Healthy, Wealthy, and Wise?

Tests for Direct Causal Paths between Health and Socioeconomic Status

by

Peter Adams, University of California, Berkeley
Michael D. Hurd, RAND Corporation
Daniel McFadden, University of California, Berkeley*
Angela Merrill, Mathematica
Tiago Ribeiro, University of California, Berkeley

February 2002

Forthcoming, *Journal of Econometrics*

ABSTRACT: This paper provides statistical methods that permit the association of socioeconomic status and health to be partially unraveled in panel data by excluding some postulated causal paths, or delimiting their range of action. These methods are applied to the Asset and Health Dynamics of the Oldest Old (AHEAD) Panel to test for the *absence* of causal links from socioeconomic status (SES) to health innovations and mortality, and from health conditions to innovations in wealth. We conclude that in this elderly American population, where Medicare covers most acute care and pension income is not affected by ability to work, the evidence supports the hypothesis of no direct causal link from SES to mortality and to incidence of most sudden onset health conditions (accidents and some acute conditions), once initial health conditions are controlled, but there is some association of SES with incidence of gradual onset health conditions (mental conditions, and some degenerative and chronic conditions), due either to causal links or to persistent unobserved behavioral or genetic factors that have a common influence on both SES and innovations in health. There is mixed evidence for an association of health conditions and innovations in wealth. The death of a spouse appears to have a negative effect on the wealth of the survivor; this is plausibly a direct causal effect. There is evidence for some association of health conditions with increased dissaving from liquid wealth for intact couples and singles. From these findings, we conclude that SES-linked therapies for acute diseases are not, on the whole, the source of mortality differentials in this population. The question of whether SES-linked preventative care influences onset of chronic and mental diseases remains open. The Appendix to this paper containing the detailed model estimates, the data, and the programs used for data preparation and estimation, can be found at http://elsa.berkeley.edu/wp/hww/

We gratefully acknowledge financial support from the National Institute on Aging through a grant to the NBER Program Project on the Economics of Aging. We thank Laura Chioda, Victor Fuchs, Rosa Matzkin, James Poterba, and Jim Powell for useful comments.

* Corresponding Author: Department of Economics, University of California, Berkeley CA 94720-3880, e-mail mcfadden@econ.berkeley.edu, home page http:/elsa.berkeley.edu/~mcfadden.

Healthy, Wealthy, and Wise?
Tests for Direct Causal Paths between Health and Socioeconomic Status

Peter Adams, Michael Hurd, Daniel McFadden, Angela Merrill, and Tiago Ribeiro

1. Introduction

1.1. The Issue. The links between health, wealth, and education have been studied in a number of populations, with the general finding that higher socioeconomic status (SES) is associated with better health and longer life.[1] In a survey of this literature, Goldman (2001) notes that this association has been found in different eras, places, genders, and ages, and occurs over the whole range of SES levels, so that it is not linked solely to poverty. The association holds for a variety of health variables (most illnesses, mortality, self-rated health status, psychological well-being, and biomarkers such as allostatic load) and alternative measures of SES (wealth, education, occupation, income, level of social integration).[2] There has been considerable discussion of the causal mechanisms behind this association, but there have been relatively few natural experiments that permit causal paths to be definitively identified.[3] In this paper, we test for the *absence* of direct causal links in an elderly population by examining whether *innovations* in health and wealth in a panel are influenced by features of the historical state.

Figure 1 depicts possible causal paths for the health and SES innovations that occur over a short period. An individual's life history is built from these period-by-period transitions. First, low SES may lead to failures to seek medical care and delay in detection of conditions, reduced access to medical services, less effective therapies, or failures to maintain treatment regimens.[4] Also, increased risk of health problems may result from increased stress or frustration, or increased exposure to environmental hazards, that are associated with low SES.[5] These factors could provide direct causal links from SES history to health events. Second, poor health may reduce the ability to work or look after oneself, and increase medical care expenditures, leading to reduced income and less opportunity to accumulate assets. This could provide a direct causal link from health to changes in SES.

There may also be hidden *common factors* that lead to ecological association of health and SES. For example, unobserved genetic heterogeneity may influence both resistance to disease and ability to work. Causal links may be reinforced or confounded by behavioral response. Behavioral factors such as childhood nutrition and stress, exercise, and smoking may influence both health and economic activity level. For example, tastes for work and for "clean living", whether genetic or learned, may influence both health and earnings. Finally, rational economic decision-making may induce robust consumers to accumulate in order to finance consumption over a long expected retirement, or unhealthy individuals to spend down assets.

Preston and Taubman (1994), Smith and Kington (1997), and Smith (1999) give detailed discussions of the various causal mechanisms that may be at work, and the role of behavioral response from economic consumers. The epidemiological literature (Goldman, 2001) uses a different terminology for the causal paths: links from SES to health innovations are termed *causal mechanisms*, while links from health to SES are termed *selection* or *reverse causation* mechanisms. Apparent association due to measurement errors, such as overstatement of the SES of the healthy or under-detection of illnesses among the poor, are called *artifactual mechanisms*.[6] This literature classifies all common factors in terms of their implicit initial action as either causal or selection mechanisms.[7]

1.2. This Study. We study the population of elderly Americans aged 70 and older, and in this population test for the *absence* of direct causal paths from SES to innovations in health, and from health status to innovations in SES. These hypotheses will in general be accepted *only* if no causal link is present *and* there are no persistent hidden factors that influence both initial state and innovations. Rejection of one of these hypotheses does not demonstrate a direct causal link, since this may be the result of common hidden factors. However, in an elderly population persistent hidden factors will often be manifest in observed covariates, so that once these covariates are controlled, the residual impact of the hidden factors on innovations will be small. For example, genetic frailty that is causal to both health problems and low wages, leading to low wealth, may be expressed through a health condition such as diabetes. Then, onset of new health conditions that are also linked to genetic frailty may be only weakly associated with low wealth, once diabetic condition has been entered as a covariate. Thus, in this population, *rejection* of the hypotheses *may* provide useful diagnostics for likely causal paths.

The objectives and conclusions of this paper are limited. We study only elderly Americans, for whom Medicare provides relatively homogeneous and comprehensive health care at limited out-of-pocket cost to the individual. This population is retired, so that new health problems do not impact earnings. Statements about the presence or absence of direct causal mechanisms in this population, given previous health and SES status, say nothing about the structure of these mechanisms in a younger population, where associations of health and SES emerge as a result of some pattern of causation and operation of common factors.

Our tests for the *absence* of causality do *not* address the question of how to identify invariant models and causal links when these tests fail. If a test for the absence of a direct causal path is rejected, it may be possible through natural or designed experiments to separate causal and ecological effects; see Angrist, Imbens, and Rubin (1996), Heckman (2000, 2001). Suppose a strictly exogenous variable is causal to SES, and clearly not itself directly causal to health or causal to common factors. Then, an association of this variable with innovations in health conditions can only be through a direct causal link from SES to health. A variable with these properties is termed an *proper instrument* or *control variable* for SES. Proper instruments are hard to find. They can be obtained through designed experiments, where random treatment assignment precludes the possibility of confounding by common factors, provided recruitment and retention of experimental subjects does not re-introduce confounding. For example, an experiment that randomly assigned co-payment rates and coverage within Medicare for prescription drugs or assisted living could provide evidence on direct causal links from SES to health conditions, provided attrition and compliance are not problems. Occasionally, natural experiments may provide random treatment assignment. Economic events that impact individuals differently and that are not related to their prior SES or health are potentially proper instruments. For example, a tax change that affects wealth differently in different States is arguably a proper instrument, as is a change in mandated Medicaid coverage that has a differential impact across States. Individual events such as receipt of inheritances may be proper instruments, although they would be confounded if they are anticipated, or if the probability of their occurrence is linked to health status; see Meer, Miller, and Rosen (2001). Weak association between SES and a proper instrument for it makes it difficult to obtain precise estimates of direct causal effects; see Staiger and Stock (1997).

