Details of the Future Elderly Model

Online Supplement

Jeffrey Sullivan
Precision Health Economics, LLC
▪
Pierre-Carl Michaud
University of Quebec at Montreal
▪
Desi Peneva
Precision Health Economics, LLC
▪
Dana P. Goldman
University of Southern California

2013

Table of Contents

1 Functioning of the Dynamic Model

1.1 *Background*

The Future Elderly Model (FEM) is a microsimulation model originally developed out of an effort to examine health and health care costs among the elderly Medicare population (age 65+). A description of the previous incarnation of the model can be found in Goldman et al. [1] The original work was funded by the Centers for Medicare and Medicaid Services and carried out by a team of researchers at the RAND Corporation.

Since that time, various extensions have been implemented to the original model. The most recent versions now projects health outcomes for all Americans aged 51 and older and uses the Health and Retirement Study (HRS) as a host dataset rather than the Medicare Current Beneficiary Survey (MCBS). [2] The work has also been extended to include economic outcomes such as earnings, labor force participation and pensions. This work was funded by the National Institute on Aging through its support of the RAND Roybal Center for Health Policy Simulation (P30AG024968), the Department of Labor through contract J-9-P-2-0033, the National Institutes of Aging through the R01 grant "Integrated Retirement Modeling" (R01AG030824) and the MacArthur Foundation Research Network on an Aging Society. Finally, the computer code of the model was transferred from Stata to C++. This report incorporates these new development efforts in the description of the model.

All tables referenced in the following sections are available as an easy-to-read companion Excel workbook. Figures are available in the text.

1.2 Overview

The defining characteristic of the model is the modeling of real rather than synthetic cohorts, all of whom are followed at the individual level. This allows for more heterogeneity in behavior than would be allowed by a cell-based approach. Also, since the HRS interviews both respondent and spouse, we can link records to calculate household-level outcomes such as net income and Social Security retirement benefits, which depend on the outcomes of both spouses. The omission of the population younger than age 51 sacrifices little generality, since the bulk of expenditure on the public programs we consider occurs after age 50. However, we may fail to capture behavioral responses among the young.

The model has three core components:

• The initial cohort module predicts the economic and health outcomes of new cohorts of 51/52 year-olds. This module takes in data from the Health and Retirement Study (HRS) and trends calculated from other sources. It allows us to "generate" cohorts as the simulation proceeds, so that we can measure outcomes for the age 51+ population in any given year.

• The transition module calculates the probabilities of transiting across various health states and financial outcomes. The module takes as inputs risk factors such as smoking, weight, age and education, along with lagged health and financial states. This allows for a great deal of heterogeneity and fairly general feedback effects. The transition probabilities are estimated from the longitudinal data in the Health and Retirement Study (HRS).

• The policy outcomes module aggregates projections of individual-level outcomes into policy outcomes such as taxes, medical care costs, pension benefits paid, and disability benefits. This component takes account of public and private program rules to the extent allowed by the available outcomes. Because we have access to HRS-linked restricted data from Social Security records and employer pension plans, we are able to realistically model retirement benefit receipt.

Figure 1. Architecture of the FEM Model

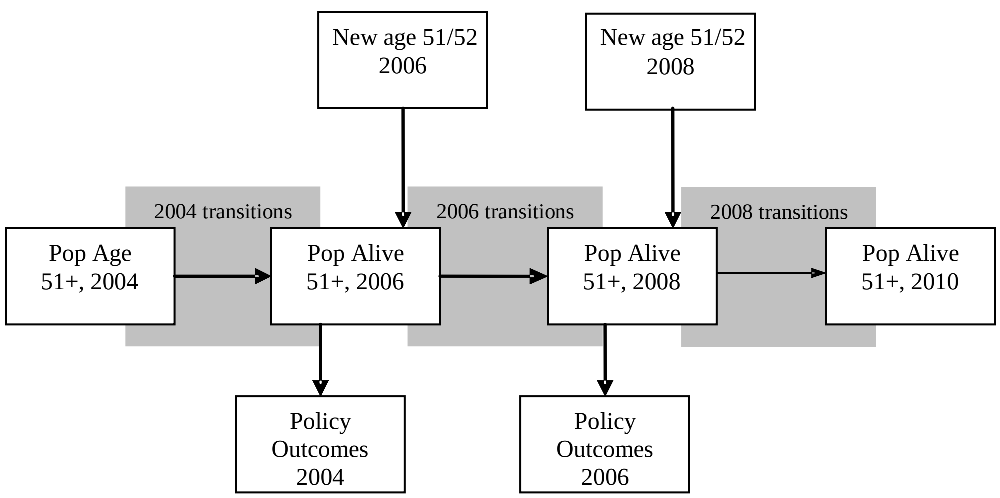


Figure 1 provides a schematic overview of the model. We start in 2004 with an initial population aged 51+ taken from the HRS. We then predict outcomes using our estimated transition probabilities (see section 4.1). Those who survive make it to the end of that year, at which point we calculate policy outcomes for the year. We then move to the following time period (two years later), when a new cohort of 51 and 52 year-olds enters (see section 5.1). This entrance forms the new age 51+ population, which then proceeds through the transition model as before. This process is repeated until we reach the final year of the simulation.

1.3 Comparison with Other Prominent Microsimulation Models

The FEM is unique among existing models that make health expenditure projections. It is the only model that projects health trends rather than health expenditures. It is also the only model that generates mortality out of assumptions on health trends rather than historical time series.

1.3.1 CBOLT Model (CBO)

The Congressional Budget Office (CBO) uses time-series techniques to project health expenditure growth in the short term and then makes an assumption on long-term growth. They use a long term growth of excess costs of 2.3 percentage points starting in 2020 for Medicare. They then assume a reduction in excess cost growth in Medicare of 1.5% through 2083, leaving a rate of 0.9% in 2083. For non-Medicare spending they assume an annual decline of 4.5%, leading to an excess growth rate in 2083 of 0.1%.

1.3.2 Centers for Medicare and Medicaid Services (CMS)

The Centers for Medicare and Medicaid Services (CMS) performs an extrapolation of medical expenditures over the first ten years, then computes a general equilibrium model for years 25 through 75 and linearly interpolates to identify medical expenditures in years 11 through 24 of their estimation. The core assumption they use is that excess growth of health expenditures will be one percentage point higher per year for years 25-75 (that is if nominal GDP growth is 4%, health care expenditure growth will be 5%).

2 Data Sources for Estimation

The Health and Retirement Study is the main data source for the model. We supplemented this data with merged Social Security covered earnings histories and data on health trends and health care costs coming from 3 major health surveys in the U.S. We describe these surveys below and the samples we selected for the analysis. We first list the variables used in the analysis. We then give details on the data sources.

Estimated Outcomes in Initial Conditions Model

Economic Outcomes	Health Outcomes
Employment	Hypertension
Earnings	Heart Disease
Wealth	Self-Reported Health
Defined Contribution Pension Wealth	BMI Status
Pension Plan Type	Smoking Status
AIME	Functional Status
Social Security Quarters of Coverage	
Health Insurance	

Estimated Outcomes in/from Transition Model

Economic Outcomes	Health Outcomes	Other Outcomes
Employment	Death	Income Tax Revenue
Earnings	Heart	Social Security Revenue
Wealth	Stroke	Medicare Revenue
Demographics	Cancer	Medical Expenses
Health Insurance	Hyper-tension	Medicare Part A Expenses
Disability Insurance Claim	Diabetes	Medicare Part B Expenses
Defined Benefit Claim	Lung Disease	Social Security Outlays
SSI Claim	Nursing Home	
Social Security Claim	BMI	
	Smoking Status	
	ADL Status	
	IADL Status	

2.1 Health and Retirement Study

The Health and Retirement Study (HRS), waves 1992-2004 are used to estimate the transition model. Interviews occur every two years. We use the dataset created by RAND (RAND HRS, version K) as our basis for the analysis. We use all cohorts in the analysis and consider sampling weights whenever appropriate. When appropriately weighted, the HRS in 2004 is representative of U.S. households where at least one member is at least 51. The HRS is also used as the host data for the simulation (pop 51+ in 2004) and for new cohorts (aged 51 and 52 in 2004).

The HRS adds new cohorts every six years, with the latest available cohort added in 2004, which is why that is used as the base year. When the 2010 data are finalized and released, the FEM will be updated to use the new cohort as its base population.

2.2 Social Security Covered Earnings Files

To get information on Social Security entitlements of respondents, we match the HRS data to the Social Security Covered Earnings files of 1992, 1993, 1998 and 2004 which provides information on earning histories of respondents as well as their entitlement to future Social Security benefits. We then construct the average indexed monthly earnings (AIME), the basis for the determination of benefit levels, from these earning histories. The AIME is constructed by first indexing using the National Wage Index (NWI) to the wage level when the respondent turns age 60. If this occurs after 2008, we project the evolution of the NWI using the average annual rate of change of the last 20 years (2.9% nominal). We then take the 35 highest years (if less than 35 years are available, remaining years are considered zero earning years) and take the average. We then convert back this annual amount on a monthly basis and convert back to $2004 U.S. dollars using the CPI. Quarters of coverage, which determine eligibility to Social Security are defined as the sum of posted quarters to the file. A worker is eligible to Social Security if he has accumulated at least 40 quarters of coverage. A worker roughly accumulates a quarter of coverage for every $4000 of coverage earnings up to a maximum of 4 per year. Not all respondents agree to have their record matched. Hence, there is the potential for non-representativeness. However, recent studies show that the extent of non-representativeness is quite small and that appropriate weighting using HRS weights mostly corrects for this problem. [3]

2.3 National Health Interview Survey (NHIS)

The NHIS contains individual-level data on height, weight, smoking status, self-reported chronic conditions, income, education, and demographic variables. It is a repeated cross-section done every year for several decades. But the survey design has been significantly modified several times. Before year 1997, different subgroups of individuals were asked about different sets of chronic conditions, after year 1997, a selected sub-sample of the adults were asked a complete set of chronic conditions. The survey questions are quite similar to that in HRS. As a result, for projecting the trends of chronic conditions for future 51/52 years olds, we only use data from 1997 to 2010. A review of survey questions is provided in Table 2. Information on weight and height were asked every year, while information on smoking was asked in selected years before year 1997, and has been asked annually since year 1997.

2.4 The Medical Expenditure Panel Survey (MEPS)

The MEPS, beginning in 1996, is a set of large-scale surveys of families and individuals, their medical providers (doctors, hospitals, pharmacies, etc.), and employers across the United States. The Household Component (HC) of the MEPS provides data from individual households and their members, which is supplemented by data from their medical providers. The Household Component collects data from a representative sub sample of households drawn from the previous year's National Health Interview Survey (NHIS). Since NHIS does not include the institutionalized population, neither does MEPS: this implies that we can only use the MEPS to estimate medical costs for the non-elderly population. Information collected during household interviews include: demographic characteristics, health conditions, health status, use of medical services, sources of medical payments, and body weight and height. Each year the household survey includes approximately 12,000 households or 34,000 individuals. Sample size for those aged 51-64 is about 4,500. MEPS has comparable measures of social-economic (SES) variables as those in HRS, including age, race/ethnicity, educational level, census region, and marital status.

2.5 Medicare Current Beneficiary Survey (MCBS)

The MCBS is a nationally representative sample of aged, disabled and institutionalized Medicare beneficiaries. The MCBS attempts to interview each respondent twelve times over three years, regardless of whether he or she resides in the community, a facility, or transitions between community and facility settings. The disabled (under 65 years of age) and oldest-old (85 years of age or older) are over-sampled. The first round of interviewing was conducted in 1991. Originally, the survey was a longitudinal sample with periodic supplements and indefinite periods of participation. In 1994, the MCBS switched to a rotating panel design with limited periods of participation. Each fall a new panel is introduced, with a target sample size of 12,000 respondents and each summer a panel is retired. Institutionalized respondents are interviewed by proxy. The MCBS contains comprehensive self-reported information on the health status, health care use and expenditures, health insurance coverage, and socioeconomic and demographic characteristics of the entire spectrum of Medicare beneficiaries. Medicare claims data for beneficiaries enrolled in fee-for-service plans are also used to provide more accurate information on health care use and expenditures.

3 Data Sources for Trends and Baseline Scenario

Two types of trends need to be projected in the model. First, we need to project trends in the incoming cohorts (the future new age 51/52 individuals). This includes trends in health and economic outcomes. Second, we need to project excess aggregate growth in real income and excess growth in health spending.

3.1 Data for Trends in Entering Cohorts

We used a multitude of data sources to compute U.S. trends. First, we used NHIS for chronic conditions and applied the methodology discussed by Goldman et al. [1] The method consists of projecting the experience of younger cohorts into the future until they reach age 51. The projection method is tailored to the synthetic cohorts observed in NHIS. For example, we observe a representative sample of age 35 individuals born in 1945 in 1980. We follow their disease patterns in 1980 to 1981 surveys by then selecting those aged 36 in 1981, accounting for mortality, etc.

We then collected information on other trends, i.e. for obesity and smoking, from other studies. [4-9] Table 3 presents the sources and Table 4 presents the trends we use in the baseline scenario. Table 5 presents the prevalence of obesity, hypertension, diabetes, and current smokers in 1978 and 2004, and the annual rates of change from 1978 to 2004. We refer the readers to the analysis in Goldman et al. for information on how the trends were constructed. [1]

3.2 Data for Other Projections

We make two assumptions relating to real growth in wages and medical costs. Firstly, as is done in the social security trustees report intermediate scenario, we assume a long term real increase in wages (earnings) of 1.1% per year. As is done by The Centers for Medicare and Medicaid Services, we assume excess real growth in medical costs (that is additional cost growth to GDP growth), as 1.5% in 2004, reducing linearly to 1% in 2033, .4% in 2053, and -.2% in 2083. We also include the Affordable Care Act cost growth targets as an optional cap on medical cost growth. Baseline medical spending figures presented assume those targets are met. GDP growth in the near term (through 2019) is based on CBO projections, with the OASDI Trustees assumption of 2% yearly afterwards.

3.3 Demographic Adjustments

We make two adjustments to the weighting in the Health and Retirement Study to match population counts from the Census. First, we post-stratify the HRS sample by 5 year age groups, gender and race and rebalance weights using the 2004 Current Population Survey (CPS). The CPS is itself matched to the decennial Census. Since we deleted some cases from the data and only considered the set of respondents with matched Social Security records, this takes account of selectivity based on these characteristics. We do this for both new cohort and host data set. The second adjustment we make is to scale up weights for future new cohorts using population projections from the Census Bureau. Again, we do this by race and gender. We use the intermediate net migration scenario produced by SSA in our simulation.

4 Estimation

In this section we describe the approach used to estimate the transition model, the core of the FEM, and the initial cohort model which is used to rejuvenate the model.

4.1 Transition Model

We consider a large set of outcomes for which we model transitions. Table 6 gives the set of outcomes considered for the transition model along with descriptive statistics and the population at risk when estimating the relationships.

Since we have a stock sample from the age 51+ population, each respondent goes through an individual-specific series of intervals. Hence, we have an unbalanced panel over the age range starting from 51 years old. Denote by j_{i0} the first age at which respondent *i* is observed and j_{iT_i} the last age when he is observed. Hence we observe outcomes at ages $j_i = j_{i0},..., j_{iT_i}$.

We first start with discrete outcomes which are absorbing states (e.g. disease diagnostic, mortality, benefit claiming). Record as $h_{i,j_i,m}$=1 if the individual outcome *m* has occurred as of age j_i. We assume the individual-specific component of the hazard can be decomposed in a time invariant and variant part. The time invariant part is composed of the effect of observed characteristics x_i and permanent unobserved characteristics specific to outcome *m*, $\eta_{i,m}$. The time-varying part is the effect of previously diagnosed outcomes $h_{i,j_i-1,-m}$, (outcomes other than the outcome *m*) on the hazard for *m*.[1] We assume an index of the form $z_{m,j_i} = x_i\beta_m + h_{i,j_i-1,-m}\gamma_m + \eta_{i,m}$. Hence, the latent component of the hazard is modeled as

$$\begin{aligned} &h^*_{i,j_i,m} = x_i\beta_m + h_{i,j_i-1,-m}\gamma_m + \eta_{i,m} + a_{m,j_i} + \varepsilon_{i,j_i,m}, \\ &m = 1,..., M_0,\ j_i = j_{i0},..., j_{iT_i},\ i = 1,..., N \end{aligned} \quad . \tag{1}$$

We approximate a_{m,j_i} with an age spline. After several specification checks, a node at age 75 appears to provide the best fit. This simplification is made for computational reasons since the joint estimation with unrestricted age fixed effects for each condition would imply a large number of parameters.

The outcome, conditional on being at risk, is defined as

$$\begin{aligned} &h_{i,j_i,m} = \max(I(h^*_{i,j_i,m} > 0), h_{i,j_i-1,m}) \\ &m = 1,..., M_0,\ j_i = j_{i0},..., j_{iT_i},\ i = 1,..., N \end{aligned} . \tag{2}$$

As mentioned in the text we consider 8 outcomes which are absorbing states. The occurrence of mortality censors observation of other outcomes in a current year. Mortality is recorded from exit interviews.

A number of restrictions are placed on the way feedback is allowed in the model. Table 7 documents restrictions placed on the transition model. We also include a set of other controls. A list of such controls is given in Table 8 along with descriptive statistics..

We have three other three other types of outcomes.

First, we have binary outcomes which are not an absorbing state. We specify latent indices as in (1) for these outcomes as well but where the lag dependent outcome also appears as a right-hand side variable. This allows for state-dependence.

---

[1] With some abuse of notation, $j_i - 1$ denotes the previous age at which the respondent was observed.

Second, we have ordered outcomes. These outcomes are also modeled as in (1) recognizing the observation rule is a function of unknown thresholds ς_m. Similarly to binary outcomes, we allow for state-dependence by including the lagged outcome on the right-hand side.

The third type of outcomes we consider are censored outcomes, earnings and financial wealth. Earnings are only observed when individuals work. For wealth, there is a non-negligible number of observations with zero and negative wealth. For these, we consider two part models where the latent variable is specified as in (1) but model probabilities only when censoring does not occur. In total, we have M outcomes.

The term $\varepsilon_{i,j_i,m}$ is a time-varying shock specific to age j_i. We assume that this last shock is normally distributed and uncorrelated across diseases. Unobserved difference η_{im} are persistent over time and are allowed to be correlated across diseases $m = 1,...,M$. We assume that these have a normal distribution with covariance matrix Ω_η.

The parameters $\theta_1 = (\{\beta_m, \gamma_m, \varsigma_m\}_{m=1}^{M}, vech(\Omega_\eta))$, can be estimated by maximum simulated likelihood. Given the normality distribution assumption on the time-varying unobservable, the joint probability of all time-intervals until failure, right-censoring or death conditional on the individual frailty is the product of normal univariate probabilities. Since these sequences, conditional on unobserved heterogeneity, are also independent across diseases, the joint probability over all disease-specific sequences is simply the product of those probabilities.

For a given respondent with frailty η_i observed from initial age j_{i0} to a last age j_{T_i}, the probability of the observed health history is (omitting the conditioning on covariates for notational simplicity)

$$l_i^{-0}(\theta;\eta_i,h_{i,j_{i0}}) = \left[\prod_{m=1}^{M-1}\prod_{j=j_{i1}}^{j_{T_i}} P_{ij,m}(\theta;\eta_i)^{(1-h_{ij-1,m})(1-h_{ij,M})}\right]\times\left[\prod_{j=j_{i1}}^{j_{T_i}} P_{ij,M}(\theta;\eta_i)\right] \quad (3)$$

We make explicit the conditioning on $h_{i,j_{i0}} = (h_{i,j_{i0},0},...,h_{i,j_{i0},M})'$, we have limited information on outcomes prior to this age.

To obtain the likelihood of the parameters given the observables, it remains to integrate out unobserved heterogeneity. The complication is that $h_{i,j_{i0},-m}$, the initial outcomes in each hazard is not likely to be independent of the common unobserved heterogeneity term which needs to be integrated out. A solution is to model the conditional probability distribution $p(\eta_i \mid h_{i,j_{i0}})$. [10] Implementing this solution amounts to including initial outcomes at baseline each hazard. This is equivalent to writing

$$\eta_i = \Gamma h_{i0} + \alpha_i$$
$$\alpha_i \sim N(0,\Omega_\alpha)$$

Therefore, this allows for permanent differences in outcomes due to differences in baseline outcomes. The likelihood contribution for one respondent's sequence is therefore given by

$$l_i(\theta; h_{i,j_{i0}}) = \int l_i(\theta; \alpha_i, h_{i,j_{i0}}) dF(\alpha_i) \quad (4)$$

To estimate the model, we make use of maximum simulated likelihood. We replace (4) with a simulated counterpart based on *R* draws from the distribution of α. We them optimize over this simulated likelihood using the BFGS algorithm. We could not obtain convergence of the joint estimator. So we assumed the distribution of α_i to be degenerate. This yielded the simpler estimation problem where each equation can be estimated separately.

