and second columns of Table 9B reveals that, for individuals in our sample with bad health, the policy causes the probability of choosing N at age 64 to fall by approximately .02 (from .158 under the baseline to .138 under policy 1) and the probability of choosing A to fall by .002 (from .013 to .011). Thus, this policy has a differential impact by health status by changing the work status of people in bad health more than the work status of people in good health; the probability of C+B at age 64 increases by .088 for individuals in bad health but by only .021 for individuals in good health).

Column 3 of Table 9 (Policy 2) shows the choice probabilities 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. Under this policy, while individuals still become eligible for benefits at the age of 62, the amount of the benefits that is received if one retires at age 62 or at any other age is reduced.[44] This change in the normal retirement age is currently being phased in. A comparison of Column 1 and Column 3 of Table 9a reveals that this policy will have relatively small effects on individual decisions. The change has virtually no effect before age 62. At ages 62 or higher, the policy causes a decrease in the probability of the non-work option (N) of one percentage point or approximately seven percent for people in our sample.

Column 4 of Table 9a (Policy 3) shows the choice probabilities associated with a third policy change in which the Disability Insurance program is removed entirely. Comparing the entries associated with C+B in Column 4 to the entries associated with C+B in Column 1 reveals that the removal of the DI program leads to very little change in the proportion of people who are working at any particular age. Instead, the change leads primarily to an increase in the proportion of people who choose to leave the workforce altogether (N). Intuitively this occurs because individuals are typically in bad health when they apply for DI benefits, and, as a

[44]Under this policy, a person who retires at the age of 67 receives the benefits he would have received under baseline if he would have retired at age 65. A similar reduction in benefits takes place at all other ages.

result, tend to find the non-work option (N) relatively appealing when the DI option is not available. More formal evidence about this matter and about the differential effect of the policy change by health status can be seen in Tables 9b and 9c. The "Pooled Ages" entry in Column 1 of Table 9b shows that, under the baseline simulation, individuals in good health choose to apply for DI benefits (A) in only .012 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 9C shows that, under the baseline simulation, individuals in bad health choose to apply for DI (A) benefits in .057 of the pooled decision years. For these individuals, comparing the "Pooled Ages" entries under the baseline (Column 1 of Table 9c) to that under policy 3 (Column 4 of Table 9C) shows that the removal of the DI option results in an increase in the probability of N from .193 to .255 but very little change in the probability of C+B (from .750 to .745).[45] Thus, the differential impact of this policy change by health status comes not from a change in work status, but rather, from a reduction in the benefits associated with being out of the workforce that arises because the person does not receive income from DI benefits.

Column 5 of Table 9 shows the choice probabilities associated with a final policy change in which individuals are given access to medicare at all ages. We find that this change has very small effects both overall and for different health groups. Given the small effect that health insurance seems to have on out of pocket medical expenses and give the small direct effect it has on utility (at least for this group), this result is not at all surprising.

---

[45]As in Table 9a for the entire sample, removing the DI option for the bad health individuals leads to a seemingly counterintuitive decrease in the probability associated with the working options (C+B). The reason for this is simply that the policy change has an effect on the number of pooled decision years for which choices are simulated. Consistent with what is seen in the data, the likelihood contribution for a person assumes that a person who applies for DI and is accepted makes no subsequent decisions (i.e., he remains on DI forever). Analogously, when constructing Table 9, subsequent choices are not simulated for a person who applies for and is accepted for DI benefits. Thus, removing the DI option leads to more simulated choices for the types of people who apply for and are awarded DI benefits. The decrease in the probability of C+B under the policy change occurs because these types of people tend to be in bad health,and find the option N relatively appealing. This can be formalized by recomputing the choice probabilities in Table 9 replacing the DI application category (A) with a new category (call it A*) which includes both people who are applying for DI in the present period ***and*** people who have been accepted in the past (and as a result are on DI in the current period but are not making decisions which enter the numbers in Table 9). In this case, for the entire sample the probability associated with C+B is .840 under the baseline and .854 under policy 3. The probability of A* is .038 under the baseline and is zero by definition under policy 3. For individuals in bad health, the probability associated with C+B is .712 under the baseline and .743 under policy 3. The probability of A* is .111 under the baseline and is zero by definition under policy 3. Thus, the probability of C+B increases as expected when the DI is removed. In addition, note that for individuals in bad health, the increase in the probability of C+B (.032) is relatively small compared to the number of people who are forced to leave DI under the policy (i.e., .111) so that the removal of the DI option leads primarily to an increase in the proportion of people that are in the option N.