Section 2 of this paper discusses the foundation for econometric causality tests, and sets out the models for the dynamics of health and SES that will be used for our analysis. Section 3 describes the panel study and data that we use. Section 4 describes the association of SES and prevalence of health conditions in the initial wave of the panel. Section 5 analyzes incidence of new health conditions, and presents tests for non-causality of SES for the incidence of health innovations. Section 6 tests for the

absence of a causal link from health conditions to wealth accumulation and other SES indicators. Section 7 uses our estimated models for prevalence and incidence to simulate life histories for a current population aged 70 under counterfactual (and unrealistically simplistic) interventions that assume a major health hazard can be removed, or SES shifted for the entire population. This simulation accounts consistently for co-morbidities and competing hazards over the life course. The purpose of this exercise is to demonstrate the feasibility of using our modeling approach for policy applications when the models pass the causality tests described in next section. Finally, Section 8 gives conclusions and outlines topics for future research.

2. Association and Causality in Panel Data

2.1. Testing Causality. The primary purpose of this study is to test for direct causal links between SES and health. There is a large literature on the nature of causality and the interpretation of "causality tests".[8] Our analysis fits generally within the approach of Granger (1969), Sims (1972), and Hoover (2001), but our panel data structure permits some refinements that are not available in a pure time series setting.

Let Y_t denote a K-vector of demographic, health, and socioeconomic random variables for a household at date t, and interpret a realization of these variables as an observation in one wave of a panel survey. Let $\boldsymbol{Y}_t$ be the information set containing the history of this vector through date t. Let

$$f(Y_t|\boldsymbol{Y}_{t-1}) \equiv f_1(Y_{1t}|\boldsymbol{Y}_{t-1})\cdot f_2(Y_{2t}|Y_{1t},\boldsymbol{Y}_{t-1})\cdot\ldots\cdot f_K(Y_{Kt}|Y_{1t},\ldots,Y_{K-1,t},\boldsymbol{Y}_{t-1}) \tag{1}$$

denote a model of the conditional distribution of Y_t given $\boldsymbol{Y}_{t-1}$. Without loss of generality, we have written this model as a product of one-dimensional conditional distributions, given history and given components of Y_t determined previously. Writing the model in this way does not imply that the components of Y_t form a causal chain, as they may be simultaneously determined, or determined in some causal sequence other than the specified sequence. However, the model structure simplifies if the current components of Y_t in the specified order do form a causal chain or are conditionally independent. If one takes Wold's view that causal action takes time, then for sufficiently brief time intervals, $f_K(Y_{Kt}|Y_{1t},\ldots,Y_{K-1,t},\boldsymbol{Y}_{t-1})$ will not depend on contemporaneous variables, and what Granger calls "instantaneous causality" is ruled out. In practice, time aggregation to observation intervals can introduce apparent simultaneous determination. Conversely, in applications where time aggregation is an issue, one can treat observed variables as indicators for some latent causal chain structure defined for very short time intervals.

We shall focus on first-order Markov processes, specializations of (1) in which only the most recent history conveys information,

$$f(Y_t|\boldsymbol{Y}_{t-1}) \equiv f(Y_t|Y_{t-1}) \equiv f_1(Y_{1t}|Y_{t-1})\cdot f_2(Y_{2t}|Y_{1t},Y_{t-1})\cdot\ldots\cdot f_K(Y_{Kt}|Y_{1t},\ldots,Y_{K-1,t},Y_{t-1}). \tag{2}$$

Note that if (1) is a higher-order Markov process, then (2) can be obtained by expanding the variables in Y_t to include higher-order lags. Greater generality could be achieved via a hidden Markov structure in which the observed Y_t are deterministic functions of a latent first-order Markov process.[9] We leave this extension for future research.

The model (2) is *valid* for a given history Y_{t-1} if it is the true conditional distribution of Y_t given this history. Term *f* a *structural* or *causal* model, or a (*probabilistic*) *law*, for Y_t *relative* to a family of histories if it has the *invariance* property that it is valid for each history in the family. Operationally, this means that within specified domains, *f* has the *transferability* property that it is valid in different populations where the marginal distribution of Y_{t-1} changes, and the *predictability* or *invariance under treatments* property that it remains valid following policy interventions that alter the marginal distribution of Y_{t-1}. By including temporal or spatial variables in Y, it is possible to weaken invariance requirements to fit almost any application. Done indiscriminately, this creates a substantial risk of producing an "over-fitted" model that will be invalid for any "out-of-sample" policy interventions. Then, proposed models should be as generic as possible. However, it may be necessary in some applications to model "regime shifts" that account for factors that are causal for some populations or time periods, and not for others.

Suppose the vector $Y_t = (H_t,S_t,X_t)$ is composed of subvectors H_t, S_t, and X_t, which will later be interpreted as health conditions, SES status, and strictly exogenous variables, respectively. We say that S is *conditionally non-causal* for H, given X, if $f(H_t|H_{t-1},X_{t-1})$ is a valid model; i.e., given H_{t-1} and X_{t-1}, knowledge of S_{t-1} is *not needed* to achieve the invariance properties of a causal model. Conversely, if $f(H_t|H_{t-1},S_{t-1},X_{t-1}) \neq f(H_t|H_{t-1},X_{t-1})$, then knowledge of S_{t-1} *contributes to the predictability* of H_t. Note that either one or both conditional non-causality of S for H and conditional non-causality of H given S may hold. If either holds, then H and S can be arrayed in a (block) causal chain, and if both hold, then H and S are conditionally independent. Writing the model (2) as a product of univariate conditional probabilities $f_i(H_{it}|H_{1t},...,H_{i-1,t},H_{t-1},S_{t-1},X_{t-1})$, one can test for conditional non-causality of S for each component H_i. It is possible to have a causal chain in which S is conditionally causal to a previous component of H, and this component is in turn "instantaneously" causal to H_i, yet there is no direct causal link from S to H_i. Placing S after H in the vector Y, we have conditional probabilities $f(S_t|H_t,H_{t-1},S_{t-1},X_{t-1})$. There may be instantaneous conditional independence of H and S, with the conditional distribution of S_t not depending on H_t, or conditional non-causality of H for S, with the conditional distribution not depending on H_{t-1}, or both. The statement that X is *strictly exogenous* in a valid model (2) is equivalent to the condition that H and S are conditionally non-causal for X in this model.[10]

The conventional definition of a *causal model* or *probabilistic law* requires that *f* be valid for the universe of possible histories (except possibly those in a set that occurs with probability zero); see Pearl (2000). It is possible to reject statistically a proposed causal model by showing that it is highly improbable that an observed sample with a given history was generated by this model. It is far more difficult using statistical analysis to conclude inductively that a proposed model is valid for the universe of possible histories. We have the far more limited objective of providing a foundation for policy analysis, where it is the invariance property under policy interventions that is crucial to predicting policy consequences. We have defined *validity* and *non-causality* as properties of a *model*, and of the outcomes of a *process* of statistical testing that could in principle be conducted on this model. Only within the domain where the model is valid, and invariance confirms that the model is accurately describing the true data generation process, can these limited positivistic model properties be related to the causal structure embedded in the true data generation process. Further, we can choose the domain over which invariance will be tested to make the definition operational and relevant for a specific analysis of policy interventions. Similarly, our definition of conditional non-causality is a positivistic construct in the spirit of the purely statistical treatment of "causality" by Granger (1969), and the test we will use is simply Granger's test for the absence of causality, augmented with invariance conditions. Thus for example, if our analysis using this framework concludes that SES is not conditionally causal for new

health events within the domain where the Medicare system finances and delivers health care, then this finding would support the conclusion that policy interventions in the Medicare system to increase access or reduce out-of-pocket medical expenses will not alter the conditional probabilities of new health events, given the health histories of enrollees in this system. It is unnecessary for this policy purpose to answer the question of whether the analysis has uncovered a causal structure in any deeper sense. Econometric analysis is better matched to the modest task of testing invariance and non-causality in limited domains than to the grander enterprise of discovering universal causal laws. However, our emphasis on invariance properties of the model, and on tests for Granger causality within invariant families, is consistent with the view of philosophers of science that causality is embedded in "laws" whose validity as a description of the true data generation process is characterized by their invariance properties; see Pearl (2000), Nozick (2001).