One problem fitting the wealth and earnings distribution is that they have a long left-tail and wealth has some negative values. We use a generalization of the inverse hyperbolic sine transform (IHT), presented by MacKinnon and Magee. [11] First denote the variable of interest y. The hyperbolic sine transform is

$$y = \sinh(x) = \frac{\exp(y) - \exp(-y)}{2} \quad (5)$$

The inverse of the hyperbolic sin transform is

$$x = \sinh^{-1}(y) = h(y) = \log(y + (1 + y^2)^{0.5}) \quad (6)$$

Consider the inverse transformation. We can generalize such transformation, first allowing for a shape parameter θ,

$$r(y) = h(\theta y)/\theta \quad (7)$$

Such that we can specify the regression model as

$$r(y) = x\beta + \varepsilon, \varepsilon \sim N(0, \sigma^2). \quad (8)$$

A further generalization is to introduce a location parameter ω such that the new transformation becomes

$$g(y) = \frac{h(\theta(y + \omega)) - h(\theta\omega)}{\theta h'(\theta\omega)} \quad (9)$$

where $h'(a) = (1 + a^2)^{-1}$.

We specify (8) in terms of the transformation g. The shape parameters can be estimated from the concentrated likelihood for θ, ω. We can then retrieve β, σ by standard OLS.

Upon estimation, we can simulate

$$\tilde{g} = x\hat{\beta} + \sigma\tilde{\eta} \quad (10)$$

where η is a standard normal draw. Given this draw, we can retransform using (9) and (5)

$$h(\theta(y+\omega)) = \theta h'(\theta\omega)\tilde{g} + h(\theta\omega)$$

$$\tilde{y} = \frac{\sinh[\theta h'(\theta\omega)\tilde{g} + h(\theta\omega)] - \theta\omega}{\theta}$$

Tables 10-14 give parameter estimates for the transition models.

4.2 Goodness-of-Fit

To judge the goodness-of-fit of the model, we estimated parameters on the 1992-2004 estimation sample and simulated outcomes of 1992 HRS respondents up to 2004. We then compared simulated and actual outcomes in 2004. Table 15 presents the results. Some differences exist but in general the fit is satisfactory.

4.3 Quality Adjusted Life Years

As an alternative measure of life expectancy, we compute a quality adjusted life year based on the EQ-5D instrument, a health-related quality of life measure. The scoring system for EQ-5D was first developed by Dolan using a UK sample. [12] Later a scoring system based on a US sample was generated. [13] Since the HRS does not ask the appropriate questions for compute EQ-5D, but the MEPS does, we use a crosswalk from the MEPS to the HRS for persons not living in a nursing home. The final OLS regression used to compute QALY in the FEM is shown in Table 32. If a person is living in a nursing home, then an additional 0.10 is subtracted from the computed QALY.

5 Model for New Cohorts

We first discuss the empirical strategy, then present the model and estimation results. The model for new cohorts integrates information coming from trends among younger cohorts with the joint distribution of outcomes in the current population of age 51 respondents in the HRS.

5.1 Information Available and Empirical Strategy

For the transition model, we need to first to obtain outcomes listed in Table 16. Ideally, we need information on

$$f_t(y_{i1},...,y_{iM}) = f_t(y_i) \tag{11}$$

where t denotes calendar time, and $y_i = (y_{i1},...,y_{iM})$ is a vector of outcomes of interest whose probability distribution at time *t* is $f_t()$. Information on how the joint distribution evolves over time is not available. Trends in conditional distributions are rarely reported either.

Generally, we have from published or unpublished sources good information on trends for some moments of each outcome (say a mean or a fraction). That is, we have information on

$$g_{t,m}(y_{i,m}) \tag{12}$$

where $g_{t,m}()$ denotes the marginal probability distribution of outcome *m* at time *t.*
For example we know from the NHIS repeated cross-sections that the fraction obese is increasing by roughly 2% a year among 51 year olds. In statistical jargon this means we have information on how the mean of the marginal distribution of y_{im}, an indicator variable that denotes whether someone is obese, is evolving over time.

We also have information on the joint distribution at one point in time, say year t_0. For example, we can estimate the joint distribution on age 51 respondents in the 1992 wave of the HRS, $f_{t_0}(y_i)$.

We make the assumption that only some part of $f_t(y_i)$ evolves over time. In particular, we will model the marginal distribution of each outcome allowing for correlation across these marginals. The correlations will be assumed fixed while the mean of the marginals will be allowed to change over time.

5.2 Model and Estimation

Assume the latent model for $y_i^* = (y_{i1}^*,...,y_{iM}^*)'$,

$$y_i^* = \mu + \varepsilon_i \tag{13}$$

where ε_i is normally distributed with mean zero and covariance matrix Ω. It will be useful to write the model as

$$y_i^* = \mu + L_\Omega \eta_i \tag{14}$$

where L_Ω is a lower triangular matrix such that $L_\Omega L_\Omega' = \Omega$ **and** $\eta_i = (\eta_{i1},...,\eta_{iM})'$ are standard normal. We observe $y_i = \Gamma(y_i^*)$ which is a non-invertible mapping for a subset of the M outcomes. For example, we have binary, ordered and censored outcomes for which integration is necessary.

Because the mapping is non-invertible, integration needs to be performed to calculate the likelihood contributions $L_i(\theta \mid y_i)$. Integration needs to be done over a large number of dimensions. We will use maximum simulated likelihood to estimate the parameters of the model . The estimator is given by

$$\theta_{MSL} = \arg\max_{\theta=(\mu,\Omega)} \frac{1}{N}\sum_{i=1}^{N} \log \frac{1}{R}\sum_{r=1}^{R} \tilde{\Pr}(y_i \mid \theta)_r \qquad (15)$$

where $\frac{1}{R}\sum_{r=1}^{R} \tilde{\Pr}(y_i \mid \theta)_r$ is a consistent estimate of $\tilde{\Pr}(y_i \mid \theta)$. This estimator is consistent if both N, R tend to infinity. In practice, one can vary R to assess the bias of the estimator for smaller R. It is asymptotically efficient for $R/\sqrt{N}$ tending to infinity.

The vector μ can depend on some variables which have a stable distribution over time z_i (say race, gender and education). This way, estimation preserves the correlation with these outcomes without having to estimate their correlation with other outcomes. Hence, we can write

$$\mu_i = z_i \beta \qquad (16)$$

and the whole analysis is done conditional on z_i.

For binary and ordered outcomes, we fix $\Omega_{m,m}$=1 which fixes the scale. Also we fix the location of the ordered models by fixing thresholds as $\tau_0 = -\infty, \tau_1 = 0, \tau_K = +\infty$ where K denotes the number of categories for a particular outcome. Because some of the binary outcomes are rare, we fix correlations to zero between two outcomes if both fraction positive are below 10%. Furthermore, we fix to zero the correlation between selected outcomes (say earnings) and their selection indicator. Hence, we consider two-part models for these outcomes.

For exposition, we order the observed outcomes as binary, ordered, continuous and finally censored. The GHK simulator can be used to simulate $\Pr(y_i \mid \theta)$.

We start with the first outcome y_{i1}^*, a discrete outcome.

1. A draw of η_{i1} consistent with observed choice y_{i1} is

$$\tilde{\eta}_{i1} = \Phi^{-1}[\tilde{u}_{i1}\Phi(\tfrac{\bar{c}_{i1}-\mu_{i1}}{L_{\Omega,11}}) + (1-\tilde{u}_{i1})\Phi(\tfrac{\underline{c}_{i1}-\mu_{i1}}{L_{\Omega,11}})] \qquad (17)$$

   where $\bar{c}_{i1} = \begin{cases} +\infty & \text{if } y_{i1}=1 \\ 0 & \text{if } y_{i1}=0 \end{cases}, \underline{c}_{i1} = \begin{cases} 0 & \text{if } y_{i1}=1 \\ -\infty & \text{if } y_{i1}=0 \end{cases}$ and $\tilde{u}_{i1}$ is a uniform draw. The bounds are slightly different for ordered outcomes where thresholds are also estimated. In particular we have
   $\bar{c}_{i1} = \tau_k, \underline{c}_{i1} = \tau_{k-1}$ if $y_{i1} = k$
   where τ_k are parameters to be estimated.
2. The probability of that first outcome is $\tilde{\Pr}(y_{i1} \mid \theta) = \Phi(\frac{\bar{c}_{i1}-\mu_i}{L_{\Omega,11}}) - \Phi(\frac{\underline{c}_{i1}-\mu_i}{L_{\Omega,11}})$

3. Now a draw of η_{i2} consistent with y_{i2} and the draw $\tilde{\eta}_{i1}$ is given by

$$\tilde{\eta}_{i2} = \Phi^{-1}[\tilde{u}_{i2}\Phi(\frac{\bar{c}_{i2}-\mu_{i2}-L_{\Omega,21}\tilde{\eta}_{i1}}{L_{\Omega,22}}) + (1-\tilde{u}_{i2})\Phi(\frac{c_{i2}-\mu_{i2}-L_{\Omega,21}\tilde{\eta}_{i1}}{L_{\Omega,22}})]$$

4. Then the probability is given by

$$\tilde{\Pr}(y_{i1}, y_{i2} \mid \theta) = \tilde{\Pr}(y_{i1} \mid \theta)[\Phi(\frac{\bar{c}_{i2}-\mu_{i2}-L_{\Omega,21}\tilde{\eta}_{i1}}{L_{\Omega,22}}) - \Phi(\frac{c_{i2}-\mu_{i2}--L_{\Omega,21}\tilde{\eta}_{i2}}{L_{\Omega,22}})] \qquad (18)$$

5. Cycle trough 3 and 4 until end of discrete outcomes. Denote by $m_0 - 1$ the number of discrete outcomes.
6. An error consistent with the first continuous outcome is

$$\tilde{\eta}_{im_0} = \frac{y_{i,m_0} - \mu_{i,m_0} - \sum_{s=1}^{m_0-1} L_{\Omega,m_0,s}\tilde{\eta}_{is}}{L_{\Omega,m_0,m_0}}$$

7. The probability is $\tilde{\Pr}(y_{i,m_0} \mid \theta) = \frac{1}{L_{\Omega,m_0,m_0}}\phi(\frac{y_{i,m_0} - \mu_{i,m_0} - \sum_{s=1}^{m_0-1} L_{\Omega,m_0,s}\tilde{\eta}_{is}}{L_{\Omega,m_0,m_0}})$
8. Hence $\tilde{\Pr}(y_{i1},...,y_{i,m_0} \mid \theta) = \tilde{\Pr}(y_{i1},...,y_{i,m_0-1} \mid \theta)\tilde{\Pr}(y_{i,m_0} \mid \theta)$
9. Cycle trough 6 to 8 until reach $m_1 - 1$, the last continuous outcome.
10. Denote by m_1 the first censored outcome. Denote by y_{ij} the binary outcome that records whether y_{im_1} can be observed. A draw consistent with y_{i,m_1} is given by

$$\tilde{\eta}_{im_1} = \Phi^{-1}[\tilde{u}_{im_1}] \text{ if } y_{ij} = 0$$

and

$$\tilde{\eta}_{im_1} = \frac{y_{i,m1} - \mu_{i,m1} - \sum_{s=1}^{m_1-1} L_{\Omega,m1,s}\tilde{\eta}_{is}}{L_{\Omega,m_1,m_1}} \text{ if } y_{ij} = 1$$

If y_{im1} is continuous and given by a draw similar to (7) if a binary outcome.

11. The probability is then

$$\tilde{\Pr}(y_{i,m_1} \mid \theta) = \left[\frac{1}{L_{\Omega,m_1,m1}}\phi(\frac{y_{i,m1} - \mu_{i,m_1} - \sum_{s=1}^{m_1-1} L_{\Omega,m1,s}\tilde{\eta}_{is}}{L_{\Omega,m_1,m_1}})\right]^{I(y_{i,j}=1)}$$

For continuous and cumulative normal similar to (8) for discrete.

12. Cycle 10-11 until reach *M*.
13. Repeat 1-9 R times and calculate $\frac{1}{R}\sum_{r=1}^{R}\tilde{\Pr}(y_i \mid \theta)_r$.
14. Repeat for each $i = 1,...,N$

We use draws from Halton sequences to generate uniform random draws. [14] Note that draws $\{\{\{u_{im,r}\}_{m=1,...,M}\}_{r=1,...,R}\}_{i=1,...,N}$ are kept fixed trough estimation. For the first past, we used 10 draws along each dimension.

Because some parameters are naturally bounded, we reparametrize the problem to guarantee an interior solution. In particular, we parametrize

$$\Omega_{m,m} = \exp(\delta_m),\ m = m_0 - 1,...,M$$

$$\Omega_{m,n} = \tanh(\xi_{m,n})\sqrt{\Omega_{m,m}\Omega_{n,n}},\ m,n = 1,...,M$$

$$\tau_{m,k} = \exp(\gamma_{m,k}) + \tau_{k-1},\ k = 2,...,K_m - 1, m \text{ ordered}$$

and estimate the $(\delta_{m,m}, \xi_{m,n}, \gamma_k)$ instead of the original parameters. Table 17 gives parameter estimates for the indices while Table 18 gives parameter estimates of the covariance matrix in the outcomes.

The latent model is written as

$$y_i^* = \mu + L_\Omega \eta_i$$

Each marginal as a mean change equal to $E(y \mid \mu) = (1+\tau)g(\mu)$ where τ is the percent change in the outcome and $g()$ is a non-linear but monotone mapping. Since it is invertible, we can find the vector μ^* where $\mu^* = g^{-1}(E(y \mid \mu)/(1+\tau))$. We use these new intercepts to simulate new outcomes.

6 Government Revenue and Expenditures

This gives a limited overview of how revenues and expenditures of the government are computed. These functions are based on 2004 rules but we include predicted changes in program rules such changes based on year of birth (e.g. Normal retirement age).

We cover the following revenues and expenditures:

Revenues	**Expenditures**
Federal Income Tax	Social Security Retirement benefits
State and City Income Taxes	Social Security Disability benefits
Social Security Payroll Tax	Supplementary Security Income (SSI)
Medicare Payroll Tax	Medical Care Costs

6.1 Social Security Benefits

Workers with 40 quarters of coverage and of age 62 are eligible to receive their retirement benefit. The benefit is calculated based on the Average Indexed Monthly Earnings (AIME) and the age at which benefits are first received. If an individual claims at his normal retirement age (NRA) (65 for those born prior to 1943, 66 for those between 1943 and 1957, and 67 thereafter), he receives his Primary Insurance Amount (PIA) as a monthly benefit. The PIA is a piece-wise linear function of the AIME. If a worker claims prior to his NRA, his benefit is lower than his PIA. If he retires after the NRA, his benefit is higher. While receiving benefits, earnings are taxed above a certain earning disregard level prior to the NRA. An individual is eligible to half of his spouse's PIA, properly adjusted for the claiming age, if that is higher than his/her own retirement benefit. A surviving spouse is eligible to the deceased spouse's PIA. Since we assume prices are constant in our simulations, we do not adjust benefits for the COLA (Cost of Living Adjustment) which usually follows inflation. We however adjust the PIA bendpoints for increases in real wages.

6.2 Disability Insurance Benefits

Workers with enough quarters of coverage and under the normal retirement age are eligible for their PIA (no reduction factor) if they are judged disabled (which we take as the predicted outcome of DI receipt) and earnings are under a cap called the Substantial Gainful Activity (SGA) limit. This limit was $9720 in 2004. We ignore the 9 month trial period over a 5 year window in which the SGA is ignored.

6.3 Supplemental Security Income Benefits

Self-reported receipt of supplemental security income (SSI) in the HRS provides estimates of the proportion of people receiving SSI under what administrative data would suggest. To correct for this bias, we link the HRS with administrative data from the social security administration identifying those receiving SSI. In the linked administrative data, 3.96% of the population receives supplementary security income, while only 2.79% of the sample reports social security income. We therefore estimate a probit of receiving SSI as a function of self-reporting social security income, as well as demographic, health, and wealth. Table 19 contains the estimates for this model.

The benefit amount is taken from the average monthly benefits found in the 2004 Social Security Annual Statistical Supplement. We assign monthly benefit of $450 for person aged 51 to 64, and $350 for persons aged 65 and older.

6.4 Medical Costs Estimation

In the FEM, a cost module links a person's current state—demographics, economic status, current health, risk factors, and functional status—to 4 types of individual medical spending. The FEM models: total medical spending (medical spending from all payment sources), Medicare spending[2], Medicaid spending (medical spending paid by Medicaid), and out of pocket spending (medical spending by the respondent). These estimates are based on pooled weighted least squares regressions of each type of spending on risk factors, self-reported conditions, and functional status, with spending inflated to constant dollars using the medical component of the consumer price index. We use the 2002-2004 Medical Expenditure Panel Survey (n = 14,098) for these regressions for persons not Medicare eligible, and the 2002-2004 Medicare Current Beneficiary Survey (n = 33, 231) for spending for those that are eligible for Medicare. Those eligible for Medicare include persons eligible due to age (65+) or due to disability status. A comparison across these different sources is provided in Table 2.

In the baseline scenario, this spending estimate can be interpreted as the resources consumed by the individual given the manner in which medicine is practiced in the United States at the beginning of the 21st century. Table 20 shows the model estimation

[2] We estimate annual medical spending paid by specific parts of Medicare (Parts A, B, and D) and sum to get the total Medicare expenditures.

results for total, Medicaid, and out of pocket spending, while Table 21 shows the model estimation results for the Medicare spending. These estimation results only use the MCBS dataset.

Since Medicare spending has numerous components (Parts A and B are considered here), models are needed to predict enrollment. In 2004, 98.4% of all Medicare enrollees, and 99%+ of aged enrollees, were in Medicare Part A, and thus we assume that all persons eligible for Medicare take Part A. We use the 1999-2004 MCBS to model take up of Medicare Part B for both new enrollees into Medicare, as well as current enrollees without Part B. Estimates are based on weighted probit regression on various risk factors, demographic, and economic conditions. The HRS starting population for the FEM does not contain information on Medicare enrollment. Therefore another model of Part B enrollment for all persons eligible for Medicare is estimated via a probit, and used in the first year of simulation to assign initial Part B enrollment status. Estimation results are shown in Table 22. The MCBS data over represents the portion enrolled in Part B, having a 97% enrollment rate in 2004 instead of the 93.5% rate given by Medicare Trustee's Report. In addition to this baseline enrollment probit, we apply an elasticity to premiums of -0.10, based on the literature and simulation calibration for actual uptake through 2009. [15] [16] The premiums are computed using average Part B costs from the previous time step and the means-testing thresholds established by the ACA.

Since both the MEPS and MCBS are known to under-predict medical spending, we applied adjustment factors to the predicted three types of individual medical spending so that in year 2004, the predicted per-capita spending in FEM equal the corresponding spending in National Health Expenditure Accounts (NHEA), for age group 55-64 and 65 and over, respectively. Table 23 shows how these adjustment factors were determined by using the ratio of expenditures in the NHEA to expenditures predicted in the FEM.

The Medicare Current Beneficiaries Survey (MCBS) 2006 contains data on Medicare Part D. The data gives the capitated Part D payment and enrollment. When compared to the summary data presented in the CMS 2007 Trustee Report, the per capita cost is comparable between the MCBS and the CMS. However, the enrollment is underestimated in the MCBS, 53% compared to 64.6% according to CMS.

Since only one year of Part D enrollment is available in the MCBS, only a cross sectional model of Part D enrollment is estimated, rather than a transition model as with Part B enrollment. A probit model is estimated to link demographics, economic status, current health, and functional status to Part D enrollment - see Tables 24 and 25 for estimates. To account for both the initial under reporting of Part D enrollment in the MCBS, as well as the CMS prediction that Part D enrollment will rise to 75% by 2012, the constant in the probit model is increased by 0.22 in 2006, to 0.56 in 2012 and beyond. The per capita Part D cost in the MCBS matches well with the cost reported from CMS. An OLS regression using demographic, current health, and functional status is estimated for Part D costs.

The Part D enrollment and cost models are implemented in the Medical Cost module. The Part D enrollment model is executed conditional on the person being eligible for Medicare, and the cost model is executed conditional on the enrollment model leading a true result, after the Monte Carlo decision. Otherwise the person has zero Part

D cost. The estimated Part D costs are added with Part A and B costs to obtain total Medicare cost, and any medical cost growth assumptions are then applied.