V. 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. Our simulations indicate that both health and economic resources play an important role in determining labor supply decisions of older workers. 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, simulations based on our model estimates show that our representative 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 minimal effects on applications for the Disability Insurance Program.[46] 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 of the application for DI benefits we should see observe 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 apply for DI will not rise by much.

[46]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. Their estimates imply a larger effect on the application for DI benefits of reducing the incentives to retire early. The modeling approach we take is so different than the one taken by Mitchell and Phillips that it is hard to know what to attribute differences to.

References

Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg (2001): "The Hyperbolic Buffer Stock Model: Calibration, Simulation, and Empirical Evaluation," *Journal of Economic Perspectives*, 15(3), Summer, 47-68.

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

Benitez-Silva, Hugo, Moshe Buchinsky, Hiu Man Chan, Sofia Cheidvasser and John Rust, (2000): "How Large is the Bias is Self-Reported Disability?" NBER Working Paper # 7526.

Benitez-Silva, Hugo, Moshe Buchinsky, and John Rust, (2004): "How Large Are the Classification Errors in the Social Security Disability Award Process?" NBER Working Paper # 10219.

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

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

Bound, J. and R. Burkhauser (1999): "Economic Analysis of Transfer Programs Targeted on People with Disabilities," in *Handbook of Labor Economics, Vol 3.,* eds., D. Card and O. Ashenfelter, Amsterdam; New York and Oxford: Elsevier Science, North-Holland, pp 3417-3528.

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

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

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), Spring 2002, pages 231-50.

Blau, D. and D. Gilleskie (2001): "Health Insurance and Retirement of Married Couples," unpublished manuscript, Department of Economics, the University of North Caroline.

Brien, Michael, Steven Stern, and Lee Lillard (2001): "Cohabitation, Marriage, and Divorce in a Model of Match Quality," Unpublished Manuscript, December.

Burkhauser, R.V., D Dwyer, M. Lindeboom, J. Theeuwes, I. Woittiez (1999): "Health, Work, and Economic Well-Being of Older Workers, Aged Fifty-One to Sixty-One: A Cross-National Comparison Using the U.S. HRS and the Netherlands CERRA Data Sets," Smith,J. P and R. J. Willis, eds. *Wealth, work, and health: Innovations in measurement in the social sciences: Essays in honor of F. Thomas Juster*. Ann Arbor: University of Michigan Press, pages 233-66.

Charles, Kerwin K. (1999): "Sickness in the Family: Health Shocks and Spousal Labor Supply," Unpublished Manuscript, Gerald R. Ford School of Public Policy Working Paper # 00-011.

Citro, Constance F., and Robert T. Michael, eds. (1995): *Measuring Poverty: A New Approach.* Washington, D.C.: National Academy Press.

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

Daly, Mary C., and John Bound (1996):. "Worker Adaptation and Employer Accommodation Following the Onset of a Health Impairment." *Journal of Gerontology*, 51B: S53-S60.

Diermeier, Daniel, Michael P. Keane, and Antonio Merlo (2002): “The Career Decisions of Politicians,” Unpublished Manuscript.

French, Eric (2000): “The Effects Of Health, Wealth, And Wages On Labor Supply And Retirement Behavior,” Federal Reserve Bank of Chicago working paper # 2000-02.

French, Eric and John Bailey Jones (2001): “The Effects of Health Insurance and Self-Insurance on Retirement Behavior,” Federal Reserve Bank of Chicago working paper # 2000-19.

French, Eric and John Bailey Jones (2002): “On the Distribution and Dynamics of Health Costs” Federal Reserve Bank of Chicago working paper # 2000-21.

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.

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

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

Juster, F . T., and R. Suzman (1995): "An Overview of the Health and Retirement Study," *Journal of Human Resources*, 30(0), Suppl., S7-56.

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, Michael P., and Kenneth 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, Brent (1999): “Latent Work Disability and Reporting Bias,” *Journal of Human Resources*; 34, 734-69

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

Mitchell Olivia and John W.R. Phillips (2000): "Retirement Responses to Early Social Security Benefit Reductions," *NBER Working Paper 7963*, National Bureau of Economic Research, Inc.