. **2.2. Some Specific Formulations**. Starting from the class (2), we consider operational models of the linear latent variable form

$$Y^*_{it} = Y_{1,t}\alpha_{1i} + ... + Y_{i-1,t}\alpha_{i-1,i} + Y_{t-1}'\beta_i + \delta_i - \sigma_i\varepsilon_{it}, \tag{3}$$

with

$$Y_{it} = \psi_i(Y^*_{it}, Y_{t-1}), \tag{4}$$

where Y^* is a latent variable, ε_{it} is an unobserved disturbance that is standard normal and independent across i and t, and ψ is a partial observability mapping that depends on the latent variable, and possibly on the lagged variables. For example, for a chronic health condition such as diabetes, Y_{it} will indicate whether there has ever been a diagnosis of the disease, with $Y_{it} = \max\{Y_{i,t-1}, \mathbf{1}(Y^*_{it} \geq 0)\}$. For an acute condition such as a heart attack, $Y_{it} = \mathbf{1}(Y^*_{it} \geq 0)$ indicates a new occurrence. Components of Y may be binomial or ordered discrete variables such as health status, or continuous variables such as household income. In this model, the α's, β, δ, and σ are parameters; restrictions are imposed as necessary for identification. In (3), the linearity in variables and parameters, the first-order Markov property, and the triangular dependence of Y^*_{it} on previous components of Y_t are not, in themselves, particularly restrictive, as one can approximate any continuous Markov model of the form (2) by a form (3) in which Y_t is expanded to include transformations and interactions to sufficient order. The normality assumption is also not restrictive in principle. A latent random variable with conditional CDF $F_1(Y^*_{1t}|Y_{t-1})$ and the partial observability transformation (4) can be redefined using the standard normal CDF Φ as $Y^{**}_{1t} = \Phi^{-1}(F(Y^*_{1t}|Y_{t-1}))$ and $Y_{1t} = \psi_1(F^{-1}(\Phi(Y^{**}_{1t})|Y_{t-1}), Y_{1,t-1})$; this gives a version of the model (3)-(4) in which the disturbance is standard normal.[11] The same construction can be applied to the remaining components of Y_t. The causal chain assumption is innocuous when the time interval is too short for most causal actions to operate, and the components of Y_t are conditionally independent. However, the causal chain assumption is much more substantive restriction when the time interval is long enough so that multiple events have time to occur, as it rules out even the feedbacks that would appear in multiple iterations of a true causal chain. Of course, the generally non-restrictive approximation properties of the model (3)-(4) do not imply that a particular specification chosen for an application is accurate, and failures of tests for invariance can also be interpreted as diagnostics for inadequate specifications.

In the model (3)-(4), a binomial component i of Y_t with the partial observability mapping $\max\{Y_{i,t-1}, \mathbf{1}(Y_{it}^* \geq 0)\}$ and the identifying restriction $\sigma_i = 1$ satisfies $Y_{it} = 1$ if $Y_{i,t-1} = 1$, and otherwise is one with the probit probability

$$f_i(Y_{it} = 1|Y_{1t},...,Y_{i-1,t},Y_{t-1}) = \Phi(Y_{1,t}\alpha_{1i} + ... + Y_{i-1,t}\alpha_{i-1,i} + Y_{t-1}'\beta_i + \delta i). \quad (5)$$

Analogous expressions can be developed for ordered or continuous components.

2.3. Measurement Issues. A feature of the panel we use is that the waves are separated by several years and the interviews within a wave are spread over many months, with the months between waves differing across households. If the model (2) applies to short intervals, say months, then the transition from one wave in month t to another in month t+s is described by the probability model

$$f(Y_{t+s}|Y_t) = \sum_{Y_{t+1},...,Y_{t+s-1}} f(Y_{t+1}|Y_t)\cdot...\cdot f(Y_{t+s}|Y_{t+s-1}). \quad (6)$$

Direct computation of these probabilities will generally be intractable, although analysis using simulation methods is possible.

A major additional complication in our panel is that interview timing appears to be related to health status, with household or proxy interviews delayed for individuals who have died or have serious health conditions. This introduces a spurious correlation between apparent time at risk and health status that will bias estimation of structural parameters. To study empirical approximations to (6) and corrections for spurious correlation, we consider a simple model of interview delay. Let $p = \Phi(\alpha+\beta x)$ denote the monthly survival probability for an individual who was alive at the previous wave interview, where x is a single time-invariant covariate that takes the values -1, 0, +1, each with probability 1/3. Counting from the time of the previous wave interview, let k denote the number of months this individual lives, and c denote the month that interviews begin for the current wave.[12] There is a distribution of initial contact times; let q denote the probability of a month passing without being contacted, and let m denote the month of initial contact. Assume that an individual who is living at the time of initial contact is interviewed immediately, but for individuals who have died by the time of initial contact, there is an interview delay, with r denoting the probability of an additional month passing without a completed interview with a household member or proxy. Let n denote the number of months of delay in this event. Assume that m and n are not observed, but the actual inter-wave interval t, equal to m if the individual is alive at time of initial contact, and equal to m+n otherwise, is observed. The density of k is $p^{k-1}(1-p)$ for $k \geq 1$. The density of m is $q^{m-c}(1-q)$ for $m \geq c$. Let d be an indicator for the event that the individual is dead at the time of initial contact. The probability of t and d = 0 is $h(0,t) = q^{t-c}(1-q)p^t$, the product of the probability of contact at t and the probability of being alive at t. The probability of t and d = 1, denoted h(1,t), is the sum of the probabilities that the individual is dead at an initial contact month m, with $c \leq m \leq t$, and the subsequent interview delay is n = t-m, or

$$h(1,t) = \sum_{m=c}^{t} (1 - p^m)q^{m-c}(1-q)r^{t-m}(1-r). \quad (7)$$

If $r < pq$, then $h(1,t) = (1-q)(1-r)\{(q^{t-c+1} - r^{t-c+1})/(q-r) - p^c((pq)^{t-c+1} - r^{t-c+1})/(pq-r)\}$. Then, the probability of an observed inter-wave interval t is $h(t) = h(0,t) + h(1,t)$, and the conditional probability of $d = 1$, given t, is $P(1|t) = h(1,t)/h(t)$. Absent interview delay, the conditional probability of $d = 1$ given t would be simply $p^{t-1}(1-p)$. The parameter values $\alpha = -2.47474$, $\beta = 0.3$, $c = 22$, $q = 0.85$, and $r = 0.5$ roughly match our panel. For these values, the median inter-wave interval is 25.5 months, and at $t = 34$, 87 percent of the interviews have been completed.

Figure 2 plots the inverse normal transformations of the true death rate $p^{t-1}(1-p)$ and the apparent death rate $P(1|t)$ against $\log(t)$ for each value of the covariate x. The *true* relationship is to a reasonable empirical approximation linear in x and in $\log(t)$. Then, in the absence of interview delay, one could approximate p, given x, with reasonable accuracy by estimating a probit model for death of the form $\Phi(\theta + \gamma x + \lambda \log(t))$, and then estimating p using the transformation $p = (1 - \Phi(\theta + \gamma x + \lambda \log(t)))^{1/t}$ for the observed inter-wave interval t.[13] However, the figure shows that interview delay induces a sharp gradient of apparent mortality hazard with inter-wave interval, so that an estimated model will not extrapolate to realistic mortality hazards over shorter periods. A simple imputation of time at risk up to initial contact leads again to models that work well with the procedure just outlined for estimation of p conditioned on x. A simple imputed time of initial contact for those who have died is the observed inter-wave interval less the difference in the mean inter-wave interview times for dead and living respondents. This imputation can be adjusted further so that the extrapolated annual death rate, $\Phi(\theta+\lambda\log(12)+\gamma x)$, matches the sample average mortality rate. We do the additional adjustment for our panel, with results that are almost identical to the simple imputation of time of initial contact.

We conducted a Monte Carlo calculation of the approximations above in a sample of 50,000. In this simulation, the empirical approximation to observed mortality in the absence of interview delay is $\Phi(-3.1053+0.5327\cdot x+0.652\cdot\log(t))$. With interview delay and the simple imputation described above, $\Phi(-2.9900+0.5349\cdot x+0.6174\cdot 1og(t))$ is the empirical approximation.[14] Table 1 gives the annual mortality rates implied by these approximations. From these results, we conclude first that in the absence of interview delay, the probit model $\Phi(\theta+\lambda\log(t)+\gamma x)$ provides an adequate approximation to exact annual mortality rates as a function of time at risk, across values of the covariate x that substantially change relative risk, and second that this remains true in the presence of interview delay when one imputes the initial contact time for dead subjects.