6.5 Taxes

We consider Federal, State and City taxes paid at the household level. We also calculate Social Security taxes and Medicare taxes. HRS respondents are linked to their spouse in the HRS simulation. We take program rules from the OECD's Taxing Wages Publication for 2004. Households have basic and personal deductions based on marital status and age (>65). Couples are assumed to file jointly. Social Security benefits are partially taxed. The amount taxable increases with other income from 50% to 85%. Low income elderly have access to a special tax credit and the earned income tax credit is applied for individuals younger than 65. We calculate state and city taxes for someone living in Detroit, Michigan. The OECD chose this location because it is generally representative of average state and city taxes paid in the U.S. Since Social Security administrative data cannot be used jointly with Geocoded information in the HRS, we apply these hypothetical taxes to all respondents.

At the state level, there is a basic deduction for each member of the household ($3100) and taxable income is taxed at a flat rate of 4%. At the city level, there is a small deduction of $750 per household member and the remainder is taxed at a rate of 2.55%. There is however a tax credit that decreases with income (20% on the first 100$ of taxes paid, 10% on the following 50$ and 5% on the remaining portion).

We calculate taxes paid by the employee for Old-Age Social Insurance (SS benefits and DI) and Medicare (Medicaid and Medicare). It does not include the equivalent portion paid by the employer. OASI taxes of 6.2% are levied on earnings up to $97,500 (2004 cap) while the Medicare tax (1.45%) is applied to all earnings.

7 Example: Obesity Reduction Scenario

In addition the to the status quo scenario, the Future Elderly Model can be used to estimate the effects of numerous possible policy changes. One such set of policy simulations involves changing the trends of risk factors for chronic conditions. This is implemented by altering the incoming cohorts. A useful example is an obesity reduction scenario which rolls back the prevalence of obesity among 50 year-olds to its 1978 level by 2030, where it remains until the end of the scenario, in 2050. As seen in Table 27, this will decrease the prevalence of obesity among the age 50+ in 2050 from 46.5% to 23.7%, nearly a 50% drop. As compared with the status quo estimates (Table 26) the FEM predicts that by 2050, this will result in an increase in social security benefits of 2.28%, and a decrease in combined Medicare/ Medicaid expenditures of 4.37%.

8 Implementation of the FEM

The FEM is implemented in multiple parts. Estimation of the transition and cross sectional models is performed in Stata, the incoming cohort model is estimated in Ox, and the simulation is implemented in C++ to increase speed.

To match the two year structure of the Health and Retirement Study (HRS) data used to estimate the transition models, the FEM simulation proceeds in two year increments. The end of each two year step is designed to occur on July 1st to allow for easier matching to population forecasts from Social Security. A simulation of the FEM proceeds by first loading a population representative of the age 51+ US population in 2004, generated from HRS. In two year increments, the FEM applies the transition models for mortality, health, working, wealth, earnings, and benefit claiming with Monte Carlo decisions to calculate the new states of the population. The population is also adjusted by immigration forecasts from the US Census Department, stratified by race and age. If incoming cohorts are being used, the new 51/52 year olds are added to the population. The number of new 51/52 year olds added is consistent with estimates from the Census, stratified by race. Once the new states have been determined and new 51/52 year olds added, the cross sectional models for medical costs, and calculations for government expenditures and revenues are performed. Summary variables are then computed. Computation of medical costs includes the persons that died to account for end of life costs. Other computations, such as social security benefits and government tax revenues, are restricted to persons alive at the end of each two year interval. To eliminate uncertainty due to the Monte Carlo decision rules, the simulation is performed multiple times (typically 100), and the mean of each summary variable is calculated across repetitions.

FEM simulation takes as inputs assumptions regarding growth in the national wage index, normal retirement age, real medical cost growth, interest rates, cost of living adjustments, the consumer price index, significant gainful activity, and deferred retirement credit. The default assumptions are taken from the 2010 Social Security Intermediate scenario, adjusted for no price increases after 2010. Therefore simulation results are in real 2009 dollars.

Different simulation scenarios are implemented by changing any of the following components: incoming cohort model, transition models, interventions that adjust the probabilities of specific transition, and changes to assumptions on future economic conditions.

List of Tables

Table 1. Comparison of Questions on Chronic Health Conditions in HRS and SHARE

	HRS	SHARE
Question	Has a doctor ever told you that you have …	Has a doctor ever told you that you had any of the conditions on this card? Please tell me the number or numbers of the conditions
Heart Disease	… a heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems?	… A heart attack including myocardial infarction or coronary thrombosis or any other heart problem including congestive heart failure
Hypertension	… high blood pressure or hypertension?	… High blood pressure or hypertension
Stroke	… a stroke?	… A stroke or cereberal vascular disease
Diabetes	… diabetes or high blood sugar?	… Diabetes or high blood sugar
Lung Disease	… chronic lung disease such as chronic bronchitis or emphysema?	… Chronic lung disease such as chronic bronchitis or emphysema
Cancer	… cancer or a malignant tumor, excluding minor skin cancers?	… Cancer or malignant tumour, including leukaemia or lymphoma, but excluding minor skin cancers

Table 2. Comparison of Prevalence and Questions on Chronic Health Conditions in HRS, NHIS, MEPS, and MCBS

Condition	Prevalence (%)							Description HRS	Description NHIS	Description MEPS	Description MCBS
	HRS (55-64)	NHIS (55-64)	MEPS (55-64)	HRS (65+)	NHIS (65+)	MCBS (65+)	MEPS (65+)				
Cancer	7%	8%	6%	15%	15%	18%	11%	Has a doctor ever told you that you have cancer or a malignant tymor, excluding minor skin cancers?	Have you ever been told be a doctor or other health professional that you had … cancer or a malignanacy of any kind? (WHEN RECODED, SKIN CANCERS WERE EXCLUDED)	List all the conditions that have bothered (the person) from (START time) to (END time) CCS codes for the conditions list are 11-21, 24-45	Has a doctor ever told you that you had any (other) kind of cancer malignancy, or tumor other than skin cancer?
Heart Diseases	14%	17%	15%	30%	31%	41%	31%	Has a doctor ever told you that you had a heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems?	Four separate questions were asked about whether ever told by a docotor or oither health professional that had: CHD, Angina, MI, other heart problems.	have you ever been told by a doctor or health professional that you have … CHD; Angina; MI; other heart problems	Siz separate questions were asked about whether ever told by a doctor that had: Angina or MI; CHD; other heart problems (included four questions)
Stroke	4%	3%	4%	10%	9%	12%	10%	Has a doctor ever told you that you had a stroke?	Have you EVER been told by a doctor or other health professional that you had a stroke?	If Female, add: [Other than during pregnancy,] Have you ever been told by a doctor or health professional that you have a stroke or TIA (transient ischemic attack)	[Since (PREV< SUPP. RD. INT. DATE),] has a doctor (ever) told (you/SP) that (you/he/she) had a stroke, a bran hemorrhage, or a cerebrovascular accident?

Table 2. Comparison of Prevalence and Questions on Chronic Health Conditions in HRS, NHIS, MEPS, and MCBS

Diabetes	12%	13%	12%	16%	15%	20%	17%	Has a doctor ever told you that you have diabetes or high blood sugar?	If Female, add: [Other than during pregnancy,] Have you ever been told by a doctor or health professional that you have diabetes or sugar diabetes?	If Female, add: [Other than during pregnancy,] Have you hever been told by a doctor or health professional that you have diabetes or sugar diabetes?	has a doctor (ever) told (you/SP) that (you/he/she) had diaebtes, high blood sugar, or sugar in (your/his/her) urine? [DO NOT INCLUDE BOERDERLINE PREGNANCY, OR PRE-DIABETEIC DIABETES.]
Hypertension	40%	42%	42%	52%	54%	61%	58%	Has a doctor ver told you that you have high blood pressure or hypertension?	Have you EVER been told by a doctor or other health professional that you had Hypertension, also called high blood pressure?	Have you EVER been told by a doctor or other health professional that you had Hypertension, also called high blood pressure?	has a doctor (ever) told (you/SP) that (you/he/she) (still) (had) (have/has) hypertension, sometimes called high blood pressure?
Lung Disease	7%	8%	7%	10%	10%	14%	8%	Has a doctor ever told you that you have chronic lung disease such a schronic bronchitis or emphysema? [IWER: DO NOT INCLUDE ASTHMA]	Question 1: During the PAST 12 MONTHS, have you ever been told by a doctor or other health professional that you had chronic bronchitis? Question 2: Have you EVER been told by a docotor or other health professional that you had … emphysema?	List all the conditions that have bothered (the person) from (START time) to (END time) CCS codes for the conditions list are 127, 129-312	has a doctor (ever) told (you/SP) that (you/he/she) had emphysema, asthma, or COPD? [COPD=CHRONIC OBSTRUCTIVE PULMONARY DISEASE.]
Overweight	40%	38%	39%	38%	36%	38%	38%		self-reported body weight and height		
Obese	28%	31%	29%	18%	23%	21%	22%				

Table 3. Data Sources and Methods for Projecting Trends for Future Cohorts

	Data source	Projection method	Directly obtained from other sources
Chronic conditions	National Health Interview Survey 1997-2006	Assume no recovery	There are other forecasts (Honeycutt, 2003, Mainous 2007) for the trends of diabetes in the U.S population; we compare their forecasts to ours and they are reasonably close
	cohort-mortality rate from mortality.org	Use synthetic cohort approach to estimate age-specific incidence rate for each condition, using NHIS 1997-2006	
	Assumed annual mortality improvement rate for year 2005-2030: 0.8% per year. Assume relative risks of mortality for each condition: rr = 2 for cancer, diabetes and heart and rr = 1.5 for hypertension, lung and stroke	Baseline prevalence is obtained from the NHIS 2003-2005 pooled data	
Cancer		Use Markov model to model the transition into a certain condition or die from 2005 to 2030	
Diabetes			
Heart			
Hypertension			
Lung			
Stroke			
Over-weight and obese	Prevalence of over-weight and obese for aged 46-56 from year 2001 to 2030, generated by Ruhm upon request		Ruhm, Christopher J., "Current and Future Prevalence of Obesity and Severe Obesity in the United States", *Forum for Health Economics and Policy*, Vol. 10, No. 2 (Obesity), Article 6, 2007, 1-26.
Ever-smoked and smoking now	Status quo - Tobacco control policies will be frozen in place as of the beginning of 2006, with excise tax rates assumed to be adjusted for inflation.		Forecast of prevalence of ever-smoked and smoking now for aged 45-54 from year 2005 to 2025, by David Levy (2006)
Any DB from current job	Prevalence of DB entitlement from current job among aged 50-55, in HRS 1992 and 2004		Historical trends of DB participation rates among all persons by different birth cohorts and by age, by Poberta 2007 (a)
Any DC from current job	Prevalence of DC entitlement from current job among aged 50-55, in HRS 1992 and 2004		Forecast of DC participation rates among all persons by different birth cohorts and by age, by Poberta 2007 (b)
Hispanic	Projection of population from US census Bureau, Interim projection consistent with 2000 census (2004), Projection of population from US census Bureau, middle series, final projection consistent with 1990 census (2000)		
Non-Hispanic black	Since the interim projection consistent with 2000 census doesn't provide projection for all race/ethnicity categories, we cannot obtain the projection of non-Hispanic black population. As a result I turn to the final projection consistent with 1990 census and find out what proportion of the black population is non-Hispanic and the proportion is approximately 95%. http://www.census.gov/population/www/projections/natdet-D5.html		
Population size 50-52			

Table 4. Projected Trends in the U.S.

		Prevalence/Means relative to year 2004					
		2004	2010	2020	2030	2040	2050
Prevalence of binary outcomes	Hypertension	100%	104%	107%	109%	111%	113%
	Heart Disease	100%	95%	91%	88%	85%	83%
	Diabetes	100%	112%	122%	127%	131%	136%
Prevalence of highest category in ordered outcomes	BMI Status - obesity	100%	124%	172%	238%	303%	328%
	Smoking Status - smoking now	100%	94%	73%	60%	50%	41%
Prevalence of censored discrete outcomes	Any DB Plan	100%	89%	72%	59%	48%	39%
	Any DC Plan	100%	114%	141%	156%	156%	156%

Table 5. Prevalence of Obesity, Hypertension, Diabetes and Current Smokers Among Aged 47-56 in 1978 and in 2004

	1978	2004	Annual rate of change in prevalence rate from 1978 to 2004
Obesity (BMI >=30 kg/m^2)	15.70%	31.60%	0.027
Hypertension (Self-reported)	29.60%	33.00%	0.004
Diabetes (Self-reported)	4.80%	8.60%	0.022
Current smokers	39.50%	26.20%	-0.016

Prevalence in 1978 is based on NHANES II 1976-1980; Prevalence in 2004 is based on NHANES 2001-2006 pooled data.

BMI is calculated using self-reported weight and height

Table 6. Outcomes in the Transition Model

			Type	mean/fraction	At risk
Disease	heart disease		biannual incidence	3.08%	undiagnosed
	hypertension		biannual incidence	3.99%	undiagnosed
	stroke		biannual incidence	1.58%	undiagnosed
	lung disease		biannual incidence	1.39%	undiagnosed
	cancer		biannual incidence	1.96%	undiagnosed
	diabetes		biannual incidence	1.95%	undiagnosed
Risk Factors	Smoking Status	never smoked	ordered	38.61%	all
		ex smoker	ordered	40.24%	all
		current smoker	ordered	14.28%	all
	Log BMI		continuous	3.27	all
	ADL Status	no ADLs	ordered	85.01%	all
		1 ADL	ordered	7.13%	all
		2 ADLS	ordered	3.17%	all
		3+ ADLS	ordered	4.69%	all
	IADL Status	no IADLs	ordered	91.07%	all
		1 IADL	ordered	5.08%	all
		2+ IADLs	ordered	3.84%	all
LFP & Benefits	working		prevalence	40.71%	age < 75
	DB pension receipt		biannual incidence	8.59%	eligible & not receiving
	SS benefit receipt		biannual incidence	8.63%	eligible & not receiving
	DI benefit receipt		prevalence	7.34%	eligible & age < 65
	Any health insurance		prevalence	89.35%	age < 65
	SSI receipt		prevalence	3.15%	all
	Nursing Home residency		prevalence	1.98%	all
	Death		biannual incidence	6.70%	all
Financial Resources $USD 2004	financial wealth		median	$ 161,000.00	all positive wealth
	earnings		median	$ 2,930.90	all working
	wealth positive		prevalence	96.49%	all

Table 7. Restrictions on Health Outcomes

	Outcome at time T																			
Value at time T-1	heart disease	hypertension	stroke	lung disease	diabetes	cancer	disability	mortality	smoking status	BMI	Any HI	DI Claim	SS Claim	DB Claim	SSI Claim	Nursing Home	Work	Earnings	Nonzero Wealth	Wealth
heart disease			X				X	X	X	X	X	X	X	X	X	X	X	X	X	X
blood pressure	X		X				X	X	X	X	X	X	X	X	X	X	X	X	X	X
stroke							X	X	X	X	X	X	X	X	X	X	X	X	X	X
lung disease							X	X	X	X	X	X	X	X	X	X	X	X	X	X
diabetes	X	X	X				X	X	X	X	X	X	X	X	X	X	X	X	X	X
cancer			X				X	X	X	X	X	X	X	X	X	X	X	X	X	X
disability							X	X	X	X	X	X	X	X	X	X	X	X	X	X
claimed DI											X	X	X	X	X		X	X	X	X
claimed SS											X			X	X		X	X	X	X
claimed DB													X		X		X	X	X	X
claimed SSI															X					
work											X	X	X		X		X	X	X	X
earnings											X	X	X	X	X		X	X	X	X
nonzero wealth											X	X	X	X	X	X	X	X	X	X
wealth											X	X	X	X	X	X	X	X	X	X
nursing home stay															X	X			X	X

Table 8. Descriptive Statistics on Exogenous Controls

Control variable	Unweighted Statistics			
	Mean	Standard Deviation	Minimum	Maximum
age	69.13	10.71	32.08	111.7
black	0.136	0.343	0	1
hispanic	0.0783	0.269	0	1
less than high school	0.250	0.433	0	1
college education	0.388	0.487	0	1
male	0.436	0.496	0	1
ever smoked (includes current)	0.608	0.488	0	1
widowed at baseline	0.219	0.414	0	1
single at baseline	0.153	0.360	0	1
log AIME/10 at baseline	0.684	0.138	-0.116	0.904
log quarters/10 at baseline	0.437	0.0895	0	0.536
Any DB at baseline	0.174	0.379	0	1
NRA 57-61	0.0469	0.211	0	1
NRA 62-63	0.0314	0.174	0	1
NRA 64+	0.0655	0.247	0	1
Any DC at baseline	0.142	0.349	0	1
(IHT of DC wealth in 1000s)/100 if any DC, zero else	0.00484	0.0133	0	0.0782

Table 10. Transition of Mortality and Chronic Conditions (coefficient/t statistic)

	Mortality	Heart disease	Stroke	Cancer	Hypertension	Diabetes	Lung disease
Non-Hispanic black	0.0338 (1.13)	-0.176 (-4.92)	-0.0121 (-0.28)	-0.0679 (-1.65)	0.142 (3.96)	0.0696 (1.84)	-0.249 (-5.27)
Hispanic	-0.111 (-2.49)	-0.246 (-4.94)	-0.0360 (-0.59)	-0.135 (-2.33)	0.0353 (0.83)	0.185 (3.87)	-0.248 (-3.81)
Less than high school	0.0149 (0.61)	0.0539 (1.85)	0.000563 (0.02)	-0.0174 (-0.52)	0.0321 (1.11)	0.0998 (2.98)	0.0355 (0.99)
Some college and above	-0.0262 (-1.05)	0.000189 (0.01)	0.0243 (0.70)	0.0398 (1.39)	-0.0237 (-0.96)	0.00987 (0.31)	-0.0366 (-1.06)
Male	0.225 (9.18)	0.149 (5.48)	0.0471 (1.33)	0.144 (4.79)	-0.141 (-5.40)	0.137 (4.27)	-0.0146 (-0.42)
Min(Lag Age, 63)	0.0152 (3.53)	0.0175 (4.70)	0.0247 (4.38)	0.0265 (6.17)	0.0202 (6.38)	0.0224 (5.33)	0.0170 (3.61)
Min(Max(Lag Age - 63, 0), 73-63)	0.0285 (7.40)	0.0275 (6.57)	0.0162 (3.01)	0.0233 (5.05)	0.0163 (3.96)	-0.00513 (-1.03)	0.0151 (2.84)
Max(Lag Age - 73, 0)	0.0486 (19.29)	0.0215 (5.87)	0.0214 (5.18)	-0.00230 (-0.53)	0.00143 (0.36)	-0.00284 (-0.57)	-0.00358 (-0.73)
Lag of Heart disease	0.191 (5.63)		0.178 (3.44)				
Lag of Stroke	0.158 (3.64)						
Lag of Cancer	0.634 (17.04)		0.00469 (0.06)				
Lag of Hypertension	0.142 (3.98)	0.210 (5.22)	0.155 (2.91)				
Lag of Diabetes	0.175 (4.05)	0.130 (2.35)	0.226 (3.61)		0.281 (4.55)		
Lag of Lung disease	0.390 (8.78)						
Lag of Has exactly 1 IADL	0.196 (5.61)						
Lag of Has 2 or more IADLs	0.467 (10.51)						
Lag of Has exactly 1 ADL	0.203 (6.39)						
Lag of Has exactly 2 ADLs	0.326 (7.35)						
Lag of Has 3 or more ADLs	0.540 (12.87)						
Lag of Current smoking	0.0990 (2.38)	0.153 (2.90)	0.0532 (0.80)	0.0290 (0.52)	0.0260 (0.54)	-0.145 (-2.52)	0.224 (3.97)
Lag of Widowed	0.0244 (0.60)	-0.0318 (-0.57)	0.0248 (0.39)	-0.00751 (-0.12)	0.148 (2.75)	0.0247 (0.36)	0.00295 (0.04)
Init. of Heart disease	0.0270 (0.73)		-0.0447 (-0.77)	0.0595 (1.71)	0.0314 (0.88)	0.0720 (1.94)	0.167 (4.42)
Init. of Stroke	-0.00799 (-0.15)	0.00645 (0.11)		0.0281 (0.45)	0.0996 (1.38)	0.0623 (0.97)	-0.218 (-2.84)
Init. of Cancer	-0.321 (-6.98)	0.139 (3.39)	0.00568 (0.06)		0.0303 (0.72)	-0.0262 (-0.48)	0.129 (2.56)
Init. of Hypertension	-0.0201 (-0.57)	0.0101 (0.25)	0.0747 (1.41)	0.0523 (2.00)		0.195 (7.35)	0.0662 (2.18)
Init. of Diabetes	0.136 (2.82)	0.119 (1.91)	0.0111 (0.16)	-0.0608 (-1.35)	-0.0898 (-1.25)		0.00432 (0.09)
Init. of Lung disease	-0.0665 (-1.28)	0.176 (3.68)	0.146 (2.69)	0.0851 (1.64)	0.000812 (0.02)	0.0285 (0.50)	
Init. of Ever smoked	0.115	0.0409	0.0182	0.0864	-0.0303	0.0475	0.278