Olin, Gary L. and Steven 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

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, pp. 25-39.

Quinn, J.F. (1998): "New Paths to Retirement," Boston College Working Paper 406.

Quinn, J.F., and M. Kozy (1995): "Bridge Jobs in the 1990s," Presented at the HRS Wave 2 Early Results Workshop, Ann Arbor, October.

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, John (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, John, Moshe Buchinsky and Hugo 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.

Stephens, Mel (forthcoming): "Worker Displacement and the Added Worker Effect," National Bureau of Economic Research Working Paper #8260, forthcoming in *Journal of Labor Economics*.

Stern, Steven (1989): "Measuring the Effect of Disability on Labor Force Participation," *The Journal of Human Resources*, Vol. 24, 361-395

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

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

van der Klaauw, Wilbert, and Kenneth I. Wolpin (2002): “Social Security, Pensions and the Savings and Retirement Behavior of Households,” Unpublished Manuscript, Department of Economics, University of North Carolina.

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, No. 2: 253-87.

Table 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

Table 2 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 3 Income Transitions

1992-1998 Choice	Bridge				Career				Apply for DI				Retire			
N	63				86				27				77			
	1991		1999		1991		1999		1991		1999		1991		1999	
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
Total Own		**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**

Table 4 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	0.222	0.261	0.243	0.229	0.188	0.267
Some College	0.170	0.183	0.135	0.114	0.208	0.333
College Grad	0.193	0.191	0.243	0.114	0.167	0.200
Work limited at wave 1	0.097	0.296	0.095	0.143	0.083	0.067
Work limited at wave 4	0.188	0.365	0.108	0.057	0.208	0.867
Age 62 or older at wave 4	0.369	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	0.466	0.078	0.419	0.486	0.542	0.333
Eligible for DC pension ever	0.369	0.122	0.405	0.371	0.375	0.267
Expected Career Earnings at last obs	32,854	6,048	33,143	32,904	34,700	23,761
Expected Bridge Earnings at last obs	13,098	2,294	14,647	12,253	10,658	13,793
With no employer health ins	0.176		0.095	0.314	0.125	0.467
With only current coverage	0.290		0.297	0.286	0.292	0.200
With both current & retiree covg	0.534		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	0.605	0.395	0.430	0.204	0.279	0.087

Table 5A- Wage 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	

*For description of models and data used see Appendix A.

Table 5B - Health expenditure equation *

	Common Truncation		Separate Truncation	
Variable	**Coefficient**	**Standard Error**	**Coefficient**	**Standard 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	

*For description of methods and data used for these equations, see Appendix A.

Table 5C - Estimates of Health Equation and Initial Conditions Equation for Baseline Specification

	Health Equation (22)	**Initial Conditions Equation (26)**
X from equations (??) and (??)		
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 from equations 22 and 26 (Activities of Daily Living at time health is observed for first column and t=0 for second column)		
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)	
vartrue	.82 (.080)	

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

	Baseline	Baseline with $\theta = 2$	Baseline with $\beta^{Common} = .67$
	Estimate (Std. Error)	**Estimate (Std. Error)**	**Estimate (Std. Error)**
Pecuniary Utility			
τ	.336 (.060)	.273 (.080)	.586 (.111)
$U_{np}^{C}(t)$ **Non-Pecuniary Utility Career** (C)			
Constant	.546 (.338)	.694 (.324)	.904 (.505)
λ_{η}^{C} (Health, η)	-.776 (.397)	-.729 (.406)	-1.028 (.525)
λ_{HI} (Has Health Insurance)	.063 (.043)	.126 (.137)	-.038 (.207)
$U_{np}^{B}(t)$ **Non-Pecuniary Utility Bridge(B)**			
Constant	-.908(.486)	-.819 (.460)	-.826 (.600)
λ_{η}^{B} (Health, η)	-1.357(.351)	-1.342 (.367)	-1.544 (.506)
λ_{HI} (Has Health Insurance)	.063 (.043)	.126 (.137)	-.038 (.207)
$U_{np}^{A}(t)$ **Non-Pecuniary Utility DI (A)**			
Constant	-2.102(.399)	-1.849 (.382)	-2.692 (.418)
λ_{η}^{A} (Health, η)	.851 (.212)	.948 (232)	.818 (.174)
Has Health Insurance	.063 (.043)	.126 (.137)	-.038 (.207)
$U_{np}^{N}(t)$ **Non-Pecuniary Utility Non-Work (N)**			
Constant	Normalized to zero	Normalized to zero	Normalized to zero
λ_{η}^{N} (Health, η)	Normalized to zero	Normalized to zero	Normalized to zero
λ_{HI} (Has Health Insurance)	.063 (.043)	.126 (.137)	-.038 (.207)
Health Equation	See Table 5C	Not Shown	Not Shown
Initial Conditions Equation	See Table 5C	Not Shown	Not Shown
Covariance Terms			
σ_{κ}	1.423 (.322)	1.360 (.339)	1.945 (.426)
$COV(\varepsilon^{I}, \nu_{t})$	-.309 (.106)	-.3044 (.110)	-.318 (.103)
$COV(\mu_{t}, \varepsilon_{t}^{C})$	-.191 (.173)	-.185 (.178)	-.277 (.165)
$COV(\mu_{t}, \varepsilon_{t}^{B})$	.216 (.128)	.193 (.161)	.147 (.155)
$COV(\varepsilon^{I}, \kappa)$	-.784 (.412)	-.645 (.410)	–2.245 (.608)
Log Likelihood Function Value	-940.033	-943.936	-945.423