These conclusions on the accuracy of the approximation should extend to the exact inter-wave transition probabilities (6) in our Markov model, supporting use of the probit approximation

$$f_i(Y_{i,t+s}=1|Y_{1t+s},...,Y_{i-1,t+s},Y_t) = \Phi(Y_{1,t+s}\alpha_{1i} + ... + Y_{i-1,t+s}\alpha_{i-1,i} + Y_t'\beta_i + \delta_i + \lambda_i\cdot\log(s)) \quad (8)$$

where $s = t_{i2} - t_{i1}$ is the imputed months between initial contact for a wave and previous wave interview, for estimation of incidence between waves. For simulation of yearly transitions, we use the approximation

$$f_i(Y_{i,t+12}=1|Y_{1t+12},...,Y_{i-1,t+12},Y_t) = 1 - (1 - \Phi(Y_{1,t+1}\alpha_{1i} + ... + Y_{i-1,t+1}\alpha_{i-1,i} + Y_t'\beta_i + \delta_i + \lambda_i\cdot\log(s)))^{12/s}$$

$$\approx \Phi(Y_{1,t+1}\alpha_{1i} + ... + Y_{i-1,t+1}\alpha_{i-1,i} + Y_t'\beta_i + \delta_i + \lambda_i\cdot\log(s))\cdot 12/s, \quad (9)$$

where s is the median inter-wave interval, and the final approximation holds when the probability of a transition is small. Formula (9) generalizes to any probability of a transition from the status quo, with the probability of remaining at the initial state defined so that all the transition probabilities sum to one.

We expect this formula to approximate well the probabilities of no new health conditions in a sample population over periods corresponding to the observed inter-wave intervals.

Estimation of models based on (2)-(8) is straightforward. Because of the independence assumption on the disturbances and the absence of common parameters across equations, the estimation separates into a probit, ordered probit, or ordinary least squares regression for each component of Y, depending on whether the partial observability mapping is binary, ordered, or linear. Conventional likelihood ratio tests can be used for the significance of explanatory variables.

3. The AHEAD Panel Data

3.1. Sample Characteristics. Our data come from the *Asset and Health Dynamics among the Oldest-Old* (AHEAD) study.[15] This is a panel of individuals born in 1923 or earlier, and their spouses. At baseline in 1993 the AHEAD panel contained 8222 individuals representative of the non-institutionalized population, except for over-samples of blacks, Hispanics and Floridians. Of these subjects, 7638 were over age 69; the remainder were younger spouses. There were 6052 households, including individuals living alone or with others, in the sample. The Wave 1 surveys took place between October 1993 and August 1994, with half the total completed interviews finished before December 1993. The Wave 2 surveys took place approximately 24 months later, between November 1995 and June 1996, with half the total completed interviews finished by the beginning of February 1996. The Wave 3 surveys took place approximately 27 months after that, between January 1998 and December 1998, with half the total completed interviews finished near the beginning of March 1998. In each wave, there was a long but thin tail of late interviews, heavily weighted with subjects who had moved, or required proxy interviews due to death or institutionalization. Subjects never interviewed, directly or by proxy, are excluded from the calculation of the distribution of interview months. AHEAD is a continuing panel, but it has now been absorbed into the larger Health and Retirement Survey (HRS), which is being interviewed on a three-year cycle.

The AHEAD panel has substantial attrition, with death being the primary but not the only cause. A significant effort has been made to track attritors, and identify those who have died through the National Death Register. Figure 3 describes outcomes for the full age-eligible sample. For subjects where a proxy interview was possible, an "exit interview"gives information on whether decedents had a new occurrence of cancer, heart attack, or stroke since the previous wave. From the 6743 age-eligible individuals who did not attrit prior to death, we formed a working sample for analysis consisting of 6489 by excluding 254 additional individuals with critical missing information. Figure 4 describes their outcomes. In a few cases, attritors in Wave 2 rejoined the sample in Wave 3, but we treat these as permanent attritors because the missing interview makes the observation unusable.

The restriction of the AHEAD panel to the non-institutionalized elderly in Wave 1 selects against those with the highest mortality risk, particularly at the oldest ages, but the impact of this selection attenuates over time. For white females, Figure 5 compares the observed annual mortality rate in the AHEAD panel with the expected annual mortality rate from the 1997 Life Tables for the United States (U.S. Census, 1998).[16] Between Waves 1 and 2, the AHEAD mortality risk is substantially below the life table for ages above 75, reflecting the selection effect of non-institutionalization. Between Waves 2 and 3, this effect has essentially disappeared. There is a persistent divergence of the mortality risks above age 90. In this range, the AHEAD data is sparse, so the curve is imprecisely determined. However, the life tables derived from historical mortality experience may overstate current mortality risk

at advanced ages. Figure 6 makes the same mortality experience comparisons for the full AHEAD working sample, and draws the same conclusions.[17]

3.2. Descriptive Statistics. The AHEAD survey provides data on health and socioeconomic status, as well as background demographics. A list of the health conditions we study, with summary statistics, is given in Table 2. A list of the socioeconomic conditions and demographic variables we use is given in Table 3. Appendix Table A1 lists the variable transformations used in our statistical analysis.[18] In setting up causality tests, using the framework set out in Section 2, we will use the health conditions, followed by the socioeconomic conditions, in the order given in these tables. We list cancer, heart disease, and stroke first because they may be instantaneously causal for death, and because we have information from decedent's exit interviews on new occurrence of these diseases. We group the remaining health conditions by degenerative and chronic conditions, then accidents, then mental conditions, since if there is any contemporaneous causality, it will plausibly flow in this order. Similarly, if there is contemporaneous causality between health and socioeconomic conditions, it plausibly flows from the former to the latter.

3.3. Constructed Variables. The collection and processing of some of the variables requires comment. AHEAD has an extensive battery of questions about health conditions, including mental health. Most health conditions are asked for in the form "Has a doctor ever told you that you had ...?" However, for cancer, heart disease, and stroke, subjects are also asked if there was a new occurrence since the previous interview, and for some condition such as arthritis, incontinence, and falls, the questions in wave 1 ask for an occurrence in the past 12 months. We note that there are some major groups of health conditions that were not investigated in AHEAD: degenerative neurological diseases, kidney and liver diseases, immunological disorders other than arthritis, sight and hearing problems, back problems, and accidents other than falls. The BMI index is calculated from self-reported height and weight. Information is collected on the number of ADL limitations, for six activities of daily living, and on the number of IADL limitations, for five instrumental activities of daily living. A high ADL limitation count indicates that the individual has difficulty with personal self-care, while a high IADL limitation count indicates difficulty in household management. The study collects data on self-assessed health status, where the subject is asked to rate his or her health as excellent, very good, good, fair or poor. We use an indicator for a Poor/Fair response. No reference is made to other groups such as "people your age." The study contains the CESD battery of questions measuring general mood; and from this we form an indicator for depression.

AHEAD is linked to Medicare records. There is insufficient detail to permit reconciliation of self-reports on objective health conditions against diagnoses in the medical records, but errors in self-reports are an issue. We find small, but significant, inconsistencies across waves of AHEAD in reported chronic conditions. A study of Canadian data for a younger population finds substantial discrepancies between self-reported conditions and diagnoses from medical records, particularly for chronic conditions such as arthritis; see Baker, Stabile, and Derl (2001). These authors also find support for a "self-justification" hypothesis that non-workers are more likely to make false positive claims for health conditions. If this reporting behavior carries into old age, then the reduction in SES as a consequence of spotty employment would induce an artifactual positive association of self-reported health conditions and SES.