Table 10. Transition of Mortality and Chronic Conditions (coefficient/t statistic)

	(4.81)	(1.60)	(0.55)	(3.04)	(-1.24)	(1.58)	(7.51)
Init. of Current smoking	0.164	0.00980	0.0876	0.0549	0.0587	0.197	0.240
	(4.18)	(0.19)	(1.39)	(1.04)	(1.26)	(3.66)	(4.32)
Init. of Has exactly 1 IADL	-0.0123	0.0148	0.0663	-0.0694	0.0130	-0.0833	0.0411
	(-0.34)	(0.35)	(1.27)	(-1.39)	(0.32)	(-1.71)	(0.81)
Init. of Has 2 or more IADLs	0.0522	0.321	0.250	-0.121	0.0713	0.263	0.161
	(0.61)	(2.23)	(1.48)	(-0.61)	(0.40)	(1.50)	(0.95)
Init. of Has exactly 1 ADL	0.0104	0.160	0.0767	0.0974	-0.00263	0.0357	0.147
	(0.32)	(4.08)	(1.61)	(2.14)	(-0.06)	(0.79)	(3.12)
Init. of Has exactly 2 ADLs	-0.0190	0.118	-0.0237	0.132	0.0517	-0.0790	0.173
	(-0.28)	(1.07)	(-0.20)	(1.10)	(0.44)	(-0.59)	(1.41)
Init. of Has 3 or more ADLs	-0.180	0.0878	-0.316	0.184	-0.315	-0.163	0.195
	(-2.47)	(0.75)	(-1.87)	(1.44)	(-2.05)	(-1.16)	(1.52)
Init. of Widowed	0.0502	0.0353	0.0414	-0.0224	-0.114	0.0894	0.0332
	(1.15)	(0.58)	(0.61)	(-0.32)	(-1.89)	(1.20)	(0.43)
Init. of Single	0.101	-0.0131	-0.0892	0.0952	-0.0341	0.0609	0.0618
	(3.20)	(-0.36)	(-1.80)	(2.42)	(-0.99)	(1.49)	(1.41)
Init. of R working for pay	-0.146	0.0948	0.165	0.0140	-0.0699	-0.0442	0.0701
	(-2.05)	(1.40)	(1.89)	(0.18)	(-1.07)	(-0.53)	(0.79)
Init. of (IHT of earnings in 1000s)/100 if working zero otherwise	-0.591	-1.194	-6.223	0.234	1.182	-0.919	-5.198
	(-0.30)	(-0.67)	(-2.54)	(0.12)	(0.69)	(-0.42)	(-2.14)
Init. of Non-pension wlth(hatota) not zero	-0.0173	0.0472	0.106	0.00546	-0.0429	0.0532	0.0841
	(-0.33)	(0.66)	(1.25)	(0.06)	(-0.56)	(0.69)	(1.01)
Init. of (IHT of hh wlth in 1000s if positive)/100 zero otherwise	-0.781	-1.276	-2.140	1.519	-0.862	-1.044	-2.425
	(-1.54)	(-2.19)	(-3.01)	(2.16)	(-1.56)	(-1.58)	(-3.56)
AIME in ini.intw (-9=no match)	-0.128	0.301	-0.396	0.427	0.752	0.131	0.118
	(-0.62)	(1.22)	(-1.32)	(1.56)	(3.14)	(0.44)	(0.37)
Quarters of earnings in ini.intw (-9=no match)	0.447	-0.192	0.638	-0.674	-0.894	-0.329	-0.0417
	(1.45)	(-0.52)	(1.43)	(-1.65)	(-2.50)	(-0.74)	(-0.09)
Init. of Health fair/poor	0.137	0.140	0.113	-0.0136	0.0292	0.106	0.190
	(5.74)	(4.60)	(3.15)	(-0.38)	(0.90)	(3.07)	(5.22)
Init. of Any DB from current job RND VG	-0.0324	-0.0620	-0.105	0.0636	0.0506	0.0486	-0.0201
	(-0.34)	(-0.83)	(-0.86)	(0.86)	(0.81)	(0.59)	(-0.19)
Init. of Normal DB Retirement Age 60-61	0.0590	0.105	0.0372	-0.0282	-0.0624	0.00345	0.0843
	(0.53)	(1.21)	(0.26)	(-0.32)	(-0.83)	(0.04)	(0.68)
Init. of Normal DB Retirement Age 62-64	0.0374	0.0765	0.152	-0.0338	0.0265	0.141	0.169
	(0.32)	(0.81)	(1.03)	(-0.35)	(0.33)	(1.42)	(1.32)
Init. of Normal DB Retirement Age 65+	0.0833	0.112	0.0329	-0.0333	0.00991	0.0555	0.0967
	(0.80)	(1.37)	(0.24)	(-0.40)	(0.14)	(0.62)	(0.83)
Init. of Any DC from current job RND VG	-0.151	-0.0686	0.117	0.101	0.159	-0.0651	-0.00646
	(-1.59)	(-0.96)	(1.12)	(1.30)	(2.63)	(-0.76)	(-0.07)
Init. of (IHT of DC wlth in 1000s)/100 if any DC zero otherwise	1.839	1.489	-1.838	-3.498	-3.450	-0.0562	0.652
	(0.80)	(0.86)	(-0.69)	(-1.84)	(-2.28)	(-0.03)	(0.26)
Min(log(Init of BMI), log(30))		-0.120	-0.253	-0.111	0.894	0.816	-0.276
		(-0.66)	(-1.15)	(-0.55)	(5.04)	(3.45)	(-1.24)
Max(log(Init of BMI) - log(30), 0)		0.496	-0.670	0.344	0.428	1.272	0.765
		(1.88)	(-2.04)	(1.08)	(1.46)	(4.96)	(2.22)
Init. Of Min(log(Init of BMI), log(30))		0.245	0.196	0.0678	-0.0471	1.347	-0.0142
		(1.30)	(0.85)	(0.32)	(-0.26)	(5.56)	(-0.06)
Init of Max(log(Init of BMI) - log(30), 0)		-0.0759	0.725	-0.0539	-0.0814	-0.0241	0.0573
		(-0.28)	(2.28)	(-0.16)	(-0.25)	(-0.09)	(0.16)
Log(Time Between Interviews)		0.338	0.530	0.537	0.344	0.423	0.220
		(3.67)	(4.46)	(5.22)	(3.95)	(3.95)	(1.87)
Constant	-3.564	-4.001	-4.203	-4.164	-5.716	-10.94	-2.799
	(-13.49)	(-10.05)	(-7.85)	(-9.33)	(-15.56)	(-20.28)	(-5.75)

t statistics in parentheses

Table 11. Transition of Economic Outcomes (coefficient/t statistic)

	Boolean for Any ADLs	Boolean for Any IADLs	HI cov -gov/emp/other	Claiming SSDI	Claiming DB	Claiming SSI	R live in nursingh ome at interview	Non-pension wlth(hatota) not zero	Claiming OASI	R working for pay
Non-Hispanic black	0.0136	0.0866	-0.137	-0.0327	0.0892	0.0627	-0.191	-0.430	-0.0707	-0.00492
	(0.58)	(3.25)	(-4.27)	(-0.65)	(1.44)	(1.25)	(-3.10)	(-10.99)	(-1.73)	(-0.20)
Hispanic	0.0970	0.0675	-0.537	-0.217	0.0648	-0.0176	-0.517	-0.453	-0.0390	-0.0178
	(3.22)	(1.96)	(-15.14)	(-3.04)	(0.68)	(-0.27)	(-4.74)	(-9.12)	(-0.74)	(-0.56)
Less than high school	0.0353	0.136	-0.281	0.0321	0.0166	0.0662	-0.0280	-0.188	0.0510	-0.107
	(1.78)	(6.04)	(-9.84)	(0.70)	(0.25)	(1.41)	(-0.56)	(-4.64)	(1.37)	(-4.71)
Some college and above	0.00250	-0.0639	0.138	-0.125	-0.143	-0.0568	0.0722	0.00111	-0.169	0.0763
	(0.13)	(-2.76)	(4.95)	(-2.77)	(-3.18)	(-1.10)	(1.39)	(0.02)	(-5.46)	(4.32)
Male	-0.0170	0.223	-0.195	0.0122	0.0782	0.199	0.0587	-0.0402	-0.202	-0.0290
	(-0.88)	(9.91)	(-6.92)	(0.27)	(1.64)	(4.26)	(1.18)	(-0.98)	(-6.21)	(-1.55)
Min(Lag Age, 63)	0.00984	-0.0169	0.00813	-0.00289	0.137	-0.00922	0.0463			
	(3.76)	(-5.67)	(2.68)	(-0.57)	(19.96)	(-1.59)	(3.22)			
Min(Max(Lag Age - 63, 0), 73-63)	0.00769	0.0192			-0.0602	-0.0726	0.0370			
	(2.58)	(5.46)			(-3.15)	(-5.67)	(4.01)			
Max(Lag Age - 73, 0)	0.0498	0.0469				0.00322	0.0556			
	(19.67)	(17.22)				(0.30)	(13.01)			
Lag of Heart disease	0.112	0.116	0.209	0.308	0.00928	0.136	0.0419	0.00953	0.0404	-0.0978
	(3.59)	(3.23)	(2.96)	(4.07)	(0.10)	(1.84)	(0.62)	(0.14)	(0.66)	(-2.50)
Lag of Stroke	0.375	0.355	-0.00822	0.185	0.0129	0.172	0.351	-0.0784	-0.0642	-0.264
	(8.56)	(7.63)	(-0.07)	(1.52)	(0.06)	(1.73)	(4.87)	(-0.95)	(-0.62)	(-3.37)
Lag of Cancer	0.0983	0.0466	0.287	0.206	-0.193	0.194	-0.163	0.1000	0.0200	-0.0412
	(2.37)	(0.94)	(2.78)	(1.81)	(-1.57)	(1.88)	(-1.41)	(0.90)	(0.25)	(-0.87)
Lag of Hypertension	0.0142	0.0375	0.0609	0.0371	-0.122	-0.0235	0.107	0.00513	0.0480	-0.0300
	(0.47)	(1.05)	(1.25)	(0.51)	(-1.58)	(-0.31)	(1.53)	(0.08)	(1.01)	(-1.00)
Lag of Diabetes	0.124	0.103	0.00728	0.161	-0.0451	0.115	0.201	0.0579	0.0276	-0.0669
	(3.29)	(2.28)	(0.11)	(1.78)	(-0.46)	(1.35)	(2.23)	(0.75)	(0.42)	(-1.56)
Lag of Lung disease	0.262	0.0353	0.345	0.00980	-0.0521	-0.0238	-0.197	0.121	0.0770	-0.139
	(5.83)	(0.66)	(3.43)	(0.09)	(-0.37)	(-0.24)	(-1.61)	(1.27)	(0.92)	(-2.38)
Lag of Has exactly 1 IADL	0.224	0.926	0.00224	0.0935	0.0141	0.122	0.382	-0.195	-0.0258	-0.0692
	(7.20)	(31.74)	(0.04)	(1.33)	(0.11)	(1.89)	(6.23)	(-3.67)	(-0.40)	(-1.62)
Lag of Has 2 or more IADLs	0.628	1.762	0.255	-0.0860	0.262	0.0277	0.907	-0.0984	-0.0814	-0.133
	(11.52)	(33.70)	(1.94)	(-0.61)	(0.74)	(0.26)	(12.88)	(-1.22)	(-0.57)	(-1.08)
Lag of Has exactly 1 ADL	0.885	0.259	0.0460	0.373	0.242	0.143	0.214	-0.122	0.0143	-0.185
	(37.31)	(8.56)	(0.93)	(6.27)	(2.41)	(2.43)	(3.57)	(-2.37)	(0.26)	(-4.77)
Lag of Has exactly 2 ADLs	1.334	0.400	0.0483	0.326	0.0691	-0.0983	0.458	-0.164	-0.118	-0.264
	(33.05)	(8.89)	(0.57)	(3.37)	(0.29)	(-1.06)	(6.11)	(-2.26)	(-1.19)	(-3.34)
Lag of Has 3 or more ADLs	1.772	0.545	0.327	0.350	0.634	-0.0220	0.460	-0.109	-0.120	-0.540
	(35.92)	(11.82)	(3.24)	(3.61)	(2.59)	(-0.24)	(6.09)	(-1.54)	(-1.21)	(-5.19)
Lag of Current smoking	0.0192	0.0854								
	(0.53)	(1.98)								
Lag of Widowed	0.0330	-0.0368	-0.142	0.486	0.0360	0.0289	0.235	-0.353	-0.0631	0.0694
	(0.89)	(-0.85)	(-1.52)	(3.93)	(0.24)	(0.25)	(3.21)	(-4.90)	(-0.63)	(1.33)
Init. of Heart disease	-0.0294	-0.0400	-0.0600	-0.212	0.0814	-0.0521	-0.0789	0.0923	0.0443	-0.0397
	(-0.84)	(-1.01)	(-0.74)	(-2.42)	(0.71)	(-0.62)	(-1.04)	(1.21)	(0.61)	(-0.85)
Init. of Stroke	-0.189	-0.0642	-0.142	-0.262	-0.350	-0.230	-0.206	0.0415	-0.0761	0.0728
	(-3.39)	(-1.11)	(-1.03)	(-1.69)	(-1.18)	(-1.79)	(-2.19)	(0.41)	(-0.56)	(0.73)
Init. of Cancer	-0.0454	-0.0264	-0.139	-0.204	0.273	-0.207	0.0408	0.0667	-0.0685	-0.00616
	(-0.92)	(-0.45)	(-1.18)	(-1.45)	(1.79)	(-1.56)	(0.31)	(0.51)	(-0.68)	(-0.10)
Init. of Hypertension	0.0645	0.00908	-0.0729	0.0761	0.232	0.0971	-0.110	0.0233	-0.0357	-0.0166
	(2.13)	(0.25)	(-1.44)	(1.02)	(2.86)	(1.26)	(-1.57)	(0.37)	(-0.71)	(-0.52)
Init. of Diabetes	0.0337	0.0647	0.142	-0.0812	0.125	-0.101	0.103	-0.0696	-0.0233	0.0195
	(0.79)	(1.29)	(1.86)	(-0.80)	(1.01)	(-1.06)	(1.03)	(-0.82)	(-0.30)	(0.38)
Init. of Lung disease	-0.0285	0.0479	-0.445	0.0178	-0.151	0.156	0.129	-0.0743	-0.124	0.00299
	(-0.55)	(0.77)	(-3.99)	(0.14)	(-0.80)	(1.37)	(0.92)	(-0.67)	(-1.22)	(0.04)
Init. of Ever smoked	0.0149	0.0180	0.123	0.0615	0.0913	-0.0459	-0.0257	0.00990	0.0408	-0.0272
	(0.82)	(0.85)	(4.41)	(1.34)	(2.06)	(-0.95)	(-0.57)	(0.25)	(1.32)	(-1.52)
Init. of Current smoking	0.116	-0.0162	-0.229	0.113	-0.0191	0.152	0.0836	-0.125	0.0213	-0.0519
	(3.41)	(-0.39)	(-7.96)	(2.53)	(-0.36)	(3.19)	(1.36)	(-2.89)	(0.62)	(-2.53)
Init. of Has exactly 1 IADL	0.0661	0.102	0.113	0.00620	-0.0396	0.0405	0.0952	-0.0220	0.0483	0.0638
	(2.32)	(3.32)	(2.74)	(0.10)	(-0.43)	(0.72)	(1.35)	(-0.44)	(1.03)	(2.05)
Init. of Has 2 or more IADLs	-0.0725	0.251	0.382	-0.400		0.115	0.0715	-0.283	0.0311	-0.876
	(-0.66)	(2.34)	(1.45)	(-1.32)		(0.44)	(0.54)	(-2.07)	(0.05)	(-1.54)
Init. of Has exactly 1 ADL	0.358	0.0615	0.0966	0.105	0.126	0.0965	-0.0357	0.0997	-0.107	-0.0799
	(14.18)	(1.98)	(1.97)	(1.80)	(1.13)	(1.75)	(-0.55)	(1.91)	(-1.97)	(-2.06)
Init. of Has exactly 2 ADLs	0.298	-0.0263	-0.255	-0.397	0.736	-0.0420	-0.00442	0.00630	0.553	-0.348
	(4.25)	(-0.35)	(-1.37)	(-1.57)	(1.04)	(-0.20)	(-0.04)	(0.06)	(1.52)	(-1.68)
Init. of Has 3 or more ADLs	0.446	-0.00565	0.297	0.182		-0.109	-0.222	-0.00103	-0.649	-0.262
	(4.96)	(-0.07)	(1.34)	(0.84)		(-0.55)	(-1.82)	(-0.01)	(-0.90)	(-1.04)
Init. of Widowed	0.00252	0.0261	-0.0415	-0.518	-0.0590	0.0753	0.0185	0.0228	0.213	-0.0475
	(0.06)	(0.56)	(-0.40)	(-3.60)	(-0.33)	(0.59)	(0.25)	(0.31)	(1.88)	(-0.77)
Init. of Single	0.0413	0.0399	-0.194	0.00871	0.0469	0.158	0.304	-0.390	-0.127	-0.0537
	(1.67)	(1.38)	(-6.16)	(0.17)	(0.81)	(3.09)	(4.63)	(-9.04)	(-3.23)	(-2.24)
Init. of R working for pay	-0.0762	-0.0318	-0.273	0.0218		0.0957	-0.280	-0.345	0.262	0.452
	(-1.51)	(-0.54)	(-3.79)	(0.17)		(0.72)	(-1.64)	(-3.05)	(2.90)	(9.06)
Init. of (IHT of earnings in 1000s)/100 if working zero otherwise	-1.854	-2.613	3.536	1.645	5.098	-5.394	1.874	8.822	-11.47	-1.527
	(-1.34)	(-1.59)	(1.84)	(0.49)	(1.35)	(-1.44)	(0.37)	(2.50)	(-5.08)	(-1.21)
Init. of Non-pension wlth(hatota) not zero	0.0641	0.0637	-0.129	0.299	-0.196	-0.00840	0.0531	0.481	-0.0957	0.0211
	(1.46)	(1.33)	(-1.68)	(2.58)	(-0.79)	(-0.09)	(0.52)	(9.67)	(-0.96)	(0.28)
Init. of (IHT of hh wlth in 1000s if positive)/100 zero otherwise	-3.185	-2.574	1.616	-2.300	-0.154	-2.131	-2.150	7.328	0.687	-2.104
	(-8.31)	(-5.81)	(2.49)	(-2.39)	(-0.11)	(-2.08)	(-1.69)	(9.45)	(0.82)	(-4.01)

Table 11. Transition of Economic Outcomes (coefficient/t statistic)