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 6A Choice Probabilities Representative Person at Different Ages - Average Health Baseline Specification

	Working (C+B)	Labor Force Exit (N+A)	Non-Work (N)	Apply DI (A)
AGE=55	0.981	0.019	0.018	0.001
AGE=60	0.974	0.025	0.024	0.001
AGE=62	0.951	0.049	0.046	0.003
AGE=64	0.947	0.053	0.050	.003
AGE=65	0.947	0.053	0.0533	NA*

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

Table 6B Choice Probabilities Representative Person at Different Ages - Health 1 Std. Dev. Below Average Baseline Specification

	Working (C+B)	Labor Force Exit (N+A)	Non-Work (N)	Apply DI (A)
AGE=55	0.902	0.098	0.057	0.041
AGE=60	0.893	0.106	0.082	0.024
AGE=62	0.831	0.168	0.138	0.030
AGE=64	0.828	0.172	0.146	0.026
AGE=65	0.839	0.161	0.161	NA*

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

Table 6C Choice Probabilities Representative Person at Different Ages - Health 1.5 Std. Dev. Below Average Baseline Specification

	Working (C+B)	Labor Force Exit (N+A)	Non-Work (N)	Apply DI (A)
AGE=55	0.805	0.195	0.084	0.111
AGE=60	0.802	0.198	0.127	0.071
AGE=62	0.731	0.269	0.209	0.060
AGE=64	0.721	0.279	0.207	0.072
AGE=65	0.745	0.254	0.254	NA*

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

Table 7A Choice Probabilities Representative Person at Different Ages - Average Health
Baseline Specification with θ=2

	Working (C+B)	Labor Force Exit (N+A)	Non-Work (N)	Apply DI (A)
AGE=55	0.972	0.028	0.027	0.01
AGE=60	0.966	0.034	0.032	0.002
AGE=62	0.937	0.063	0.059	0.004
AGE=64	0.941	0.059	0.055	.004
AGE=65	0.941	0.059	0.059	NA*

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

Table 7B Choice Probabilities Representative Person at Different Ages - Health 1 Std. Dev. Below Average
Baseline Specification with θ=2

	Working (C+B)	Labor Force Exit (N+A)	Non-Work (N)	Apply DI (A)
AGE=55	0.879	0.121	0.080	0.041
AGE=60	0.867	0.133	0.109	0.023
AGE=62	0.808	0.192	0.164	0.028
AGE=64	0.808	0.192	0.156	0.035
AGE=65	0.823	0.177	0.177	NA*

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

Table 7C Choice Probabilities Representative Person at Different Ages - Health 1.5 Std. Dev. Below Average
Baseline Specification with θ=2

	Working (C+B)	Labor Force Exit (N+A)	Non-Work (N)	Apply DI (A)
AGE=55	0.778	0.222	0.117	0.105
AGE=60	0.774	0.226	0.154	0.072
AGE=62	0.706	0.294	0.239	0.055
AGE=64	0.699	0.301	0.216	0.085
AGE=65	0.728	0.272	0.272	NA*