The study measures cognition using a battery of questions which test several domains (Herzog and Wallace, 1997): learning and memory are assessed by immediate and delayed recall from a list of 10 words that were read to the subject; reasoning, orientation and attention are assessed from Serial 7's,

counting backwards by 1 and the naming of public figures, dates and objects.[19] This score reflects both long-term native ability and health-related impairments due to health events. We carry out the following statistical analysis to reduce the effect of native ability so that we can concentrate on health-related loss of cognitive function. First, we analyze a "non-impaired" sample of younger individuals, born between 1942 and 1947, who were administered the same cognitive battery in the 1998 Health and Retirement Survey as were the AHEAD subjects. For these younger individuals, where health-related impairment of cognitive function is rare, we carry out a LAD regression of the cognitive score on education level, sex, and race. We use this fitted regression to predict a "baseline" non-impaired cognitive score for each member of the AHEAD sample. An additional adjustment is required because average education levels were rising rapidly early in the Twentieth Century, due to changes in child labor laws and introduction of compulsory education. We assign each AHEAD subject a "1923 cohort equivalent" education level by first regressing education on sex, race, and birth cohort, using a specification search to find interactions and non-linearities, and then adding to their actual years of education the difference in the mean years of education for their sex-race cohort and the corresponding 1923 sex-race cohort. We then calculate for each AHEAD sample member the deviation of their cognitive score from this adjusted baseline. As a normalization, we assign a threshold such that fifteen percent of AHEAD subjects aged 70-74 in Wave 1 fall below the threshold. We then use the same threshold in other age groups and other waves to define an indicator for cognitive impairment.[20]

3.4. Measurement of Wealth. AHEAD individuals and couples are asked for a complete inventory of assets and debts, and about income sources. Subjects are asked first if they have any assets in a specified category, and if so, they are asked for the amount. A non-response to the amount is followed by unfolding bracket questions to bound the quantity in question, and this may result in complete or incomplete bracket responses. Through the use of unfolding brackets, full non-response to asset values was reduced to levels usually less than 5 percent, much lower than would be found in a typical household survey. Generally, median responses among full respondents for an asset category are comparable to other economic surveys, such as the Survey of Consumer Finance. However, changes in reported assets between waves contain outliers that suggest significant response errors between waves. For couples, where both members are asked the questions on assets, there is also substantial inter-subject response variation. It is possible that these repeated reports could be used to control statistically for response error in couples. However, there are systematic differences between respondents, and we use the asset responses only from the individual that a couple says manages the household finances. There may also be an issue of bias in responses recovered by unfolding brackets. Hurd *et al* (1997) used experimental variation in the bracket sequences for two financial questions on Wave 2 of AHEAD, and found that anchoring to the bracket quantities was significant.

For complete or incomplete bracket responses in an asset category, we impute continuous quantities using hot deck methods. In Wave 1, if information on ownership of an asset is missing for a subject, but this subject does give ownership status in Wave 2, then we impute Wave 1 ownership by drawing from the conditional empirical distribution of those who have the same response in Wave 2 and give a response in Wave 1. For subjects missing ownership in both Waves 1 and 2, we draw an ownership pair from the empirical distribution of ownership pairs for those giving responses. Given ownership and complete or incomplete bracket information, we draw from the empirical distribution of Wave 1 continuous responses that are consistent with the subject's bracket. In later waves, we have adopted a first-order Markov cross-wave hot deck imputation procedure that assigns a continuous quantity within the given response bracket. First, missing ownership is imputed by choosing randomly from

respondents, conditional on ownership in the previous wave. Then, given ownership, we impute a quantitative *change* in the item from the previous wave by drawing from the empirical distribution of subjects with complete responses that fall in the corresponding brackets in the current and the previous wave. This assures that imputed changes will have the same empirical distribution as observed changes, given the conditioning information available. This procedure does not revise previous wave imputations, so analyses based on earlier waves are not affected. We have experimented with cross-item imputation methods, where bracket information on some asset categories would be used to refine the conditioning used in the imputation of other asset categories. We have found that this has very little effect on the imputed variables or on the results obtained from analyses that use these variables. Therefore, we carry out all imputations one item at a time.

Measured wealth is accumulated over eleven asset categories, including imputed items. We distinguish *liquid wealth*, composed of IRA balances, stocks, bonds, checking accounts, certificates of deposit, less debt, and *non-liquid wealth*, composed of net homeowner equity, other real estate, vehicles and other transportation equipment, businesses, and other assets. The variation in reported wealth of AHEAD households by asset category is substantial from wave to wave, suggesting that in addition to real volatility and reallocation of wealth portfolios, there are serious reporting problems with assets. Values of businesses owned and real estate are problematic items, since current market valuations may be unavailable to respondents, and subjective valuations may be unreliable. Suppose that W_t is measured wealth of household in wave t, in 1997 dollars, and that $W_t = W^*_t + \eta_t$, where W^*_t is true wealth and η_t is reporting error. To minimize the impact of extreme outliers in wealth and wealth changes, which we believe are a particular problem due to gross reporting errors, our statistical analysis will use bounded transformations of measured wealth.

The equations of motion for real wealth satisfy $dW_i^*/dt = rW_i^* + S_i$, where i = T,N,L indexes total wealth or its non-liquid and liquid components, r is the instantaneous real rate of return, including unrealized capital gains, and S_i is the flow of savings to the wealth component. Make the logistic transformation $Z_i = 1/(1+\exp(-c_i W_i+d_i))$, where c_i and d_i are chosen so that in AHEAD wave 1 the median and the semi-interquartile range of Z_i are one-half. Then, Z_i is a monotone transformation of measured wealth that is less sensitive to extremes. The equation of motion for Z_i is $dZ_i/dt - r\cdot Z_i(1-Z_i)\cdot(\log(Z_i/(1-Z_i))+d) = c\cdot Z_i(1-Z_i)\cdot S + c\cdot Z_i(1-Z_i)\cdot(d\eta/dt - r\eta)$. We assume that over an inter-wave interval, this equation of motion can be approximated by

$$\frac{Z_{it} - Z_{i,t-1} - R_{t-1}Z_{i,t-1}(1-Z_{i,t-1})\cdot(\log(Z_{i,t-1}/(1-Z_{i,t-1}))+d_i)}{m_t} = S_{it}^{\#} + \sigma\nu_{it}, \qquad (10)$$

where t indexes the wave, m_t is the interval in months between waves t-1 and t, R_t is the S&P real rate of return over the given interval, $S_{it}^{\#}$ is the measured part of $c_i\cdot Z_{i,t-1}(1-Z_{i,t-1})\cdot S_{i,t-1}$, attenuated at extreme values of $Z_{i,t-1}$, and the disturbance ν_{it} includes the measurement error $c_i\cdot Z_{i,t-1}(1-Z_{i,t-1})\cdot((\eta_{it}-\eta_{i,t-1})/m_{t-1}-r\eta_{t-1})$ and the unmeasured part of $c_i\cdot Z_{i,t-1}(1-Z_{i,t-1})\cdot S_{i,t-1}$. We assume ν is homoskedastic. This is consistent with a measurement error in observed wealth that is heteroskedastic, with gross measurement errors more likely when true wealth is near its extremes.[21] The disturbance in (10) may be serially correlated; however, we have not incorporated this into our analysis. The effect of the transformation is to substantially reduce the influence of outliers in the distribution of changes in measured wealth. In application, we specify $S_{it}^{\#}$ to be a linear function of transformed nonliquid and liquid wealth, $Z_{T,t-1}(1-Z_{T,t-}$

$_1)\log(Z_{N,t-1}/(1-Z_{N,t-1}))$ and $Z_{T,t-1}(1-Z_{T,t-1})\log(Z_{L,t-1}/(1-Z_{L,t-1}))$, and of other SES, demographic, and health variables, scaled by $Z_{T,t-1}(1-Z_{T,t-1})$

3.5. Mortality and Observed Wealth Change. A problem with the analysis of wealth changes is that terminal wealth is not observed following the death of a single, or the death of both members of a household, introducing a selection effect. A second problem is that a household death may have a direct impact on the wealth of a survivor, due to the expenses associated with a death and the disposition of the estate. There are also severe wealth measurement problems following a household death, since a death typically requires a valuation of assets, and in many cases changes the financially responsible respondent. For this reason, we will analyze separately wealth changes for singles and for couples, allow a regime shift following the death of one member of a couple, and account for the selection that occurs when there are no survivors.