AIME in ini.intw (-9=no match)	-0.289	-0.780	2.018	-0.144	-0.417	-2.315	-0.0757	1.122	-0.393	0.934
	(-1.72)	(-4.12)	(7.38)	(-0.30)	(-0.76)	(-5.25)	(-0.21)	(3.49)	(-1.17)	(4.92)
Quarters of earnings in ini.intw (-9=no match)	0.218	0.718	-2.199	2.455	-0.431	2.986	-0.156	-0.794	4.210	0.0114
	(0.88)	(2.56)	(-5.47)	(3.30)	(-0.49)	(4.44)	(-0.29)	(-1.70)	(8.29)	(0.04)
Init. of Health fair/poor	0.324	0.180	-0.00796	0.315	0.0607	0.191	0.0412	-0.106	-0.0347	-0.168
	(16.90)	(7.98)	(-0.25)	(6.88)	(0.83)	(4.14)	(0.86)	(-2.78)	(-0.84)	(-6.68)
Init. of Any DB from current job RND VG	-0.0433	-0.0558	0.498	-0.139		-0.193	0.0509	0.132	0.137	-0.0220
	(-0.73)	(-0.75)	(6.36)	(-1.13)		(-1.01)	(0.33)	(0.72)	(1.93)	(-0.57)
Init. of Normal DB Retirement Age 60-61	0.0640	0.0878	0.203	0.0774	-0.191	0.206	-0.0126	0.168	-0.156	0.0688
	(0.92)	(1.01)	(2.01)	(0.54)	(-3.19)	(0.98)	(-0.05)	(0.76)	(-1.93)	(1.52)
Init. of Normal DB Retirement Age 62-64	0.0291	0.167	0.0147	0.252	-0.213	0.556	-0.303	-0.0967	0.00548	0.0290
	(0.39)	(1.84)	(0.14)	(1.74)	(-3.23)	(2.73)	(-0.86)	(-0.46)	(0.06)	(0.58)
Init. of Normal DB Retirement Age 65+	0.0678	0.0533	-0.0356	0.231	-0.249	0.307		0.286	-0.0670	0.0629
	(1.04)	(0.65)	(-0.41)	(1.77)	(-4.33)	(1.54)		(1.35)	(-0.87)	(1.48)
Init. of Any DC from current job RND VG	0.0486	0.0662	0.151	0.233	-0.196	0.208	0.490	-0.248	-0.0371	0.238
	(0.87)	(0.98)	(2.29)	(2.39)	(-2.21)	(1.76)	(2.24)	(-1.66)	(-0.53)	(6.07)
Init. of (IHT of DC wlth in 1000s)/100 if any DC zero otherwise	-2.886	-1.676	7.781	-6.461	4.488	-1.151	-15.62	19.43	2.390	-4.690
	(-2.00)	(-0.96)	(3.99)	(-2.44)	(2.12)	(-0.35)	(-2.16)	(3.04)	(1.41)	(-4.89)
Min(log(Init of BMI), log(30))	-0.259	-0.914								
	(-2.10)	(-6.63)								
Max(log(Init of BMI) - log(30), 0)	1.183	-0.111								
	(6.79)	(-0.49)								
Init. Of Min(log(Init of BMI), log(30))	0.594	0.447								
	(4.61)	(3.09)								
Init of Max(log(Init of BMI) - log(30), 0)	-0.00928	0.175								
	(-0.05)	(0.76)								
Log(Time Between Interviews)	0.158	0.205	-0.0619	0.0564	0.569	0.0777	1.098	-0.139	1.101	-0.385
	(2.46)	(2.73)	(-0.66)	(0.37)	(3.38)	(0.49)	(6.57)	(-1.01)	(9.79)	(-5.98)
Lag of R working for pay			-0.380	-0.0461		-0.105		0.214	0.366	1.330
			(-5.16)	(-0.33)		(-0.66)		(1.45)	(4.02)	(28.63)
Lag of (IHT of earnings in 1000s)/100 if working zero otherwise			15.99	-11.18	18.02	-5.589		0.372	-17.73	5.007
			(8.25)	(-3.16)	(11.43)	(-1.30)		(0.09)	(-7.97)	(4.22)
Lag of Non-pension wlth(hatota) not zero			-0.222	-0.307	0.124	-0.103	0.0666	0.940	-0.167	0.326
			(-2.90)	(-2.79)	(0.38)	(-1.19)	(0.70)	(19.80)	(-1.66)	(4.10)
Lag of (IHT of hh wlth in 1000s if positive)/100 zero otherwise			6.392	-2.161	3.840	-2.030	-4.526	11.39	1.894	-4.004
			(9.87)	(-2.18)	(2.45)	(-1.94)	(-3.71)	(14.75)	(2.20)	(-7.48)
Lag of Claiming SSDI			1.075	2.986	-0.141	-0.119		0.0222	-1.023	-0.812
			(16.14)	(56.25)	(-0.61)	(-1.85)		(0.35)	(-18.40)	(-12.41)
Lag of Claiming OASI			-0.0878		-0.0186	-0.294		0.163		-0.530
			(-1.86)		(-0.24)	(-4.01)		(2.27)		(-21.39)
Lag of Claiming SSI						3.022				
						(53.70)				
Lag of Claiming DB						-0.181		-0.285	0.146	-0.185
						(-1.24)		(-1.80)	(2.07)	(-4.48)
Lag of R live in nursingh ome at interview							2.059	-0.562		
							(19.94)	(-5.31)		
Spline Lag of Age Knot at 58								0.00411		
								(0.46)		
Splined Lag of Age Knot at 73								-0.0155		
								(-2.33)		
Splined Lag of Age > 73								-0.0131		
								(-3.02)		
Dummy for Wave 3								-0.0276		
								(-0.42)		
Dummy of Wave 4								-0.140		
								(-2.07)		
Dummy of Wave 5								0.000346		
								(0.01)		
Dummy for Wave 6								-0.0733		
								(-1.02)		
Dummy for Wave 7								-0.0560		
								(-0.76)		
Lag of Age in [58, 59]									1.624	-0.171
									(19.11)	(-7.10)
Lag of Age == 60									3.152	-0.457
									(37.20)	(-14.37)
Lag of Age == 61									3.283	-0.477
									(38.50)	(-15.11)
Lag of Age == 62									2.984	-0.211
									(33.10)	(-5.90)
Lag of Age in [63, 64]									4.369	-0.264
									(47.88)	(-8.83)
Lag of Age in [65, 68]									3.367	-0.0929
									(30.56)	(-3.11)
Constant	-3.355	0.548	0.274	-2.792	-10.02	-1.366	-6.446	0.130	-4.910	-1.318
	(-12.04)	(1.76)	(1.36)	(-8.29)	(-17.60)	(-3.73)	(-7.31)	(0.26)	(-29.33)	(-13.57)

t statistics in parentheses

Table 12. OLS regression for Log(BMI) (coefficient/t statistic)

	Log(BMI)
Non-Hispanic black	-0.000502 (-0.52)
Hispanic	0.000617 (0.49)
Less than high school	-0.000583 (-0.72)
Some college and above	-0.00180 (-2.56)
Male	-0.00151 (-2.03)
Min(Lag Age, 63)	0.0000764 (0.64)
Min(Max(Lag Age - 63, 0), 73-63)	-0.000794 (-5.40)
Max(Lag Age - 73, 0)	-0.00109 (-7.30)
Lag of Heart disease	-0.000527 (-0.39)
Lag of Stroke	-0.00369 (-1.79)
Lag of Cancer	0.00146 (0.85)
Lag of Hypertension	0.00364 (3.11)
Lag of Diabetes	-0.00201 (-1.24)
Lag of Lung disease	-0.00171 (-0.86)
Lag of Has exactly 1 IADL	-0.000942 (-0.66)
Lag of Has 2 or more IADLs	-0.00687 (-2.75)
Lag of Has exactly 1 ADL	-0.000758 (-0.61)
Lag of Has exactly 2 ADLs	-0.000351 (-0.17)
Lag of Has 3 or more ADLs	0.00190

Table 12. OLS regression for Log(BMI) (coefficient/t statistic)

	(0.87)
Lag of Current smoking	-0.0118
	(-8.31)
Lag of Widowed	-0.000714
	(-0.45)
Init. of Heart disease	-0.000807
	(-0.53)
Init. of Stroke	-0.00107
	(-0.41)
Init. of Cancer	-0.00316
	(-1.54)
Init. of Hypertension	-0.00349
	(-2.88)
Init. of Diabetes	0.000617
	(0.33)
Init. of Lung disease	-0.00233
	(-0.99)
Init. of Ever smoked	0.00110
	(1.60)
Init. of Current smoking	0.00568
	(4.18)
Init. of Has exactly 1 IADL	-0.000720
	(-0.60)
Init. of Has 2 or more IADLs	0.00286
	(0.56)
Init. of Has exactly 1 ADL	-0.0000598
	(-0.05)
Init. of Has exactly 2 ADLs	-0.00476
	(-1.37)
Init. of Has 3 or more ADLs	-0.000269
	(-0.07)
Init. of Widowed	0.00121
	(0.69)
Init. of Single	0.000173
	(0.18)
Init. of R working for pay	0.00358
	(1.92)
Init. of (IHT of earnings in 1000s)/100 if working zero otherwise	-0.0746

Table 12. OLS regression for Log(BMI) (coefficient/t statistic)

	(-1.53)
Init. of Non-pension wlth(hatota) not zero	0.00459 (2.28)
Init. of (IHT of hh wlth in 1000s if positive)/100 zero otherwise	0.0129 (0.81)
AIME in ini.intw (-9=no match)	-0.00116 (-0.17)
Quarters of earnings in ini.intw (-9=no match)	0.00923 (0.90)
Init. of Health fair/poor	-0.00301 (-3.48)
Init. of Any DB from current job RND VG	0.0000519 (0.03)
Init. of Normal DB Retirement Age 60-61	0.00173 (0.82)
Init. of Normal DB Retirement Age 62-64	0.00298 (1.29)
Init. of Normal DB Retirement Age 65+	-0.000836 (-0.42)
Init. of Any DC from current job RND VG	0.00169 (0.94)
Init. of (IHT of DC wlth in 1000s)/100 if any DC zero otherwise	-0.0492 (-1.12)
Min(log(Init of BMI), log(30))	0.787 (158.15)
Max(log(Init of BMI) - log(30), 0)	0.789 (102.94)
Init. Of Min(log(Init of BMI), log(30))	0.177 (34.76)
Init of Max(log(Init of BMI) - log(30), 0)	0.153 (19.04)
Log(Time Between Interviews)	-0.0106 (-4.30)
Birth Year	0.000649 (6.23)
Constant	-1.135 (-5.50)

Table 13. Ordered Probits for Transition of Smoking, and Functional Status (coefficient/t statistic)

	Number of ADLs if any	Number of IADLs if any	Smoking Status
main			
Non-Hispanic black	0.14	0.071	-0.0464
	-3.77	-1.26	(-1.45)
Hispanic	0.0763	0.229	0.00936
	-1.56	-3.14	-0.22
Less than high school	0.0116	-0.01	0.00258
	-0.36	(-0.20)	-0.1
Some college and above	0.037	0.15	-0.0668
	-1.03	-2.62	(-2.74)
Male	0.00633	-0.00131	0.0501
	-0.19	(-0.03)	-2.01
Min(Lag Age, 63)	-0.0058	0.0172	-0.0172
	(-1.14)	-2.16	(-5.77)
Min(Max(Lag Age - 63, 0), 73-63)	0.00267	0.0188	-0.0246
	-0.51	-2.31	(-5.94)
Max(Lag Age - 73, 0)	0.0269	0.0157	-0.0153
	-7.52	-3.28	(-3.15)
Lag of Heart disease	0.132	-0.0155	0.0879
	-2.76	(-0.22)	-1.94
Lag of Stroke	0.205	0.368	0.0816
	-3.69	-4.86	-1.19
Lag of Cancer	-0.0747	-0.0614	-0.059
	(-1.05)	(-0.57)	(-0.98)
Lag of Hypertension	-0.0342	-0.03	-0.0657
	(-0.67)	(-0.39)	(-1.64)
Lag of Diabetes	0.0538	-0.0892	-0.0955
	-0.93	(-0.94)	(-1.71)
Lag of Lung disease	0.0545	0.0672	-0.0532
	-0.86	-0.63	(-0.95)
Lag of Has exactly 1 IADL	0.195	0.212	0.0172
	-4.78	-4.18	-0.36
Lag of Has 2 or more IADLs	0.556	0.994	-0.116
	-10.18	-15.97	(-1.33)
Lag of Has exactly 1 ADL	0.214	0.114	0.0132
	-6.32	-1.95	-0.32
Lag of Has exactly 2 ADLs	0.477	0.0539	-0.0362
	-10.83	-0.72	(-0.52)
Lag of Has 3 or more ADLs	1.044	0.181	-0.105

Table 13. Ordered Probits for Transition of Smoking, and Functional Status (coefficient/t statistic)

	-23.09	-2.59	(-1.42)
Lag of Current smoking	-0.0955	-0.117	2.031
	(-1.62)	(-1.22)	-66.63
Lag of Widowed	-0.0459	0.0461	0.102
	(-0.79)	-0.54	-1.82
Init. of Heart disease	-0.149	-0.0043	-0.0306
	(-2.83)	(-0.06)	(-0.59)
Init. of Stroke	0.000142	-0.356	-0.198
	0	(-3.75)	(-2.25)
Init. of Cancer	0.0579	0.0499	0.0685
	-0.69	-0.4	-0.95
Init. of Hypertension	0.0307	-0.0107	-0.0551
	-0.6	(-0.14)	(-1.31)
Init. of Diabetes	-0.00104	0.156	0.103
	(-0.02)	-1.51	-1.61
Init. of Lung disease	-0.0489	-0.146	0.0886
	(-0.67)	(-1.18)	-1.33
Init. of Ever smoked	-0.0188	-0.0528	5.161
	(-0.59)	(-1.12)	-113.17
Init. of Current smoking	0.0878	0.0967	0.914
	-1.56	-1.06	-28.83
Init. of Has exactly 1 IADL	0.0133	0.13	0.0335
	-0.32	-2.26	-0.83
Init. of Has 2 or more IADLs	-0.0681	0.183	-0.164
	(-0.61)	-1.51	(-0.88)
Init. of Has exactly 1 ADL	0.168	0.123	0.0351
	-4.88	-2.07	-0.85
Init. of Has exactly 2 ADLs	0.187	0.102	0.171
	-2.57	-0.86	-1.33
Init. of Has 3 or more ADLs	0.206	-0.0175	0.11
	-2.74	(-0.15)	-0.76
Init. of Widowed	0.0248	-0.0492	-0.0487
	-0.41	(-0.55)	(-0.77)
Init. of Single	-0.0000932	-0.0253	0.0267
	(-0.00)	(-0.38)	-0.86
Init. of R working for pay	-0.0949	-0.404	0.0283
	(-0.95)	(-2.61)	-0.44
Init. of (IHT of earnings in 1000s)/100 if working zero otherwise	0.606	4.758	-1.989
	-0.21	-1.05	(-1.19)

Table 13. Ordered Probits for Transition of Smoking, and Functional Status (coefficient/t statistic)

Init. of Non-pension wlth(hatota) not zero	-0.0289	0.0616	-0.0432
	(-0.48)	-0.69	(-0.69)
Init. of (IHT of hh wlth in 1000s if positive)/100 zero otherwise	-1.429	-0.834	-2.091
	(-2.34)	(-0.84)	(-4.16)
AIME in ini.intw (-9=no match)	-0.076	-0.423	-0.017
	(-0.29)	(-1.10)	(-0.07)
Quarters of earnings in ini.intw (-9=no match)	-0.00647	0.853	0.138
	(-0.02)	-1.51	-0.38
Init. of Health fair/poor	0.133	0.0619	0.0167
	-4.35	-1.34	-0.58
Init. of Any DB from current job RND VG	-0.511	-0.568	-0.0138
	(-3.02)	(-1.62)	(-0.22)
Init. of Normal DB Retirement Age 60-61	0.39	0.547	-0.081
	-2.09	-1.44	(-1.10)
Init. of Normal DB Retirement Age 62-64	0.517	0.936	-0.0416
	-2.61	-2.46	(-0.53)
Init. of Normal DB Retirement Age 65+	0.411	0.76	-0.135
	-2.29	-2.08	(-1.96)
Init. of Any DC from current job RND VG	0.226	0.0607	0.00705
	-1.82	-0.27	-0.12
Init. of (IHT of DC wlth in 1000s)/100 if any DC zero otherwise	-5.392	-4.062	-0.283
	(-1.57)	(-0.64)	(-0.19)
Min(log(Init of BMI), log(30))	-0.404	-0.702	-0.896
	(-2.23)	(-2.94)	(-5.55)
Max(log(Init of BMI) - log(30), 0)	0.109	-0.594	-0.56
	-0.44	(-1.26)	(-2.10)
Init. Of Min(log(Init of BMI), log(30))	0.0737	0.186	0.539
	-0.38	-0.72	-3.25
Init of Max(log(Init of BMI) - log(30), 0)	0.0263	0.6	0.171
	-0.1	-1.26	-0.63
Log(Time Between Interviews)	0.438	0.482	-0.0563
	-4.07	-2.98	(-0.67)
cut1			
Constant	-0.651	0.688	-0.00471
	(-1.34)	-0.97	(-0.01)
cut2			
Constant	0.027		4.661
	-0.06		-13.52

t statistics in parentheses

Table 14. Estimating Earnings and Household Wealth Using Inverse Hyperbolic Sine Transformation Preserving Shape

	Household Wealth if nonzero	Individual earnings if working
Non-Hispanic black	-4.026	0.0473
	(-18.21)	(0.52)
Hispanic	-5.329	-0.385
	(-18.44)	(-3.17)
Less than high school	-2.213	-0.145
	(-11.93)	(-1.62)
Some college and above	4.170	0.893
	(26.41)	(13.88)
Male	-0.804	0.817
	(-4.78)	(11.64)
Spline Lag of Age Knot at 58	0.357	
	(10.99)	
Splined Lag of Age Knot at 73	0.116	
	(4.20)	
Splined Lag of Age > 73	-0.00494	
	(-0.19)	
Lag of Heart disease	-1.505	-0.791
	(-4.94)	(-4.75)
Lag of Stroke	-2.259	-0.764
	(-4.75)	(-2.00)
Lag of Cancer	0.695	-0.295
	(1.81)	(-1.46)
Lag of Hypertension	-0.597	0.0493
	(-2.25)	(0.42)
Lag of Diabetes	-1.085	-0.0777
	(-2.97)	(-0.43)
Lag of Lung disease	-0.716	-0.620
	(-1.58)	(-2.43)
Lag of Has exactly 1 IADL	-0.480	0.0874
	(-1.45)	(0.52)
Lag of Has 2 or more IADLs	-1.767	-1.436
	(-2.92)	(-2.20)
Lag of Has exactly 1 ADL	-1.549	-0.350
	(-5.42)	(-2.03)

Table 14. Estimating Earnings and Household Wealth Using Inverse Hyperbolic Sine Transformation Preserving Shape

Lag of Has exactly 2 ADLs	-1.215	-0.463
	(-2.51)	(-1.18)
Lag of Has 3 or more ADLs	-0.873	-0.842
	(-1.66)	(-1.38)
Lag of Widowed	-4.214	0.468
	(-11.71)	(2.20)
Lag of R working for pay	-2.424	-7.236
	(-4.88)	(-37.11)
Lag of (IHT of earnings in 1000s)/100 if working zero otherwise	15.49	259.1
	(1.21)	(56.14)
Lag of Non-pension wlth(hatota) not zero	-7.567	-0.444
	(-10.63)	(-1.27)
Lag of (IHT of hh wlth in 1000s if positive)/100 zero otherwise	617.5	7.633
	(126.94)	(3.82)
Lag of Claiming SSDI	-1.773	-2.209
	(-4.56)	(-4.61)
Lag of Claiming OASI	-0.742	-2.529
	(-2.63)	(-19.68)
Lag of Claiming DB	-0.887	-2.821
	(-2.17)	(-14.39)
R live in nursingh ome at interview	-3.451	
	(-3.02)	
Init. of Heart disease	1.249	0.621
	(3.59)	(3.12)
Init. of Stroke	1.106	0.942
	(1.85)	(2.00)
Init. of Cancer	0.134	0.302
	(0.29)	(1.24)
Init. of Hypertension	-0.654	-0.119
	(-2.38)	(-0.96)
Init. of Diabetes	-0.455	-0.0445
	(-1.07)	(-0.21)
Init. of Lung disease	-0.768	0.485
	(-1.43)	(1.58)

Table 14. Estimating Earnings and Household Wealth Using Inverse Hyperbolic Sine Transformation Preserving Shape

Init. of Ever smoked	-0.200	-0.0189
	(-1.29)	(-0.29)
Init. of Current smoking	-1.948	-0.0871
	(-10.47)	(-1.15)
Init. of Has exactly 1 IADL	-0.218	-0.451
	(-0.79)	(-3.68)
Init. of Has 2 or more IADLs	-1.648	2.182
	(-1.31)	(0.56)
Init. of Has exactly 1 ADL	0.412	0.0329
	(1.43)	(0.19)
Init. of Has exactly 2 ADLs	-1.025	0.515
	(-1.24)	(0.53)
Init. of Has 3 or more ADLs	-0.000672	0.444
	(-0.00)	(0.30)
Init. of Widowed	-0.625	-0.356
	(-1.55)	(-1.42)
Init. of Single	-3.829	0.225
	(-17.08)	(2.59)
Init. of R working for pay	-2.335	-2.929
	(-4.99)	(-13.93)
Init. of (IHT of earnings in 1000s)/100 if working zero otherwise	65.59	99.12
	(5.37)	(19.74)
Init. of Non-pension wlth(hatota) not zero	-4.625	0.415
	(-6.85)	(1.28)
Init. of (IHT of hh wlth in 1000s if positive)/100 zero otherwise	270.6	-1.347
	(56.35)	(-0.70)
AIME in ini.intw (-9=no match)	23.57	17.02
	(14.94)	(21.79)
Quarters of earnings in ini.intw (-9=no match)	-32.61	-22.23
	(-13.83)	(-18.40)
Init. of Health fair/poor	-1.509	-0.0493
	(-7.57)	(-0.48)
Init. of Any DB from current job RND VG	0.167	0.655
	(0.41)	(5.20)
Init. of Normal DB Retirement Age 60-61	-0.502	-0.191