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

Table 8A Choice Probabilities Representative Person at Different Ages - Average Health
Baseline Specification with β^{Common}=.67

	Working (C+B)	Labor Force Exit (N+A)	Non-Work (N)	Apply DI (A)
AGE=55	0.997	0.003	0.003	0.000
AGE=60	0.994	0.007	0.006	0.000
AGE=62	0.977	0.023	0.023	0.000
AGE=64	0.976	0.024	0.024	.000
AGE=65	0.971	0.029	0.029	NA*

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

Table 8B Choice Probabilities Representative Person at Different Ages - Health 1 Std. Dev. Below Average
Baseline Specification with β^{Common}=.67

	Working (C+B)	Labor Force Exit (N+A)	Non-Work (N)	Apply DI (A)
AGE=55	0.962	0.038	0.022	0.015
AGE=60	0.961	0.039	0.028	0.011
AGE=62	0.903	0.097	0.085	0.012
AGE=64	0.898	0.102	0.090	0.012
AGE=65	0.902	0.097	0.097	NA*

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

Table 8C Choice Probabilities Representative Person at Different Ages - Health 1.5 Std. Dev. Below Average
Baseline Specification with β^{Common}=.67

	Working (C+B)	Labor Force Exit (N+A)	Non-Work (N)	Apply DI (A)
AGE=55	0.925	0.075	0.036	0.039
AGE=60	0.917	0.083	0.050	0.028
AGE=62	0.817	0.183	0.147	0.036
AGE=64	0.811	0.189	0.152	0.037
AGE=65	0.823	0.177	0.177	NA*

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

Table 9a Policy Simulations - All Health Combined

	Baseline	**Policy 1**	**Policy 2**	**Policy 3**	**Policy 4**
Pooled Ages	C+B=.859 N+A=.141 N=.125 A=.016	C+B=.866 N+A=.134 N=.117 A=.017	C+B=.862 N+A=.138 N=.121 A=.017	C+B=.854 N+A=.146 N=.146 A=.000**	C+B=.860 N+A=.140 N=.125 A=.015
AGE=55	C+B=.863 N+A=.137 N=.115 A=.022	C+B=.865 N+A=.135 N=.113 A=.023	C+B=.864 N+A=.136 N=.114 A=.022	C+B=.862 N+A=.138 N=.138 A=0.000**	C+B=.863 N+A=.133 N=.112 A=.021
AGE=60	C+B=.858 N+A=.141 N=.131 A=.010	C+B=.862 N+A=.138 N=.125 A=.013	C+B=.860 N+A=.140 N=.128 A=.011	C+B=.852 N+A=.148 N=.148 A=.000**	C+B=.857 N+A=.143 N=.133 A=.010
AGE=62	C+B=.840 N+A=.160 N=.147 A=.013	C+B=.866 N+A=.134 N=.121 A=.013	C+B=.849 N+A=.151 N=.137 A=.015	C+B=.837 N+A=.163 N=.163 A=.000**	C+B=.839 N+A=.161 N=.148 A=.013
AGE=64	C+B=.806 N+A=.195 N=.172 A=.023	C+B=.831 N+A=.169 N=.150 A=.019	C+B=.815 N+A=.185 N=.160 A=.024	C+B=.795 N+A=.205 N=.205 A=.000**	C+B=.804 N+A=.196 N=.173 A=.023
AGE=65	C+B=.808 N+A=.192 N=.192 A=.000*	C+B=.816 N+A=.184 N=.184 A=.000*	C+B=.820 N+A=.180 N=.180 A=.000*	C+B=.782 N+A=.218 N=.218 A=.000*	C+B=.807 N+A=.193 N=.193 A=.000*

* person cannot apply for Disability Insurance at age 65 or older.
** 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 giving individuals medicare in all periods, regardless of age.