For a single female, we adopt a bivariate selection model,

$$Y^*_{ft} = Y_{f,t-1}\beta_f + \varepsilon,\ y_{ft} = \mathbf{1}(Y^*_{ft} > 0),\ Y_{wt} = Y_{w,t-1}\beta_w + Y'_{w,t}\gamma_w + \lambda\varepsilon + \kappa\eta,\ Y_{wt} \text{ observed if } y_{ft} = 1 \qquad (11)$$

where the first latent equation determines survival, $y_{ft} = 1$, the second equation corresponds to the transformed wealth change equation (10) with dependence on the previous state $Y_{w,t-1}$ and the previously determined components $Y'_{w,t}$ of the current state, with the wealth change observed for survivors. The disturbance ε_f has mean zero, variance one, and a density $f(\varepsilon)$. The disturbance η is independent of ε, and has mean zero and variance one. The correlation of the disturbances in the selection and wealth change equations is $\rho = \lambda/(\lambda^2+\kappa^2)^{½}$, and the unconditional variance of the wealth change equation is $\sigma^2 = \lambda^2+\kappa^2$. When ϵ and η are standard normal, this is the conventional bivariate normal selection model. However, specification tests for normality fail, and for robustness we adopt a more flexible specification, approximating the density $f(\varepsilon)$ by an Edgeworth expansion,

$$f(\varepsilon) = \sum_{j=0}^{J} \gamma_j H_j(\varepsilon)\varphi(\varepsilon), \qquad (12)$$

where the γ_j are parameters and $H_j(\varepsilon)$ are Hermite orthogonal polynomials. Let $\Psi_{jk}(a) = \int_a^\infty \varepsilon^k H_j(\varepsilon)\varphi(\varepsilon)d\varepsilon$. Then, the polynomials $H_j(\varepsilon)$ and the partial moment functions $\Psi_{jk}(a)$ can be constructed using the recursions

$$\begin{aligned} &H_0(\varepsilon) = 1,\ H_1(\varepsilon) = \varepsilon, \text{ and } H_j(\varepsilon) = \varepsilon H_{j-1}(\varepsilon) - (j-1)H_{j-2}(\varepsilon)/d\varepsilon \text{ for } j > 1, \\ &\Psi_{00}(a) = \Phi(-a),\ \Psi_{01}(a) = \varphi(a), \text{ and } \Psi_{0k}(a) = a^{k-1}\varphi(a) + (k-1)\Psi_{0,k-2}(a) \text{ for } k > 1, \\ &\Psi_{j0}(a) = H_{j-1}(a)\varphi(a) \text{ and } \Psi_{jk}(a) = a^k H_{j-1}(a)\varphi(a) + k\Psi_{j-1,k-1}(a) \text{ for } k > 0, \text{ for } j > 0. \end{aligned} \qquad (13)$$

Appendix Table A1 derives these results and gives the leading terms for H_j and Ψ_{jk}: We require that f integrate to one and have unconditional mean zero and variance one; this forces $\gamma_0 = 1$, $\gamma_1 = 0$, and $\gamma_2 = 0$. The free parameters γ_j for $j > 2$ determine higher-order moments of ε. For example, skewness and kurtosis are determined by $E\varepsilon^3 = 6\gamma_3$ and $E\varepsilon^4 = 3+24\gamma_4$. With these restrictions, we have, finally

$$E(\varepsilon|\varepsilon>a) = \frac{\phi(a)+\sum_{j=3}^{J}\gamma_j\Psi_{j1}(a)}{\Phi(-a)+\sum_{j=3}^{J}\gamma_j\Psi_{j0}(a)} \text{ and } E(\varepsilon^2|\varepsilon>a) = \frac{\Phi(-a)+a\phi(a)+\sum_{j=3}^{J}\gamma_j\Psi_{j2}(a)}{\Phi(-a)+\sum_{j=3}^{J}\gamma_j\Psi_{j0}(a)} \quad . \tag{14}$$

We use the Edgeworth approximation and these conditional expectations with J = 4. We then have

$$E(Y^*_{wt}|\varepsilon>-Y_{f,t-1}\beta_f) = Y_{w,t-1}\beta_w + Y'_{w,t}\gamma_w + \lambda \frac{\phi(Y_{f,t-1}\beta_f)+\sum_{j=3}^{4}\gamma_j\Psi_{j1}(-Y_{f,t-1}\beta_f)}{\Phi(Y_{f,t-1}\beta_f)+\sum_{j=3}^{4}\gamma_j\Psi_{j0}(-Y_{f,t-1}\beta_f)} \tag{15}$$

We estimate this conditional expectation in a two-step procedure. First, the parameters β_f of the selection equation are estimated by maximum likelihood, and substituted into the expression (15) for the expectation of the wealth change equation.[22] Then, the parameters in this conditional expectation are estimated using nonlinear least squares. The disturbance $\zeta = \lambda\varepsilon + \kappa\eta - E(|\varepsilon>-Y_{f,t-1}\beta_f)$ has mean zero and variance

$$E(\zeta^2|\varepsilon>-Y_{f,t-1}\beta_f)) = \kappa^2 + \lambda^2 \left(\frac{\sum_{j=0}^{J}\gamma_j\Psi_{j2}(-Y_{f,t-1}\beta_f)}{\sum_{j=0}^{J}\gamma_j\Psi_{j0}(-Y_{f,t-1}\beta_f)} - \left[\frac{\sum_{j=0}^{J}\gamma_j\Psi_{j1}(-Y_{f,t-1}\beta_f)}{\sum_{j=0}^{J}\gamma_j\Psi_{j0}(-Y_{f,t-1}\beta_f)}\right]^2 \right) \quad . \tag{16}$$

We regress the squared residuals from the estimation of (15) on the right-hand-side variables in (16) to obtain an estimate of κ^2. Finally, we estimate the covariance matrix for the parameter estimates using the generalized method of moments "sandwich" formula, with the Eicker-White procedure used for robustness against heteroskedasticity of unknown form, and the delta method used to incorporate the effects of variance in the first-stage selection parameter estimates.

Next consider the effects of death and selection on couples. We adopt a trivariate selection model with selection equations

$$Y^*_{ft} = Y_{f,t-1}\beta_f + \varepsilon_f, \; Y^*_{mt} = Y_{m,t-1}\beta_m + \varepsilon_m, \; y_{ft} = \mathbf{1}(Y^*_{ft}>0), \; y_{mt} = \mathbf{1}(Y^*_{mt} > 0), \tag{17}$$

for the female and male members of the couple, respectively, where ε_f and ε_m are assumed to be independent with zero mean and unit variance, and densities $g_f(\varepsilon_f)$ and $g_m(\varepsilon_m)$. The independence assumption could fail if there are hidden common factors in mortality risk for both household members; e.g., indirect effects of smoking. However, the frequency of multiple deaths in a household between waves is sufficiently rare in the AHEAD data so that mortality risk interactions are empirically not identified. We distinguish three regimes (y_f,y_m) in which wealth change is observed: intact couples where both members survive (11), the female survives the death her spouse (10), and the male survives the death of his spouse (01). We will let $Y^0_t = [Y_{ht-1}\; Y_{ft-1}\; Y_{mt-1}\; y_{ft}\cdot Y'_{ft}\; y_{mt}\cdot Y'_{mt}]$ denote the vector of variables that explain wealth change, where Y_{ht-1},Y_{ft-1}, and Y_{mt-1} are, respectively, previous wave

common, female, and male variables, Y'_{ft} are previously determined components of the current state for females, observed only for survivors and hence zeroed out for non-survivors, and Y'_{mt} are the analogous previously determined components of the current state for males, again observed only for survivors. We assume that in an observed regime jk the wealth change model takes the form

$$Y_{wt} = Y^0_t \beta_{wjk} + (\lambda_f - \theta_f \mathbf{1}(j+k=1))\varepsilon_f + (\lambda_m - \theta_m \mathbf{1}(j+k=1))\varepsilon_m + \kappa_{jk}\eta, \tag{18}$$

where η is independent of ε_f and ε_m, with zero mean and unit variance. For intact couples, unobserved dependence of wealth change on selection is reflected in the parameters λ_f and λ_m, a direct extension of the bivariate selection model to the trivariate case with two independent selection effects. For intact couples, the correlations of the disturbances in the selection and wealth change equations are $\rho_f = \lambda_f/(\kappa_{11}^2+\lambda_f^2+\lambda_m^2)^{1/2}$ and $\rho_m = \lambda_m/(\kappa_{11}^2+\lambda_f^2+\lambda_m^2)^{1/2}$. The coefficients β_{jk} and the standard deviation κ_{jk} are allowed to vary by regime, to capture the observed and unobserved economic effects of a death on valuation and reporting of assets. In addition, the selection effects are allowed to shift, from λ_f to λ_f-θ_f for the female selection disturbance, and from λ_m to λ_m-θ_m for the male selection disturbance. We incorporate these effects to accommodate an apparent interaction in which unexpected survival of a couple (e.g., ε_f and ε_m large positive) increases dissaving, perhaps due to additional medical and living expenses linked to overcoming high mortality hazards, but the unexpected death of a spouse (e.g., ε_f large negative) also increases dissaving, perhaps because revaluations of assets tends to be more drastic in circumstances where mortality hazard is low.