Table 14. Estimating Earnings and Household Wealth Using Inverse Hyperbolic Sine Transformation Preserving Shape

	(-1.06)	(-1.30)
Init. of Normal DB Retirement Age 62-64	-0.423	-0.0855
	(-0.82)	(-0.53)
Init. of Normal DB Retirement Age 65+	-0.706	-0.413
	(-1.59)	(-2.98)
Init. of Any DC from current job RND VG	-3.760	-0.716
	(-9.29)	(-5.84)
Init. of (IHT of DC wlth in 1000s)/100 if any DC zero otherwise	145.7	28.62
	(14.72)	(9.33)
Dummy for Wave 3	-0.0418	
	(-0.17)	
Dummy of Wave 4	-0.539	
	(-2.05)	
Dummy of Wave 5	0.163	
	(0.64)	
Dummy for Wave 6	-0.281	
	(-1.00)	
Dummy for Wave 7	0.183	
	(0.64)	
Log(Time Between Interviews)	0.658	0.119
	(1.10)	(0.50)
Min(Lag Age, 63)		0.00660
		(0.83)
Min(Max(Lag Age - 63, 0), 73-63)		0.141
		(5.59)
Constant	-4.933	5.183
	(-2.47)	(8.51)

t statistics in parentheses

Table 15. Comparison of Predicted and Observed Outcomes for HRS Cohort in 2004

	1992 Observed	2004 Observed	2004 Simulated
Survival	100%	88%	97%
Cancer Prevalence	5%	16%	15%
Diabetes Prevalence	10%	18%	19%
Heart Disease Prevalence	12%	29%	25%
Hypertension Prevalence	35%	56%	57%
Lung Disease Prevalence	6%	10%	13%
Stroke Prevalence	3%	11%	9%
Any Condition Prevalence	49%	80%	78%
3+ Conditions Prevalence	5%	32%	15%
Any IADLs Prevalence	12%	14%	10%
Any ADLs Prevalence	12%	21%	9%
Overweight Prevalence	41%	31%	39%
Obesity Prevalence	22%	36%	27%
Ever Smoked Prevalence	64%	57%	61%
Current Smoking Prevalence	29%	11%	15%
Working Prevalence	62%	20%	23%
OASI Claiming	5%	75%	82%
SSDI Claiming	8%	2%	3%
SSI Claiming	7%	2%	4%
Mean Earnings (thousands)	$ 14.90	$ 7.70	$ 7.65
Median HH wealth (thousands)	$ 131.93	$ 153.00	$ 213.96

Table 16. Descriptive Statistics Initial Conditions for Estimation (1992) and Simulation (2004)

			1992	2004	Selection
Binary		working for pay	70%	74%	all
		non-zero wealth	96%	98%	all
		hypertension	32%	31%	all
		heate disease	10%	8%	all
		diabetes	8%	9%	all
		any health insurance	85%	87%	all
		SRH fair or poor	20%	14%	all
Ordered	BMI Status	normal	36%	30%	all
		overweight	42%	35%	all
		obese	23%	36%	all
	Smoking Status	never smoked	34%	41%	all
		former smoker	34%	34%	all
		current smoker	32%	25%	all
	Functional Status	no ADL	91%	93%	all
		no IADL	91%	95%	all
Continuous		log aime (nomal $USD)	0.7260021	0.7267455	all
		log quarters of coverage	0.4326305	0.4345214	all
Censored Continuous		earnings	$ 40,614.92	$ 44,242.06	if working
		wealth	$ 232,606.50	$ 292,721.71	if non-zero
		dc wealth	$ 15.76	$ 53.71	if dc plan
Censored Discrete		any db plan	28%	33%	if working
		any dc plan	24%	32%	if working
Censored Ordered	Early Age Eligible DB	<52	18%	24%	all
		52-57	58%	54%	all
		58>	24%	22%	all
	Normal Age Eligible DB	<57	18%	21%	all
		57-61	25%	25%	all
		62-63	18%	17%	all
		64>	38%	37%	all
Covariates		hispanic	6%	9%	all
		black	12%	11%	all
		male	48%	49%	all
		less high school	23%	9%	all
		college	36%	63%	all
		single	20%	26%	all
		widowed	4%	2%	all
		cancer	5%	5%	all
		lunge disease	5%	4%	all
		stroke	2%	2%	all

Table 17. Parameter Estimates for Latent Model (Conditional Mean and Thresholds)

covariate	Hypertension	Heart Disease	Diabetes	Any Health Insurance	Self-repoted Health	Weight Status	Smoking Status	Function Status	Working	Nonzero Wealth
black	0.533804	0.0266325	0.383543	-0.104687	0.55494	0.335153	-0.0877226	0.315922	-0.0927098	-1.06499
hispan	0.003414	-0.163198	0.308697	-0.661202	0.425825	0.196995	-0.296138	0.260283	-0.134074	-0.970737
hsless	0.0917565	0.162802	0.233981	-0.503555	0.512868	0.119763	0.325363	0.275916	-0.318622	-0.277265
college	-0.0438145	-0.0654016	-0.0461253	0.184565	-0.373108	-0.138809	-0.110955	-0.290062	0.244635	0.632133
male	0.0980248	0.272602	0.0743117	-0.0221166	0.0118689	0.101087	0.416308	-0.084304	0.418108	-0.08367
single	0.179446	-0.0215096	0.0941147	-0.252689	0.197503	-0.0339399	0.251647	0.0590587	0.0743597	-0.974693
widowed	0.153526	0.0165263	0.0971176	-0.41429	0.345621	0.19888	0.269208	-0.0243994	0.178656	-0.986268
Lung Disease	0.225398	0.682802	0.340743	-0.190175	1.00991	0.041741	0.594646	0.669016	-0.403559	0.185724
Cancer	0.0002446	0.327371	0.110757	0.467964	0.665257	-0.0621035	0.178155	0.351412	-0.271444	0.0777144
Stroke	0.976346	1.00614	0.504945	-0.0515253	1.1302	0.1271	0.176914	0.864126	-0.94063	-0.58408
constant	-0.700217	-1.56425	-1.71395	1.35249	-1.251	0.300168	0.0973724	-1.05315	0.418126	2.85139

Table 17. Parameter Estimates for Latent Model (Conditional Mean and Thresholds)

Log(AIME)	Log(Quarters Worked)	IHT(HH Wealth)	IHT(Earned Income)	Log(DC Wealth)	Any DC Plan	Any DB Plan	Early Retirement Age	Normal Retirement Age	
-0.019987	-0.0063732	-0.0731182	-0.0216715	-0.0724877	-0.0029314	-0.5851122	-16.84629	-0.0238185	black
-0.0471219	-0.0273668	-0.209397	-0.224873	-0.0122119	-0.148202	-2.166444	-15.66683	-0.0244621	hispan
-0.0450942	-0.0210904	-0.292743	-0.294862	0.272853	-0.0578842	-1.911611	-12.48751	-0.0475891	hsless
0.0095202	-0.005404	0.273752	0.0591213	-0.207269	-0.349554	3.183553	11.22917	0.0698589	college
0.127444	0.0641736	0.178004	0.0320516	-0.141812	0.0265351	4.748151	-2.617562	0.0990174	male
0.0128859	0.0091106	0.0721962	-0.0236959	-0.0147138	0.081574	0.4953211	-23.02886	-0.0151746	single
0.0004597	0.0030366	0.0782894	-0.132053	0.168107	0.121412	0.0305638	-17.41909	-0.0699797	widowed
-0.0146736	-0.00287	-0.209386	0.0113309	0.0630776	-0.025893	-1.422953	-15.52895	-0.0335796	lunge
0.0105398	0.0062261	0.0603327	-0.166115	0.0572551	-0.145513	1.209624	0.7208427	-0.0386802	cancre
-0.0297451	-0.0167381	-0.154346	-0.286447	0.176488	-0.160052	-0.6967615	-12.55789	0.1117583	stroke
0.668169	0.411153	-0.461799	-0.197573	0.904425	0.983369	11.69532	62.79325	0.5361765	constant

Table 18. Parameter Estimates: Paramterized Covariance Matrix

	Hypertension	Heart Disease	Diabetes	Any Health Insurance	Self-repoted Health	Weight Status	Smoking Status	Function Status	Working
Hypertension	1								
Heart Disease	0.277192	1							
Diabetes	0.300338	0.23171	1						
Any Health Insurance	0.0031552	-0.0242862	-0.0154716	1					
Self-repoted Health	0.290169	0.430257	0.324875	-0.0661165	1				
Weight Status	0.281839	0.0825386	0.156885	-0.0290745	0.134158	1			
Smoking Status	-0.0169154	0.0137628	-0.0006829	-0.0858103	0.0764173	-0.115586	1		
Function Status	0.120008	0.197944	0.103585	-0.0786604	0.415898	0.086167	-0.0202586	1	
Working	-0.0715813	-0.220485	-0.0951401	0.218957	-0.396341	0.0130773	-0.0471683	-0.360546	1
Nonzero Wealth	-0.147113	-0.167685	-0.0687014	0.116027	-0.200537	-0.0044556	-0.286169	-0.159866	0.382936
Log(AIME)	-0.0161113	-0.118696	-0.0441202	0.172663	-0.147874	-0.0134086	0.002129	-0.181199	0.380476
Log(Quarters Worked)	-0.0214464	-0.0997786	-0.0301324	0.115122	-0.116805	-0.0195221	0.0365283	-0.159176	0.363494
IHT(HH Wealth)	-0.0038586	-0.0383905	-0.0846944	0.118651	-0.113078	-0.0048792	-0.101793	-0.118847	0.0636059
IHT(Earned Income)	0.0052256	-0.0240596	-0.0869879	0.237859	-0.0874815	-0.0031914	-0.0414417	-0.104482	0
Log(DC Wealth)	0.027536	-0.000561	-0.0427334	0.215915	-0.0666775	-0.0465485	-0.0596319	-0.0938427	0.0249342
Any DC Plan	0.0263494	0.035424	0.0050752	0.228411	-0.0656421	-0.017682	-0.0619362	-0.0748389	0
Any DB Plan	0.0576736	-0.0453252	-0.128975	0.329201	-0.0517907	0.03503	-0.0528141	-0.0091015	0
Early Retirement Age	0.0584707	-0.0557669	0.0528724	0.0549995	-0.0059988	0.0337351	0.0178985	0.0486129	0.20218
Normal Retirement Age	-0.0124036	-0.0626024	-0.007316	0.175183	-0.0136174	0.0164101	0.0182109	-0.0124438	0.287243

Table 18. Parameter Estimates: Paramterized Covariance Matrix

Nonzero Wealth	Log(AIME)	Log(Quarters Worked)	IHT(HH Wealth)	IHT(Earned Income)	Log(DC Wealth)	Any DC Plan	Any DB Plan	Early Retirement Age	Normal Retirement Age
1									
0.183477	-4.4424								
0.155335	1.66693	-5.3004							
0	0.13385	0.0847593	-7.68598						
0.0996853	0.338047	0.183304	0.198855	-9.65995					
0.112154	0.376082	0.248571	0.272285	0.632512	-7.67886				
0.114475	0.288905	0.210902	0.108049	0.39555	1.0107	1			
0.0674868	0.088682	0.0076923	0.0816012	0.289428	0.181436	0.111172	1		
-0.0489416	0.067466	0.103534	-0.0590534	-0.246565	-0.0928313	-0.0818607	0	1	
0.073313	0.188392	0.210839	-0.0999346	-0.0977753	-0.0379107	0.0090438	0	0.332804	1

Table 19. Ex-Post Estimation of Receiving SSI

Probit of Receive SSI

	Administrative Data Received SSI
Self Repored SSI	1.93
	12.26
Non-Hispanic black	0.32
	2.4
Hispanic	0.84
	6.15
Less than high school	0.49
	3.92
Some college and above	-0.26
	-1.49
Male	-0.19
	-1.65
Age	0.02
	1.83
Heart disease	0.02
	0.15
Stroke	0.12
	0.69
Cancer	-0.24
	-1.4
Diabetes	0.14
	1.17
Lung disease	0.14
	0.93
IHT Wealth	-11.94
	-6.29
Wealth Non-Zero	-0.5
	-2.75
Single	0.63
	5.1
Widowed	0.12
	0.81
Constant	-2.83
	-3.97
N	3,910

Data source: HRS 2002, only those with linked social security ea

Table 20. Regression Estimates of Medical Spending (coefficient/t statistic)

	MCBS total medical costs	MEPS Total medical costs	MCBS total Medicaid costs	MEPS total Medicaid costs	MCBS out of pocket costs	MEPS out of pocket costs
Max(0, age - 75)	4.669		-4.222		29.98	
	(0.14)		(-0.45)		(2.20)	
Min(age, 75)	255.9	110.2	-7.242	-9.169	29.78	27.12
	(5.38)	(2.88)	(-0.82)	(-1.88)	(2.37)	(5.98)
Male	-237.6	-187.6	-247.2	-159.7	-354.4	-314.9
	(-0.91)	(-0.61)	(-4.36)	(-4.55)	(-3.70)	(-9.31)
Black	2586.7	-950.8	814.3	286.9	147.3	-430.2
	(3.11)	(-3.00)	(6.89)	(2.92)	(0.29)	(-11.16)
Hispanic	562.8	-1483.4	572.9	121.3	-285.2	-330.6
	(0.92)	(-6.61)	(5.62)	(1.65)	(-2.49)	(-8.50)
Less than high school	-411.6	-646.8	588.3	503.0	-707.2	-130.0
	(-1.12)	(-2.50)	(8.15)	(6.65)	(-3.84)	(-2.80)
Some college and above	441.4	371.7	-166.6	-90.43	159.5	135.0
	(1.52)	(1.17)	(-2.98)	(-2.34)	(1.30)	(3.51)
widowed	117.5	-373.2	153.3	234.3	205.0	69.05
	(0.34)	(-0.94)	(2.34)	(2.22)	(1.19)	(0.69)
single	-583.9	834.1	559.4	590.5	-431.6	108.7
	(-1.44)	(2.81)	(5.94)	(10.93)	(-2.53)	(2.67)
Cancer	3300.1	5785.3	100.8	478.4	218.4	509.1
	(10.97)	(7.84)	(1.83)	(3.62)	(2.11)	(6.19)
Diabetes	922.7	3286.3		667.1		554.6
	(1.38)	(8.51)		(7.52)		(8.20)
Hypertension	1299.5	1783.2	171.0	143.0	183.2	331.9
	(5.09)	(5.83)	(2.35)	(3.81)	(2.23)	(9.42)
Heart disease	3695.9	4544.7	162.5	430.7	722.5	373.3
	(5.92)	(10.27)	(1.57)	(5.80)	(1.96)	(6.41)
Lung disease	2429.7	2019.1	244.3	374.7	200.7	290.1
	(7.27)	(3.36)	(3.42)	(3.41)	(1.67)	(3.40)
Stroke	5748.7	3727.3	763.0	1329.1	2031.3	749.6
	(2.41)	(5.34)	(3.43)	(5.76)	(1.25)	(3.85)
Nursing Home	46206.4		21803.7		16445.6	
	(52.88)		(31.63)		(29.02)	
ADL 3+-Not in nursing home	7540.6		486.4		689.5	
	(11.86)		(5.21)		(3.36)	
Eligable for Medicare due to disablity	4848.0		284.1		1548.1	
	(3.96)		(2.00)		(2.23)	
Died	15463.1		454.4		1220.1	

Table 20. Regression Estimates of Medical Spending (coefficient/t statistic)

	(7.08)		(1.47)		(2.58)	
Diabetes & Heart Disease	2121.6		6.996		165.0	
	(2.86)		(0.06)		(0.54)	
Diabetes & Hypertension	2477.7		217.5		456.8	
	(3.19)		(2.27)		(2.29)	
Hypertension & Heart Disease	-584.0		-37.51		-452.0	
	(-0.94)		(-0.32)		(-1.45)	
Hypertension & Stroke	-2599.1		-618.0		-1651.7	
	(-1.09)		(-2.57)		(-1.04)	
Disability Claiming & Death	2842.3		-117.7		-436.2	
	(0.27)		(-0.11)		(-0.51)	
Disability Claiming & Nursing Home	3053.7		8648.9		-10287.0	
	(0.61)		(2.86)		(-11.47)	
Death & Nursing Home	-36889.0		-13140.1		-8546.9	
	(-17.62)		(-14.24)		(-9.88)	
Death & Cancer	4235.3		-874.3		-53.09	
	(1.40)		(-3.49)		(-0.13)	
Death & Diabetes	4382.5		718.3		-14.61	
	(1.83)		(1.81)		(-0.04)	
Death & Hypertension	-1349.4		-474.2		-261.7	
	(-0.65)		(-1.42)		(-0.47)	
Death & Heart Disease	827.9		116.1		-888.0	
	(0.41)		(0.35)		(-2.12)	
Death & Lung Disease	-2296.0		-83.61		-372.1	
	(-0.93)		(-0.24)		(-0.95)	
Death & Stroke	-4147.2		75.33		-981.5	
	(-1.46)		(0.19)		(-1.41)	
Constant	-12354.9	-3474.1	511.7	426.6	-639.5	-701.6
	(-3.71)	(-1.63)	(0.83)	(1.53)	(-0.72)	(-2.78)

t statistics in parentheses

Table 21. Regression Estimates of Medicare Spending in Parts A and B (coefficient/t statistic)

	total medicare costs	medicare pt a costs	medicare pt b costs
Max(0, age - 75)	-5.524	1.576	-29.25
	(-0.24)	(0.07)	(-3.03)
Min(age, 75)	231.3	124.9	79.12
	(8.28)	(5.29)	(5.72)
Male	128.7	39.68	-5.157
	(0.76)	(0.27)	(-0.06)
Black	2053.0	1121.5	1140.0
	(5.36)	(3.72)	(4.00)
Hispanic	693.5	285.3	453.7
	(1.85)	(0.71)	(2.64)
Less than high school	280.3	406.9	-70.00
	(1.33)	(2.18)	(-0.59)
Some college and above	65.56	107.6	39.45
	(0.37)	(0.70)	(0.41)
widowed	241.5	230.2	28.78
	(1.19)	(1.27)	(0.27)
single	179.8	257.2	3.530
	(0.72)	(1.23)	(0.02)
Cancer	2077.1	775.3	1784.5
	(10.45)	(4.72)	(13.97)
Diabetes	309.0		749.2
	(0.72)		(2.79)
Hypertension	502.9	321.5	287.5
	(2.88)	(2.10)	(3.24)
Heart disease	1931.9	1365.5	1051.4
	(6.83)	(6.02)	(5.33)
Lung disease	1294.0	743.7	783.9
	(5.70)	(3.76)	(6.68)
Stroke	2993.1	1707.0	1752.5
	(3.82)	(3.67)	(2.43)
nhmliv	6113.2	4498.1	1521.3
	(10.35)	(8.39)	(7.87)
ADL 3+-Not in nursing home	5044.2	3968.8	2040.7

Table 21. Regression Estimates of Medicare Spending in Parts A and B (coefficient/t statistic)

	(11.26)	(9.86)	(9.94)
Eligable for Medicare due to disablity	2327.7	777.7	1305.5
	(4.27)	(1.95)	(3.35)
Died	12977.8	13464.0	2769.3
	(7.56)	(7.90)	(6.36)
Diabetes & Heart Disease	1443.4	1139.1	275.0
	(3.06)	(2.95)	(1.23)
Diabetes & Hypertension	1504.1	984.3	471.8
	(3.11)	(3.48)	(1.74)
Hypertension & Heart Disease	-151.1	-141.4	-89.94
	(-0.44)	(-0.48)	(-0.45)
Hypertension & Stroke	-1091.5	-6.432	-1293.2
	(-1.28)	(-0.01)	(-1.82)
Disability Claiming & Death	-3750.4	-4995.2	20.03
	(-0.86)	(-1.37)	(0.02)
Disability Claiming & Nursing Home	1247.1	776.4	700.7
	(0.79)	(0.58)	(0.91)
Death & Nursing Home	-12115.9	-11093.6	-2932.9
	(-7.34)	(-6.86)	(-6.89)
Death & Cancer	2524.8	2345.4	662.6
	(1.12)	(1.05)	(1.03)
Death & Diabetes	5458.7	5557.1	756.0
	(2.73)	(2.89)	(1.45)
Death & Hypertension	-934.6	-1249.7	446.4
	(-0.56)	(-0.76)	(0.98)
Death & Heart Disease	1715.7	2113.7	-1384.9
	(1.01)	(1.26)	(-3.09)
Death & Lung Disease	329.3	172.6	-33.04
	(0.16)	(0.09)	(-0.06)
Death & Stroke	-3618.9	-2375.8	-1811.1
	(-1.80)	(-1.20)	(-3.54)
Constant	-13778.1	-8094.8	-4310.5
	(-7.05)	(-4.94)	(-4.42)

t statistics in parentheses

Table 22. Probit Regression Estimates on Medicare Part B Take up (coefficient/t statistic)