Table 9b Policy Simulations - Good Health (no worse than 1 standard deviation below average)

	Baseline	Policy 1	Policy 2	Policy 3	Policy 4
Pooled Ages	C+B=.871 N+A=.129 N=.117 A=.012	C+B=.878 N+A=.122 N=.110 A=.012	C+B=.875 N+A=.125 N=.113 A=.012	C+B=.868 N+A=.132 N=.132 A=.000**	C+B=.872 N+A=.128 N=.117 A=.011
AGE=55	C+B=.882 N+A=.118 N=.104 A=.014	C+B=.883 N+A=.117 N=.102 A=.015	C+B=.883 N+A=.117 N=.103 A=.014	C+B=.882 N+A=.118 N=.118 A=0.000**	C+B=.882 N+A=.208 N=.105 A=.013
AGE=60	C+B=.865 N+A=.134 N=.125 A=.009	C+B=.877 N+A=.123 N=.115 A=.009	C+B=.866 N+A=.134 N=.124 A=.010	C+B=.859 N+A=.141 N=.141 A=.000**	C+B=.864 N+A=.136 N=.127 A=.009
AGE=62	C+B=.848 N+A=.152 N=.141 A=.011	C+B=.874 N+A=.126 N=.115 A=.011	C+B=.858 N+A=.142 N=.130 A=.012	C+B=.849 N+A=.151 N=.151 A=.000**	C+B=.849 N+A=.151 N=.140 A=.011
AGE=64	C+B=.829 N+A=.171 N=.158 A=.013	C+B=.850 N+A=.150 N=.138 A=.011	C+B=.835 N+A=.165 N=.152 A=.014	C+B=.818 N+A=.182 N=.182 A=.000**	C+B=.824 N+A=.176 N=.163 A=.013
AGE=65	C+B=.820 N+A=.180 N=.180 A=.000*	C+B=.821 N+A=.179 N=.179 A=.000*	C+B=.827 N+A=.173 N=.173 A=.000*	C+B=.718 N+A=.283 N=.283 A=.000*	C+B=.819 N+A=.181 N=.181 A=.000*

* person cannot apply for Disability Insurance at age 65 or older.
** 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 giving individuals medicare in all periods, regardless of age.

Table 9c Policy Simulations - Bad Health (1 standard deviation or more below average)

	Baseline	Policy 1	Policy 2	Policy 3	Policy 4
Pooled Ages	C+B=.750 N+A=.25 N=.193 A=.057	C+B=.780 N+A=.220 N=.170 A=.050	C+B=.762 N+A=.238 N=.182 A=.056	C+B=.745 N+A=.255 N=.255 A=.000**	C+B=.750 N+A=.250 N=.198 A=.052
AGE=55	C+B=.668 N+A=.332 N=.228 A=.104	C+B=.670 N+A=.330 N=.223 A=.107	C+B=.689 N+A=.311 N=.216 A=.095	C+B=.688 N+A=.309 N=.309 A=0.000**	C+B=.667 N+A=.333 N=.236 A=.097
AGE=60	C+B=.789 N+A=.211 N=.189 A=.022	C+B=.767 N+A=.233 N=.196 A=.037	C+B=.822 N+A=.178 N=.159 A=.019	C+B=.800 N+A=.200 N=.200 A=.000**	C+B=.785 N+A=.215 N=.194 A=.021
AGE=62	C+B=.791 N+A=.208 N=.180 A=.028	C+B=.816 N+A=.184 N=.154 A=.031	C+B=.759 N+A=.241 N=.199 A=.042	C+B=.763 N+A=.237 N=.237 A=.000**	C+B=.746 N+A=.253 N=.217 A=.036
AGE=64	C+B=.654 N+A=.346 N=.263 A=.083	C+B=.742 N+A=.258 N=.204 A=.054	C+B=.634 N+A=.366 N=.242 A=.124	C+B=.652 N+A=.348 N=.348 A=.000**	C+B=.616 N+A=.384 N=.267 A=.117
AGE=65	C+B=.660 N+A=.340 N=.340 A=.000*	C+B=.791 N+A=.201 N=.201 A=.000*	C+B=.771 N+A=.229 N=.229 A=.000*	C+B=.656 N+A=.344 N=.344 A=.000*	C+B=.655 N+A=.345 N=.345 A=.000*

* person cannot apply for Disability Insurance at age 65 or older.
** 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 giving individuals medicare in all periods, regardless of age.