We adopt the Edgeworth approximation (12) for each of the densities $g_f(\epsilon_f)$ and $g_m(\epsilon_m)$. In regime jk with j+k > 0, letting $s_f = 2y_f-1$ and $s_m = 2y_m-1$, one has

$$\begin{aligned} E(Y_{wt}|jk) = {} & Y^0_t \beta_{wjk} + s_f(\lambda_f - y_f\theta_f)\frac{\phi(Y_{f,t-1}\beta_f) + \sum_{j=3}^4 \gamma_j \Psi_{j1}(-Y_{f,t-1}\beta_f)}{\Phi(s_f Y_{f,t-1}\beta_f) + s_f \sum_{j=3}^4 \gamma_j \Psi_{j0}(-Y_{f,t-1}\beta_f)} \\ & + + s_m(\lambda_m - y_m\theta_m)\frac{\phi(Y_{m,t-1}\beta_m) + \sum_{j=3}^4 \gamma_j \Psi_{j1}(-Y_{m,t-1}\beta_m)}{\Phi(s_m Y_{m,t-1}\beta_m) + s_m \sum_{j=3}^4 \gamma_j \Psi_{j0}(-Y_{m,t-1}\beta_m)} \quad . \end{aligned} \tag{19}$$

As in the case of singles, we estimate the conditional expectation (19) by nonlinear least squares after plugging in estimates of β_f and β_m from the earlier mortality models. A final stage, analogous to (16), regresses the squared residuals from (19) for each regime on an intercept, $E(\varepsilon_f^2|s_f) - (E(\varepsilon_f|s_f))^2$, and $E(\varepsilon_m^2|s_m) - (E(\varepsilon_m|s_m))^2$; the coefficient on the intercept is an estimate of κ_{jk}^2. Because the number of deaths among couples is relatively small, we impose the constraints $\beta_{w10} = \beta_{w01}$ for empirical identification.

Household income in AHEAD is also susceptible to measurement error, and to minimize its effect, we use income quartiles as explanatory variables. These are obtained by converting all incomes to 1997 dollars, determining the quartiles for the pooled incomes of all subjects in all waves, and using the thresholds thus established to classify each observed subject income. AHEAD respondents rate the safety of their neighborhood and the condition of their dwelling on a five-point scale, from poor to excellent; we use indicators for poor or fair responses.

4. SES and Prevalence of Health Conditions

4.1. Descriptive Statistics. We first give some descriptive statistics on the prevalence of health conditions in the AHEAD population. Table 4 shows prevalence rates in the AHEAD sample in Wave 1, classified by age and sex, for the health conditions listed in Table 2. Generally, prevalence of health conditions does not show a strong age gradient, indicating broadly that morbidity rates among survivors do not increase much with age. Selection effects from initial non-institutionalization and from mortality may be responsible. The major exception is cognitive impairment, which rises as age increases. The prevalence of acute and degenerative conditions among survivors fall after about age 80, reflecting the effect of selection due to deaths from these conditions. Males have higher prevalence of acute and degenerative diseases than do females, but females have higher prevalence of mental and chronic conditions, and accidents.

Figure 7 shows the age gradients of wealth, income, and education in wave 1 of the AHEAD sample. These gradients reflect substantial cohort effects, as well as life-cycle and composition effects. Work, income, and asset accumulation patterns of the AHEAD population were impacted by World War II, and those over age 80 experienced the Great Depression during their prime working years. The U.S. was substantially rural when the AHEAD population was born, and education was truncated for work for many members of this population. In addition to cohort effects, the curve for assets reflects life-cycle decumulation of assets through the retirement years, and the curve for income reflects the rising proportion of widows in the survivors to older ages. There is an additional compositional effect from the association of SES and mortality: higher SES is selected preferentially among survivors. However, in aggregate cross-section, the life cycle and cohort effects dominate the compositional effects.

4.2. Models of Association. To examine the association of SES and health conditions, we estimate a series of binomial probit models of the form

$$P(Y_{it}=1|Y_{1t},...,Y_{i-1,t},W_t,X_t), \tag{20}$$

where the Y_{it} are indicators for the prevalence of various health conditions, W_t denotes a vector of SES conditions, and X_t denotes other demographic variables. The health conditions appear in the same sequence as in Table 2, with previous conditions in (20) providing information on association among health conditions. Included in W_t are indicators for the top and bottom quartiles of wealth and income, indicators for 10 or more years of education, (high school) and for 14 or more years of education (college), and indicators for poor or fair neighborhood safety and dwelling condition.

The detailed estimation results are given for whites in Appendix Table A2. Sample sizes are not adequate for comparable models for non-whites. We find the expected patterns of co-morbidity, with a strong association of heart disease, stroke, lung disease, and arthritis, and a strong association of diabetes, heart disease, and high blood pressure. Incontinence is associated with cancer and stroke, and for women with diabetes, high blood pressure, and arthritis. Falls, hip fractures, and strokes are associated. Psychiatric diseases and depression are associated with arthritis and falls. BMI is positively associated with diabetes, high blood pressure, and arthritis, and negatively associated with lung disease. Current smokers have lower BMI, are less likely to have diabetes, and are more likely to be depressed. Numbers of ADL's and IADL's are positively associated with most acute diseases, arthritis, falls, hip fractures, cognitive impairment, and psychiatric disease. A poor or fair self-reported health status is associated with most acute and chronic diseases, with ADL's and IADL's, and with depression.

Some covariates are associated with health conditions, and may be risk factors for these conditions. Widowhood is associated with increased cancer and heart disease for women, increased psychiatric disease for men, and increased lung disease and depression for both men and women. For women, father's age at death is associated with heart disease, and mother's age at death is associated with high blood pressure. For men, father's age at death is associated with high blood pressure and arthritis.

We generally find a statistically significant association of SES and prevalence of health conditions, as summarized in Table 5. It is noteworthy that for males the prevalence of the acute diseases, cancer, heart disease, and stroke, are not strongly associated with SES, contrary to literature findings for younger populations. This may be the result of early onset of these diseases, particularly among the poor and among smokers, that selects out of the AHEAD population those males most at risk for these diseases.

Overall, wealth is the SES component most commonly associated with health conditions. Education, neighborhood rating, and dwelling rating are occasionally significant, and income is almost never significant. Table 6 summarizes the SES components that are individually significant in their association with various health conditions, and indicates the sign of the correlation. For a number of these conditions, prevalence rates are insufficient to detect the effects of SES components with satisfactory power. However, for heart disease, high blood pressure, arthritis, cognitive impairment, and self-rated health status, sample sizes should guarantee reliable indicators of association.

4.3. Relative Risk. To provide an indication of the direction and magnitude of the association of health conditions and SES, we calculate *relative risk* for low SES versus high SES, where the definition of low SES is bottom quartiles for income and wealth, less than a high school education, and a poor/fair neighborhood and dwelling, and the definition of high SES is top quartiles for income and wealth, a college education, and a good or better neighborhood and dwelling. Relative risk is defined as the AHEAD sample average of the ratio of the two probabilities, all other variables remaining at the observed levels for the subjects. Table 7 summarizes the relative risks for the various health conditions. Note that the prevalence models are describing only association, not causation, so that relative risk numbers cannot be interpreted causally. With the statistically insignificant exception of cancer and psychiatric conditions for females, high SES is associated with lower prevalence. Thus, we confirm in the AHEAD population the literature findings of a systematic association of SES with mortality and morbidity risk, and show that this association extends across a variety of acute, degenerative, chronic, and mental health impairments.