	New Meidcare Enrollees	Existing Medicare Enrollees not in Part B	All Medicare Eligible
Male	-0.190 (-1.82)	-0.127 (-1.47)	-0.196 (-6.59)
Black	-0.207 (-1.17)	-0.239 (-1.73)	-0.291 (-6.76)
Hispanic	-0.382 (-2.25)	-0.153 (-1.08)	-0.473 (-11.75)
Less than high school	0.104 (0.75)	-0.249 (-2.11)	0.0488 (1.44)
Some college and above	-0.0728 (-0.67)	-0.178 (-1.85)	-0.118 (-3.69)
widowed	0.130 (0.79)	0.0688 (0.57)	0.0505 (1.41)
Earned Income ($1000)	-0.00646 (-4.14)	-0.00134 (-0.88)	-0.00385 (-9.03)
work	-0.555 (-5.28)	-0.532 (-5.28)	-0.561 (-18.34)
Cancer	0.0600 (0.42)	0.0731 (0.58)	0.133 (3.58)
Hypertension	0.397 (3.20)	0.163 (1.86)	0.102 (3.68)
Heart disease	-0.156 (-1.25)	0.124 (0.73)	0.0882 (1.78)
Stroke	0.211 (1.03)	0.733 (2.89)	0.145 (1.65)
ADL 2-Not in nursing home	-0.0733 (-0.43)		
2 ADLs	-0.300 (-1.34)		
ADL 3+-Not in nursing home	0.137 (0.36)	0.0353 (0.22)	0.0629 (1.22)

Table 22. Probit Regression Estimates on Medicare Part B Take up (coefficient/t statistic)

Obese(bmi>=30)	0.190	-0.0296	0.0508
	(1.12)	(-0.29)	(1.58)
Ever smoked	0.0442	-0.0417	0.0258
	(0.43)	(-0.39)	(0.75)
Eligable for Medicare due to disablity	-0.331	-1.092	-0.146
	(-2.02)	(-4.81)	(-2.09)
Diabetes & Hypertension	-0.137		
	(-0.73)		
Diabetes & Heart Disease	0.280	-0.183	
	(1.16)	(-0.83)	
Hypertension & Obesity	-0.247		
	(-1.09)		
DI Claim and 3+ ADLs	0.670		
	(1.13)		
Max(age, 75)		-0.0433	
		(-2.89)	
Min(0, age - 75)		-0.0795	0.00979
		(-3.46)	(2.20)
Diabetes		0.0731	0.0907
		(0.49)	(2.57)
Lung disease		0.163	
		(1.46)	
hibpe_stroke		-0.633	-0.0816
		(-2.10)	(-0.80)
hearte_smokev		-0.194	-0.138
		(-1.01)	(-2.36)
Age Spline Knot at 65			0.00731
			(0.94)
Age Spline Knot at 75			0.0721
			(14.71)
Constant	1.453	2.531	1.231
	(9.70)	(2.41)	(2.42)

t statistics in parentheses

Table 23. Compare per capita medical spending by payment sources and by age group in year 2004

	Age 55-64			Age 65 and over		
	NHEA 2004 ($)	FEM 2004, unadjusted ($)	Adjustment factor	NHEA 2004 ($)	FEM 2004, unadjusted ($)	Adjustment factor
	(1)	(2)	(1)/(2)	(3)	(4)	(3)/(4)
Payment sources						
Total	7,787	6,411	1.21	14,797	13,920	1.06
Medicare	706	508	1.39	7,242	7,339	0.99
Medicaid	1,026	400	2.57	2,034	1,222	1.66

Table 24. Probit Regression Estimates on Medicare Part D Enrollment (coefficient/t statistic)

	Medicare Part D Enrollment
Max(age, 75)	0.0352 (6.66)
Min(0, age - 75)	-0.0176 (-5.25)
Male	-0.201 (-6.21)
Black	0.117 (2.16)
Hispanic	0.232 (4.01)
Less than high school	0.295 (7.93)
Some college and above	-0.0433 (-1.24)
married	-0.164 (-4.98)
Earned Income ($1000s)	-0.00439 (-6.80)
work	-0.144 (-2.98)
Cancer	-0.0218 (-0.60)
Diabetes	0.0852 (2.41)
Hypertension	0.0248 (0.78)
Stroke	0.0495 (1.12)
Heart disease	0.0220 (0.73)
ADL 3+-Not in nursing home	0.0818 (1.56)
Ever smoked	-0.0717 (-2.29)
Eligable for Medicare due to disablity	0.393 (4.36)
Constant	-2.203 (-5.80)

t statistics in parentheses

Table 25. OLS Regression Estimates on Medicare Part D Spending (coefficient/t statistic)

	Medicare Part D Expenditures
Max(age, 75)	-22.39 (-3.88)
Min(0, age - 75)	-2.978 (-0.84)
Male	-259.4 (-8.10)
Black	412.0 (6.73)
Hispanic	154.7 (2.88)
Less than high school	444.3 (11.73)
Some college and above	-74.75 (-2.19)
widowed	178.4 (5.01)
Diabetes	190.4 (5.01)
Hypertension	171.4 (5.57)
Stroke	136.1 (2.73)
Heart disease	167.2 (5.49)
Lung disease	295.6 (6.69)
Current smoking	163.8 (3.11)
Eligable for Medicare due to disablity	506.9 (4.86)
IADL 1-Not in nursing home	0 (.)
IADL 2+-Not in nursing home	-917.4 (-10.04)
ADL 2-Not in nursing home	103.9 (2.29)
2 ADLs	160.5 (2.55)
ADL 3+-Not in nursing home	234.8 (3.86)
Constant	3403.5 (8.75)

t statistics in parentheses

Table 26. Status Quo FEM Estimates

		Baseline Estimates		
	Year	2010	2030	2050
	Population size (Million)	97.26	132.24	162.91
	Population 65+ (Million)	43.24	75.35	95.46
	obesity (BMI >=30) (%)	34%	48%	55%
	over weight (25<=BMI<30) (%)	34%	30%	27%
	Ever-smoked	56%	45%	32%
Prevalence of selected conditions	Smoking now	15%	9%	5%
	Diabetes	18%	26%	31%
	Heart disease	23%	28%	31%
	Hypertension	51%	61%	65%
Labor participation	Working (%)	37%	31%	29%
	Average earnings if working ($2010)	$ 45,493.72	$ 44,965.23	$ 47,678.82
Government revenues from aged 51+ (Billion $2010)	Federal personal income taxes	$ 238.44	$ 305.22	$ 388.23
	Social security payroll taxes	$ 92.63	$ 110.00	$ 133.59
	Medicare payroll taxes	$ 22.45	$ 25.73	$ 31.24
	Total Revenue	$ 353.53	$ 440.95	$ 553.06
Government expenditures from aged 51+ (Billion $2010)	Old Age and Survivors Insurance benefits (OASI)	$ 564.54	$ 1,320.72	$ 1,899.83
	Disability Insurance benefits (DI)	$ 48.06	$ 58.04	$ 84.76
	Supplementary Security Income (SSI)	$ 22.22	$ 31.05	$ 45.99
	Medicare costs	$ 628.99	$ 1,465.27	$ 3,099.33
	Medicaid costs	$ 206.36	$ 407.24	$ 1,009.03
	Medicare + Medicaid	$ 835.36	$ 1,872.51	$ 4,108.36
	Total medical costs for aged 51+ (Billion $2010)	$ 1,621.79	$ 3,367.58	$ 7,107.61

Table 27. Obesity Reductions Scenario - FEM Estimates

	Obesity Estimates		Relative Change to Status Quo		Absolute Change to Status Quo	
	Year		Year		Year	
	2030	2050	2030	2050	2030	2050
Population size (Million)	132.503921	164.368767	0.20%	0.90%	0.27	1.46
Population 65+ (Million)	75.5230253	96.7897985	0.23%	1.40%	0.18	1.33
Prevalence of selected conditions						
obesity (BMI >=30) (%)	41%	44%	-14.15%	-19.46%	-0.07	-0.11
over weight (25<=BMI<30) (%)	32%	29%	4.40%	6.03%	0.01	0.02
Ever-smoked	45%	32%	0.11%	0.43%	0.00	0.00
Smoking now	9%	6%	2.04%	3.80%	0.00	0.00
Diabetes	22%	24%	-15.92%	-24.97%	-0.04	-0.08
Heart disease	28%	30%	-1.58%	-3.43%	0.00	-0.01
Hypertension	58%	61%	-5.09%	-6.83%	-0.03	-0.04
Labor participation						
Working (%)	31%	29%	0.57%	0.06%	0.00	0.00
Average earnings if working ($2010)	$ 44,970.08	$ 47,681.50	0.01%	0.01%	4.85	2.68
Government revenues from aged 51+ (Billion $2010)						
Federal personal income taxes	$ 307.83	$ 392.85	0.86%	1.19%	2.61	4.62
Social security payroll taxes	$ 110.85	$ 134.93	0.77%	1.00%	0.85	1.34
Medicare payroll taxes	$ 25.92	$ 31.56	0.77%	1.00%	0.20	0.31
Total Revenue	$ 444.60	$ 559.33	0.83%	1.13%	3.66	6.27
Government expenditures from aged 51+ (Billion $2010)						
Old Age and Survivors Insurance benefits (OASI)	$ 1,323.70	$ 1,924.47	0.23%	1.30%	2.98	24.64
Disability Insurance benefits (DI)	$ 56.80	$ 81.63	-2.13%	-3.69%	-1.24	-3.13
Supplementary Security Income (SSI)	$ 30.64	$ 45.67	-1.30%	-0.70%	-0.40	-0.32
Medicare costs	$ 1,450.64	$ 3,031.89	-1.00%	-2.18%	-14.64	-67.44
Medicaid costs	$ 389.56	$ 960.41	-4.34%	-4.82%	-17.68	-48.62
Medicare + Medicaid	$ 1,840.19	$ 3,992.30	-1.73%	-2.82%	-32.32	-116.06
Total medical costs for aged 51+ (Billion $2010)	$ 3,295.19	$ 6,885.00	-2.15%	-3.13%	-72.39	-222.61

Table 32. Quailty Adjusted Life Year Model

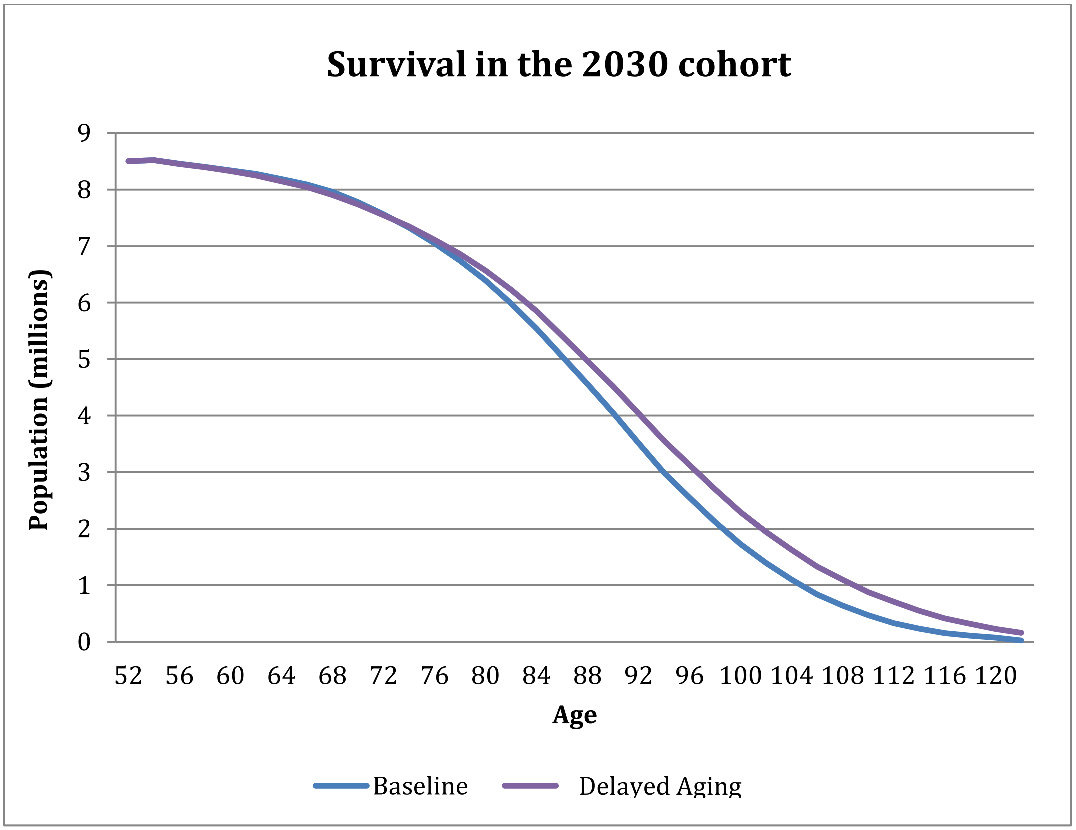
	Quality Adjusted Life Year
Has exactly 1 IADL	-0.0365
	(-14.34)
Has 2 or more IADLs	-0.0453
	(-13.26)
Has exactly 1 ADL	-0.0700
	(-34.32)
Has exactly 2 ADLs	-0.121
	(-40.98)
Has 3 or more ADLs	-0.168
	(-54.26)
Cancer	-0.0208
	(-11.32)
Diabetes	-0.0402
	(-22.64)
Heart disease	-0.0407
	(-28.75)
Hypertension	-0.0420
	(-35.78)
Lung disease	-0.0712
	(-33.35)
Stroke	-0.0374
	(-16.18)
Single	-0.0271
	(-15.06)
Widowed	-0.00320
	(-2.31)
Constant	0.897
	(996.98)

t statistics in parentheses

9 Supplemental Exhibits

EXHIBIT S1

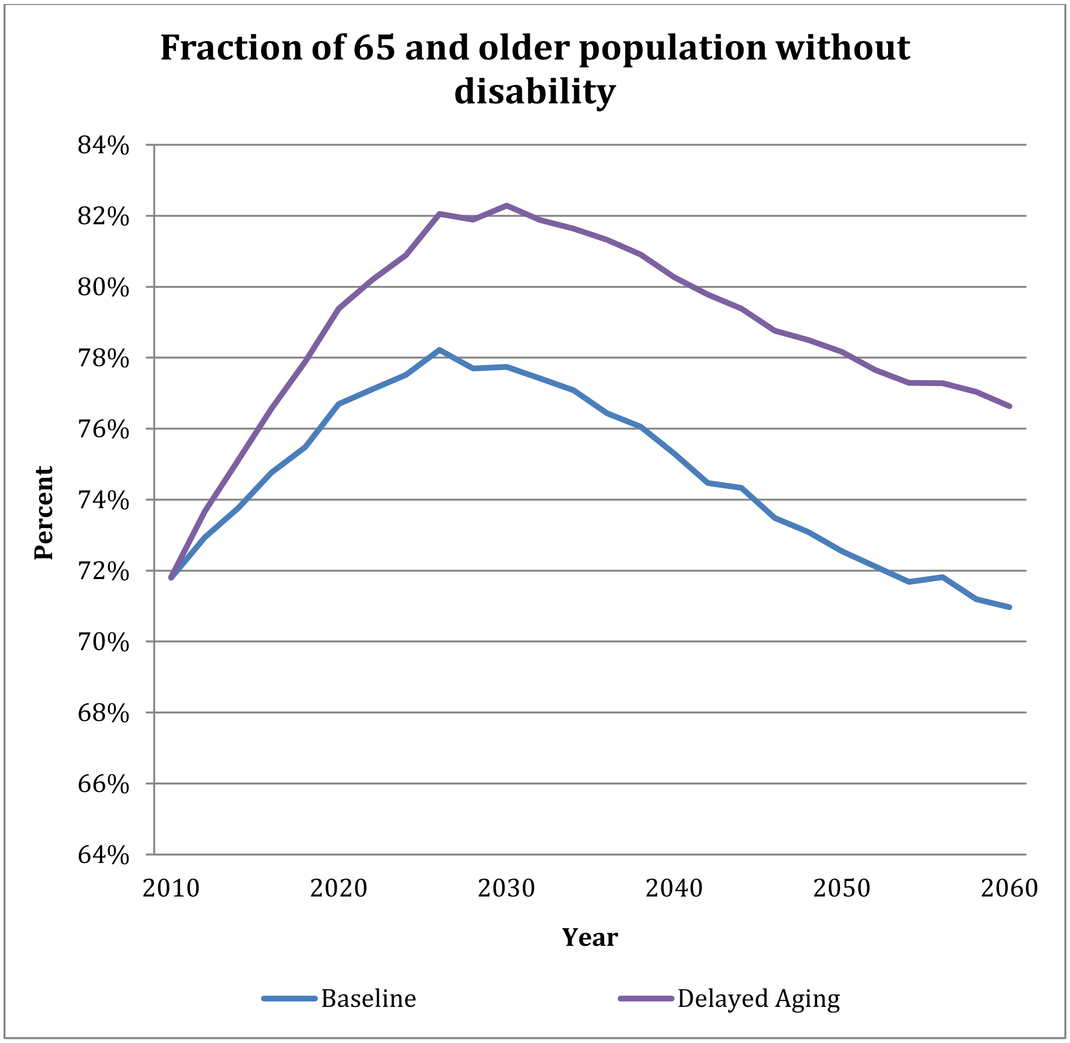
Survival in the 2030 cohort under baseline and delayed aging scenarios



SOURCE: Authors' calculations from the Future Elderly Model.

EXHIBIT S2

Fraction of 65 and older population without disability

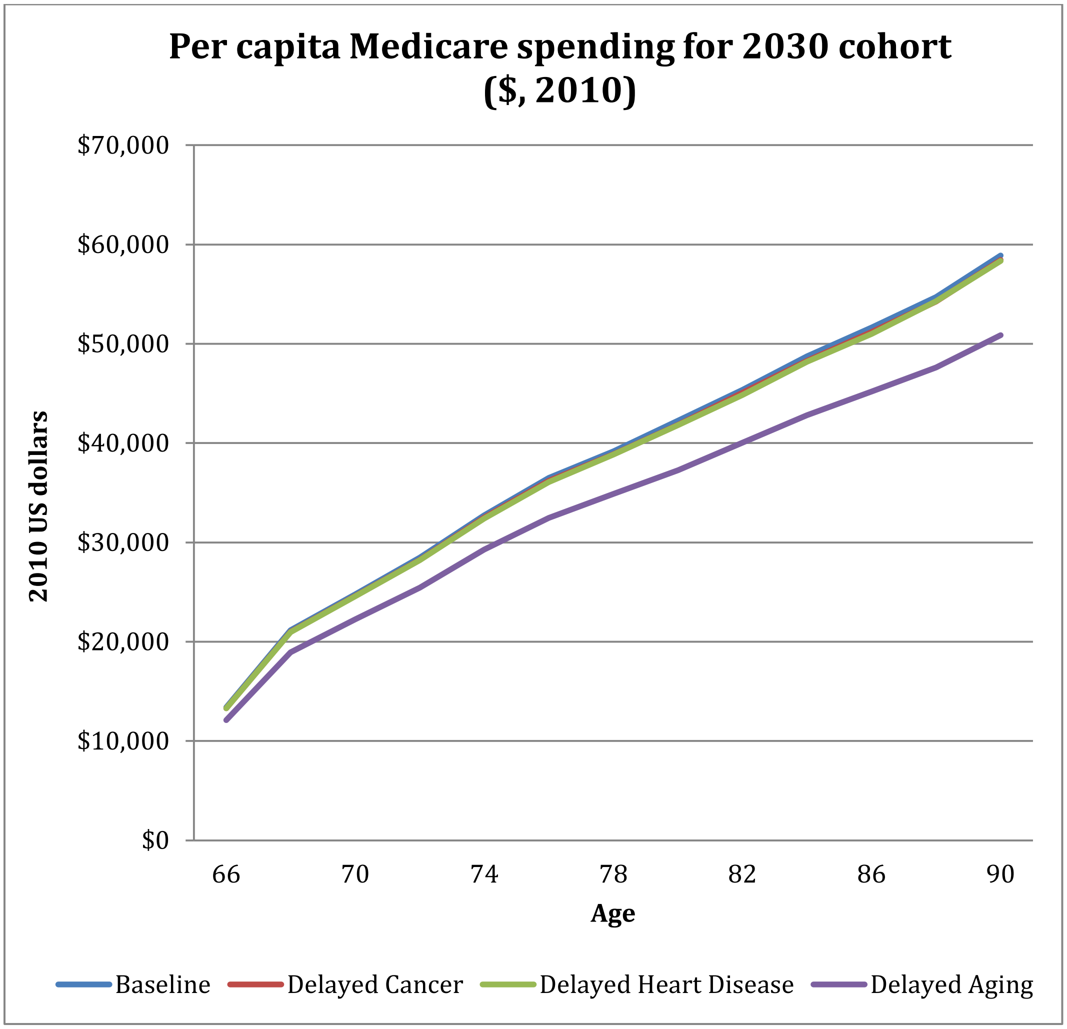


SOURCE: Authors' calculations from the Future Elderly Model.