Appendix A: **Data Elements**

Definition of the choice set

Responses to the wave 2 survey define the observed choices used in the empirical model. An individual in the behavioral sample is classified as having applied for disability benefits (n=23) if he reported applying for Social Security Disability or Supplemental Security Income sometime in the year leading up to the wave 2 interview, regardless of his work status as of wave 2. Among those not classified as disability applicants, those who report that they are doing no work for pay as of wave 2 (n=62) are classified as "non-workers." Note that this includes a few respondents (n=4) who had applied for disability benefits but did so more than a year before the wave 2 interview. Among non-applicants who reported working, they were asked if they were working for the same employer, named by the interviewer, they reported in the wave 1 interview. If they responded negatively (including 38 who report working for themselves in wave 2 but reported working for someone else at wave 1), and indicated that they had worked in the new job for less than one year they were classified as having a bridge job (n=149). Persons who had worked for a different employer (or themselves) for more than a year are not included in the behavioral sample because we have no detailed information on that job until the wave 2 interview in conducted. The remaining workers (n=813) were those who replied that their employer was the same as in wave 1 and are classified as working in their "career" job. Finally, seven respondents who either denied working for the employer given in wave 1 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.

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

$$S(t) = \exp[-\exp(X_t^{\prime}\beta)],$$

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).

<u>Income sources</u>

Wages

Wages in the career job are estimated with person fixed effects 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 wage income for the persons in our sample over the period to be 1.93% (s.e.=0.23%) per year. Future real career wages were thus assumed to grow at this rate from the base reported in 1992. The full set of estimates from this model are as shown in Table 5A.

Wages 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. Bridge earnings are estimated as a function of the 1992 (career) wage and the age at which the respondent leaves the career job. The estimates are shown in Table 5A.

Defined benefit pensions

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 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.

Defined contribution pensions

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.

Social Security (retirement and disability) benefits

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.

Non-pension wealth

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).

Issues related to actuarial adjustment for individuals who return to work after beginning to receive SS benefits

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.

Additional information related to the computation of SS 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.

Defined Benefit Private Pension

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.

Defined Contribution Private Pensions

[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.

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 θ.[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 θ 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 θ, 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 θ, 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 θ 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.

Health Care Expenditure

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 basic model is given by equation (4):

(4) Ln Expenditures(t) = $\Lambda_1^k + \Lambda_2$ (Age-55) + $\Lambda_3\, \eta_t + e^k$

---

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

where k denotes the insurance group, and η is health as it is predicted by a model with age, income, education, and functional limitations as independent variables. The dependent variable in the health equation is a two-category variable indicating the presence of any self-reported limitations in the person's ordinary daily activities. 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. 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 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 5B.

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.

	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

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 (2002), 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.

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 $S^{NS}(t+1)$, $S(t+1)=S^{NS}(t+1)\cup\{DI(t+1), \eta_{t+1}, \varepsilon(t+1), \psi(t+1), HI^B(t+1), L(t+1)\}$,[50]

(B.1) $E[V(t+1, S(t+1)|S(t),d^C(t)=1]$

$$=PR(L(t+1)=1)\cdot PR(HI^B(t+1)=1)\cdot\int\int\int [V(t+1,S^{NS}(t+1),L(t+1)=1,\psi(t+1),\varepsilon(t+1),\eta_{t+1},HI^B(t+1)=1] \; dG(\varepsilon(t+1))\, dF(\psi(t+1))\, d\, H(\eta_{t+1}|\eta_t)$$

$$+PR(L(t+1)=0)\cdot PR(HI^B(t+1)=1)\cdot\int\int\int [V(t+1,S^{NS}(t+1),L(t+1)=0,\psi(t+1),\varepsilon(t+1),\eta_{t+1}, HI^B(t=1)=1] \; dG(\varepsilon(t+1))\, dF(\psi(t+1))\, d\, H(\eta_{t+1}|\eta_t)$$

$$+PR(L(t+1)=0)\cdot PR(HI^B(t+1)=2)\cdot\int\int\int [V(t+1,S^{NS}(t+1),L(t+1)=0,\psi(t+1),\varepsilon(t+1),\eta_{t+1}, HI^B(t=1)=2] \; dG(\varepsilon(t+1))\, dF(\psi(t+1))\, d\, H(\eta_{t+1}|\eta_t)$$

$$+PR(L(t+1)=1)\cdot PR(HI^B(t+1)=2)\cdot\int\int\int [V(t+1,S^{NS}(t+1),L(t+1)=1,\psi(t+1),\varepsilon(t+1),\eta_{t+1}, HI^B(t=1)=2] \; dG(\varepsilon(t+1))\, dF(\psi(t+1))\, d\, H(\eta_{t+1}|\eta_t)$$

where H, G, and F are the distribution functions of η_{t+1} given η_t, ε, and ψ respectively.[51]