EXHIBIT S3

Per capita Medicare spending for 2030 cohort ($, 2010)

(Calculations based on 10% reduction in cancer and heart disease incidence)



SOURCE: Authors' calculations using the Future Elderly Model.

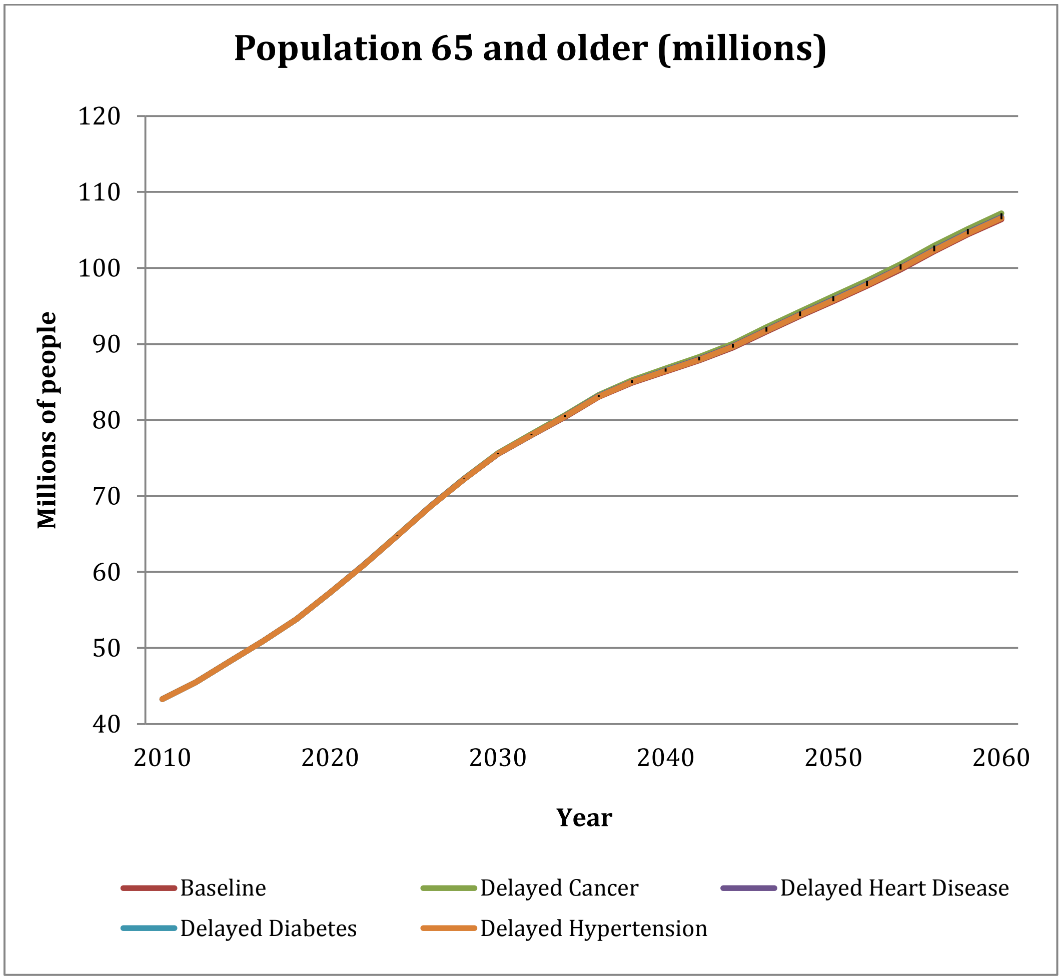
Note: Predicted lifetime per capita Medicare spending for a representative sample aged 51/52 in 2030 under four different medical progress scenarios.

10 Sensitivity Analysis

EXHIBIT S4

Population under alternative disease scenarios (delayed diabetes, delayed hypertension)

(Calculations based on 10% reduction in disease incidence)



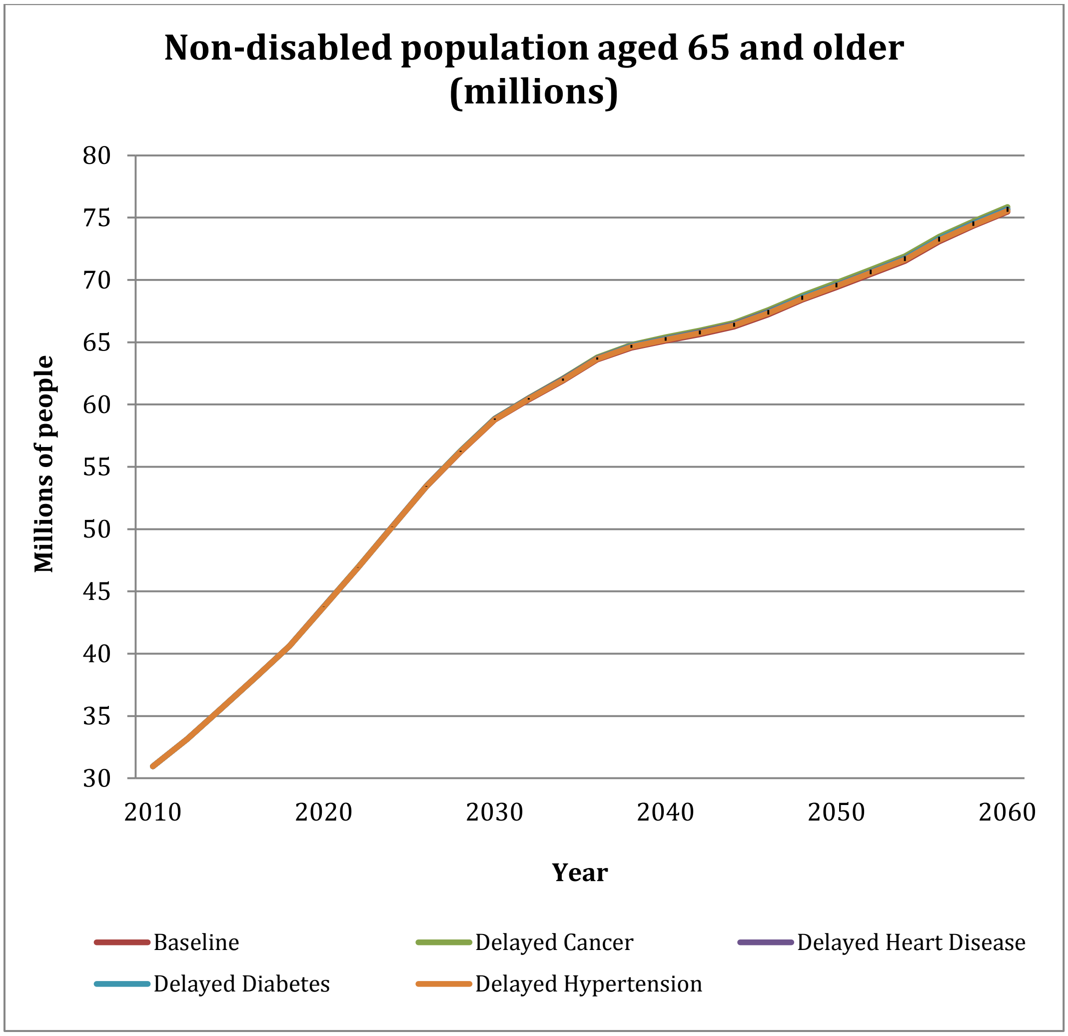
SOURCE: Authors' calculations using the Future Elderly Model.

Note: Population 65 years of age and older under various medical progress scenarios.

EXHIBIT S5

Non-disabled population under alternative disease scenarios (delayed diabetes, delayed hypertension)

(Calculations based on 10% reduction in disease incidence)



SOURCE: Authors' calculations using the Future Elderly Model.

Note: Total number of Americans aged 65 and older who are not disabled under various medical progress scenarios. Disability was defined as having any of the following: limitation in instrumental activities of daily living (IADLs), limitation in activities of daily living (ADL), or living in a nursing home.

EXHIBIT S6

Remaining life expectancy at age 50 for all four scenarios compared across race/ethnicity and education level

(Calculations based on 10% reduction in cancer and heart disease incidence)

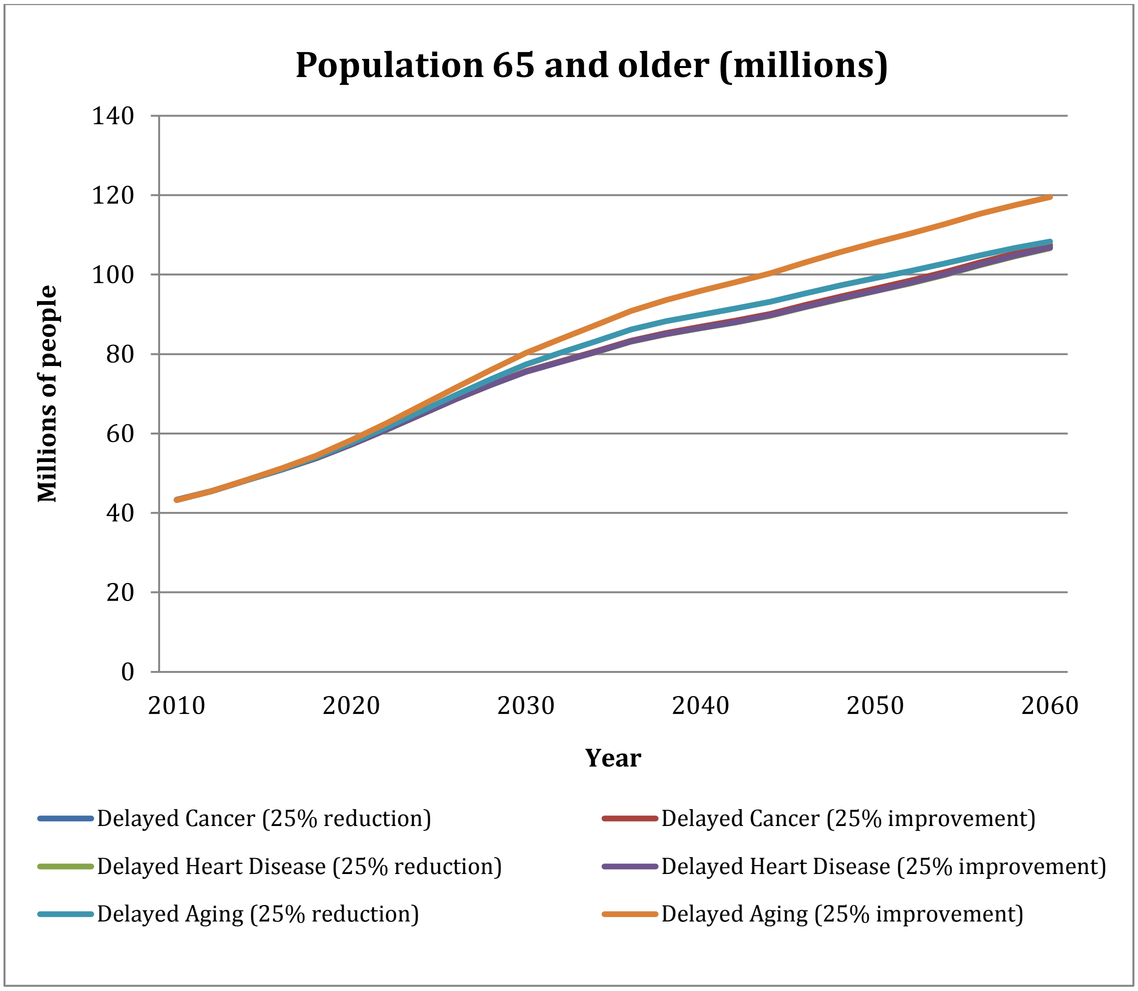
	Baseline	Delayed Cancer	Delayed Heart Disease	Delayed Aging
Black, Less Than High School	28.39	29.37	29.39	29.77
Black, High School Grad	31.90	32.86	32.75	33.49
Black, Some College	34.17	35.05	24.88	36.11
White, Less Than High School	28.58	29.71	29.58	30.38
White, High School Grad	35.24	36.11	35.93	37.52
White, Some College	37.07	37.69	37.56	39.33
Hispanic, Less Than High School	33.23	34.03	34.00	35.09
Hispanic, High School Grad	37.81	38.47	38.40	40.00
Hispanic, Some College	38.85	39.92	39.79	41.80

SOURCE: Authors' calculations using the Future Elderly Model.

EXHIBIT S7

Population 65 and older under alternative disease and delayed aging scenarios

(Calculations based on 10% reduction in cancer and heart disease incidence as the anchor scenarios for delayed cancer and delayed heart disease)



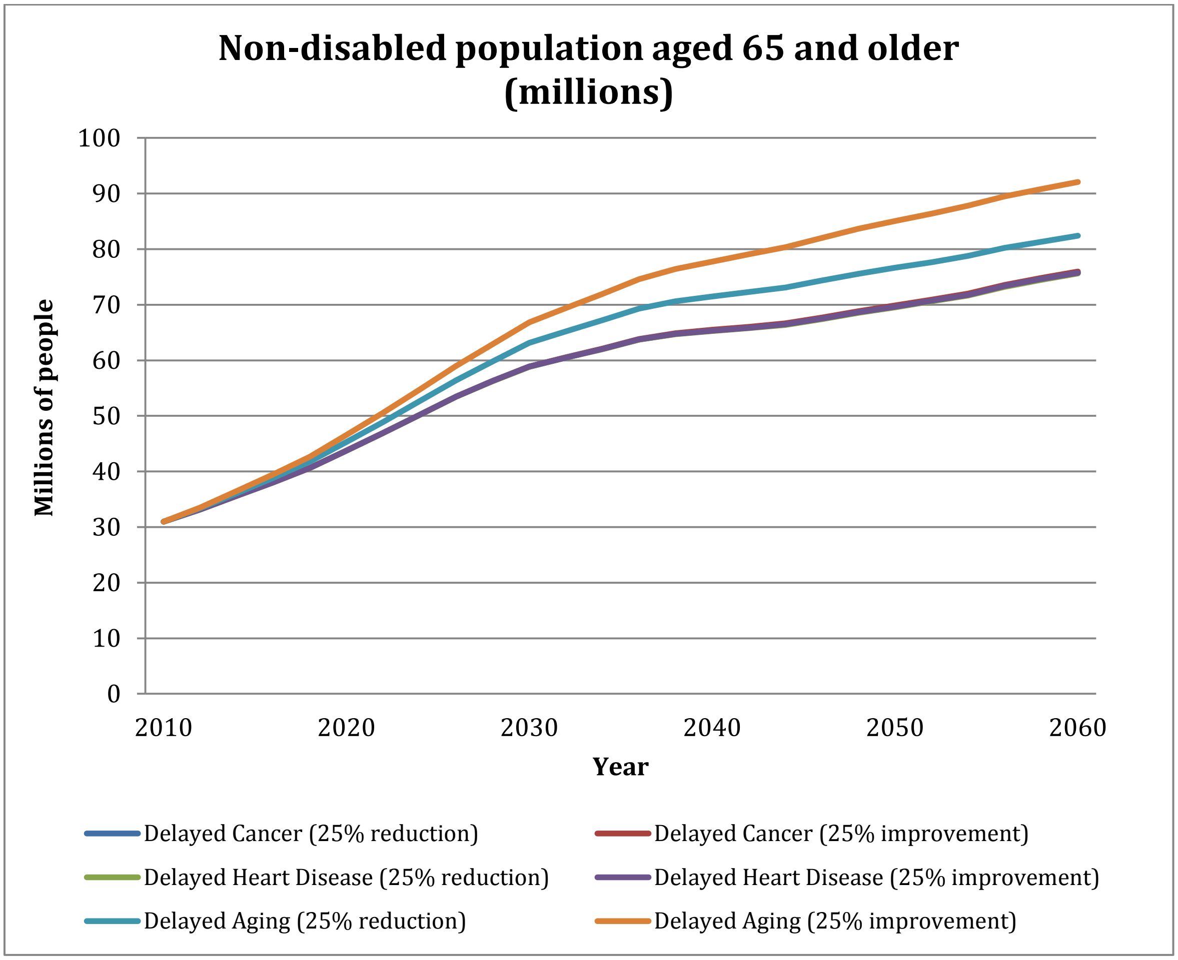
SOURCE: Authors' calculations from the Future Elderly Model.

Note: Population 65 years of age and older under various medical progress scenarios.

EXHIBIT S8

Non-disabled population aged 65 and older (millions) under alternative disease and delayed aging scenarios

(Calculations based on 10% reduction in cancer and heart disease incidence as the anchor scenarios for delayed cancer and delayed heart disease)



SOURCE: Authors' calculations from the Future Elderly Model.

Note: Total number of Americans aged 65 and older who are not disabled under various medical progress scenarios. Disability was defined as having any of the following: limitation in instrumental activities of daily living (IADLs), limitation in activities of daily living (ADL), or living in a nursing home.

References

1. Goldman, D., et al., *Health status and medical treatment of the future elderly: Final Report*, 2004, RAND Corporation: Santa Monica, CA.

2. Health and Retirement Study (RAND HRS), *Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740)*, 2012: Ann Arbor, MI, .

3. Michaud, P.-C., et al., *The Effects of Attrition and Non-Response in the Health and Retirement Study*, in *RAND Working Paper 407*2006: Santa Monica, CA.

4. Honeycutt, A.A., et al., *A dynamic Markov model for forecasting diabetes prevalence in the United States through 2050.* Health care management science, 2003. **6**(3): p. 155-64.

5. Levy, D., *Trends in Smoking Rates Under Different Tobacco Control Policies:Results from the SimSmoke Tobacco Policy Simulation Model*, 2006.

6. Poterba, J., S. Venti, and D.A. Wise, *The Decline of Defined Benefit Retirement Plans and Asset Flows*, 2007, National Bureau of Economic Research, Inc, NBER Working Papers: 12834.

7. Poterba, J., S. Venti, and D.A. Wise, *New Estimates of the Future Path of 401(k) Assets*, 2007, National Bureau of Economic Research, Inc, NBER Working Papers: 13083.

8. Ruhm, C., *Current and Future Prevalence of Obesity and Severe Obesity in the United States.* Forum for Health Economics & Policy, 2007. **10**(2).

9. Mainous, A.G., 3rd, et al., *Impact of the population at risk of diabetes on projections of diabetes burden in the United States: an epidemic on the way.* Diabetologia, 2007. **50**(5): p. 934-40.

10. Wooldridge, J.M., *A framework for estimating dynamic, unobserved effects panel data models with possible feedback to future explanatory variables.* Economics Letters, 2000. **68**(3): p. 245-250.

11. MacKinnon, J.G. and L. Magee, *Transforming the Dependent Variable in Regression Models.* International Economic Review, 1990. **31**(2): p. 315-339.

12. Dolan, P., *Modeling valuations for EuroQol health states.* Medical care, 1997: p. 1095-1108.

13. Shaw, J.W., J.A. Johnson, and S.J. Coons, *US valuation of the EQ-5D health states: development and testing of the D1 valuation model.* Med Care, 2005. **43**(3): p. 203-20.

14. Train, K., *Discrete Choice Methods with Simulation*2003, Cambridge: Cambridge University Press.

15. Atherly, A., B.E. Dowd, and R. Feldman, *The effect of benefits, premiums, and health risk on health plan choice in the Medicare program.* Health Serv Res, 2004. **39**(4p1): p. 847-864.

16. Buchmueller, T., *Price and the health plan choices of retirees.* J Health Econ, 2006. **25**(1): p. 81-101.