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 will be discussed in detail in Section 3. Unfortunately, this distributional assumption, along with the equation (10) assumption about η_{t+1} given η_t and our use

[50] $\psi_t=\{\psi_t^C,\psi_t^B\}$

[51] 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.

of the empirical distribution of ψ, implies that equation (C.1) does not have a closed form solution.[52] 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

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

where η^{d}_{t+1}, $\varepsilon^{d}(t+1)$, and $\psi^{d}(t+1)$ and represent the dth of D draws from the distributions of η_{t+1} given η_t, $\varepsilon(t+1)$, 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[53]. 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.[54] This causes problems with the derivative-based optimization 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 (6). The additional parameter produces a smooth simulator which approaches the value in equation (C.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

---

[52]As shown in Rust (1987), the estimation of dynamic programming problems can be greatly facilitated by the presence of extreme value unobservables which produce closed form solutions for integrals such as those in equation (C.1). A recent example 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.

[53]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.

[54]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.

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

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).

As mentioned in Section 2.3.3, the presence of a serially correlated health variable implies that a modification must be made to the standard backwards recursion process for solving value functions. 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 η_3,...,η_T.

Once the range of values for η_1,...,η_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.[55] Equation (21) 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 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).[56]

[55]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.

[56]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. Evidence and intuition 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: Computation of the Likelihood function

The likelihood in equation (23) can be approximated as

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

$PR(\nu_2+\mu_2<-\pi X_2-\gamma Z_2)\bullet$

$PR(\nu_3+\mu_3<-\pi X_3-\gamma Z_3 \mid (\nu_2+\mu_2)^d)\bullet$

$PR(\nu_5+\mu_5<-\pi X_5-\gamma Z_5 \mid (\nu_2+\mu_2)^d, (\nu_3+\mu_3)^d)\bullet$

$PR(\varepsilon^I>-\Pi_1X_0-\Pi_2Z_0 \mid (\nu_2+\mu_2)^d,(\nu_3+\mu_3)^d,(\nu_5+\mu_5)^d, \nu_1^d, \nu_2^d, \nu_3^d,\nu_4^d,\nu_5^d,\kappa^d)\bullet$

$PR(\kappa^d +\lambda_\eta^C\nu_1^d+\varepsilon_1^C - \kappa^d - \lambda_\eta^B\nu_1^d-\varepsilon_1^B >V^*_B(1,S(1),\psi_1^d,\eta_1^d)-V^*_C(1,S(1),\psi_1^d,\eta_1^d \mid (\nu_2+\mu_2)^d,(\nu_3+\mu_3)^d,(\nu_5+\mu_5)^d, \nu_1^d, \nu_2^d,$

$\nu_3^d,\nu_4^d,\nu_5^d,\kappa^d,\varepsilon^{Id})$

•

•

•

$PR(\lambda_\eta^N\nu_5^d+\varepsilon_5^N - \lambda_\eta^A\nu_5^d-\varepsilon_5^A >V^*_A(5,S(5),\psi_5^d,\eta_5^d)-V^*_N(5,S(5),\psi_5^d,\eta_5^d \mid \bullet)$ }

where the superscript d denotes the dth of D total draws of a particular variable. ψ_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 (23), $(\nu_2+\mu_2)^d <0$ 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 $(\nu_3+\mu_3)^d$. $\nu_1^d,...,\nu_1^5$ are drawn unconditionally from the marginal distribution of $(\nu_1^d,...,\nu_1^5)$ given $(\nu_2+\mu_2)^d$, $(\nu_3+\mu_3)^d$, and $(\nu_5+\mu_5)^d$. Given $\nu_1^d,...,\nu_1^5$, the values of $\eta_1^d,...,\eta_5^d$ can be constructed using equation (12).[57] κ^d is drawn from its unconditional distribution (it is not correlated with any of the previously simulated values). $\varepsilon^{Id}>-\Pi_1X_0-\Pi_2Z_0$ 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 (23) and (C.1) are computed.

[57]One minor complication is that the health conditions Z_t from equation (12) 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).

The probability expression on the last line of equation (C.1) 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.1) feasible. Recall from Appendix C 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 $\eta_1^{*1},...,\eta_1^{*N^\eta}$ which will not be the same as η_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.1). 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 η_1^d and the smallest grid point greater than η_1^d.