5. Incidence of Health Conditions and Tests for Causality in the AHEAD Panel

5.1. Models of Incidence. Following the format described in Section 2, we use the incidence of new health problems (or recurrence of cancer, heart disease, stroke, incontinence, falls, and hip fractures), conditioned on initial demographic, health, and SES status, to test for the absence of direct causal pathways. We define incidence for a group of health conditions to be the occurrence of a condition that was not previously reported, or a recorded reoccurrence in the case of an acute condition (cancer, heart disease, stroke). The descriptive statistics in Table 2 provide information on rates of incidence of these conditions between waves.[23]

We estimate models for incidence of each health condition, conditioned on previously considered incidences of health conditions, the prevalence of health conditions in the previous wave of the panel,

and on SES and demographic variables in the previous wave. The models are binomial probit except for BMI, which is fitted with a linear model using OLS, and numbers of ADL and IADL impairments, which are fitted as ordered probits. The estimates are given in Appendix Table A3 for whites.. Again, the data do not permit the same analysis of non-whites. The models are estimated by stacking the data for wave 1 to wave 2 transitions above the data for wave 2 to wave 3 transitions. Table 8 summarizes the health conditions and covariates that are significant risk factors for the incidence of health conditions. The associations reflect a number of known co-morbidities, but show relatively few associations of SES components and incidence of health conditions.

5.2. Causality Tests. Figure 8 gives the structure of the invariance and causality tests we report. We test only whether the model parameters are invariant between the wave 1 to 2 transitions and the wave 2 to 3 transitions. We exclude intercepts, age splines, and log of time at risk terms from the invariance test. The reason for doing so is that these terms will capture variations in survey recontact procedure across waves. However, we find in most cases that there is no significant difference in the age spline coefficients across the different transitions. The models are estimated unconstrained, and with the imposition of invariance, non-causality of SES, or both. Likelihood ratio tests are conducted for invariance, with and without non-causality imposed, and for non-causality conditioned on invariance. Since the invariance test without non-causality of SES imposed, and the non-causality test conditioned on invariance, are nested, they should give the same conclusion, at compatible significance levels, as a joint test of invariance and non-causality. In accordance with Section 2, we take acceptance of the joint hypothesis as evidence that there is *not* a direct causal path from SES to incidence of the given health condition, and take this as support for the proposition that differential access to medical care and SES-linked environmental hazards are not causing incidence rates to vary with SES.

The test results are given in Table 9. The columns of numbers in these tables are, respectively, significance levels for the invariance test with SES variables included, the invariance test with SES variables excluded, the non-causality test conditioned on invariance, and the joint test of invariance and non-causality. The final columns in the table give the relative risk for high versus low SES (see Section 4.3), and the significance level of a T-test of the null hypothesis that the relative risk is one.

In a majority of cases, our test for invariance is accepted. For cancer and heart disease incidence, it is necessary to separate models for those with and without a previous occurrence of the condition. For females, exceptions where invariance is rejected at the one percent level are mortality and ADL count. Exceptions for males are cancer with a previous occurrence, mortality, ADL count, BMI, and IADL count. The failure of the mortality models to satisfy invariance may be related to the initial selection of a non-institutionalized population in wave 1 of the AHEAD panel. Of course, in addition to the question of the power of our test to detect invariance failures, our single test of invariance across waves falls considerably short of the battery of invariance tests that would be desirable to establish that the model system has the stability and sensitivity required for policy applications.

For females, the tests for non-causality of the SES variables, conditioned on a maintained hypothesis of invariance, are rejected for arthritis and psychiatric disease at the five percent level and for cognitive impairment and self-rated health at the one percent level. Notably, these are all chronic or mental conditions where Medicare coverage is limited and the cost of drugs or assistance may be substantial. For males, this test for non-causality is rejected for cancer with a previous occurrence, heart disease with no previous occurrence, lung disease, diabetes, cognitive impairment, and self-rated health at the five percent level, and for IADL count at the one percent level. The results of the joint test for invariance and non-causality are roughly consistent with the separate tests. For conditions such as

cancer with a previous occurrence and IADL count for males, cognitive impairment for females, and mortality for both females and males, the non-causality test results may be confounded by the failure of invariance. Non-causality tests for groups of conditions are reported in Table 9. For females, the tests are accepted at the one percent level for acute diseases and mortality, degenerative and chronic diseases, and accidents, and rejected for mental diseases. For males, the tests are accepted at the one percent level for chronic diseases, accidents, and mental diseases, and rejected for acute and degenerative diseases. The mental disease test for males is marginally significant. The rejection of the tests for acute and degenerative diseases for males are the cumulative result of marginally significant rejections for cancer, heart disease, lung disease, and stroke; the pattern is consistent with a common influence of SES-linked smoking history not fully captured in the "ever smoked" covariate.

The relative risks in Table 9 should *not* be interpreted causally, since again the cases where non-causality is rejected and the relative risks are substantially different from may be due to a common unobserved effect rather than a direct causal link. The pattern of 15 relative risks exceeding one and 18 less than one suggests no broad linkage between SES and health changes, *given prior health*, and the direct links that may be indicated from the significance levels (lung disease and hip fractures for males and females, some cancers and arthritis for males) appear to be related to specific features of poverty, such as smoking history and poor dwelling environment. There are a few cases where the relative risk for high versus low SES is substantially less than one, irrespective of statistical significance, indicating an unproven link of sufficient magnitude to warrant further investigation: lung disease, hip fracture, and the mental diseases for females, and lung disease, diabetes, arthritis, and the mental diseases for males. Large deviations in relative risk from one that are not statistically significant suggest that acceptance of the joint hypothesis of invariance and non-causality could be due to low power. Notably, death shows no relation to SES, once previous health state is controlled, and the relative risks are insignificantly different from one. This indicates that there are no strong direct causal links from SES to mortality, which at the level of resolution of this study rules out differential access to medical treatment for life-threatening illness. Thus, the association of SES and mortality among the elderly appears to come primarily from variation in the prevalence of health conditions with SES, and more weakly from indirect causal links from SES to incidence of health conditions that increase mortality risk.

The pattern of failures of the non-causality test for mental diseases suggests the possibility of a direct causal link related to differential access. Medicare limits the scope of care for mental conditions, so ability to pay may indeed be an important factor in efficacy of treatments that prevent or control these conditions.

6. Tests for Non-Causality from Health Status to Asset Accumulation

6.1. Models of Incidence. Health may influence asset accumulation of elderly households because of the cost of medical treatment and related services. Medicare covers acute conditions with limited copayments, but there is the possibility of direct effects from uncovered costs of drugs and living assistance. Also, health conditions may limit the consumption of other goods, and because health status is an indicator of longevity, an individual planning consumption and precautionary reserves over remaining life may adjust target wealth based on altered perceptions of longevity and anticipated medical costs; see Alessie, Lusardi, and Kapteyn (2000), Attanasio and Hoynes (1995), Hurd (1987), Hurd and Wise (1989), Hurd, McFadden, and Gan (1998). These effects could induce an association of SES and health status even if there were no causal links from SES to health. In the elderly AHEAD

population, we will not observe the most likely direct causal link from health status to accumulation among workers, the effect of health on current labor market participation and productivity.

We analyze transitions in wealth from wave to wave using the framework of Section 2 and the model (11) for singles and (17)-(18) for couples, with demographics, previous wave health conditions, and current wave incidence of new health conditions as explanatory variables. Statistically significant selection coefficients are consistent with a direct causal link from death to a change in household wealth, but also consistent with ecological factors that induce an association of mortality risk and SES. Total, non-liquid, and liquid wealth are analyzed separately, with the transformation (10) applied to each component.

Appendix Table A4 gives the detailed incidence models for total, non-liquid, and liquid wealth change. As in previous studies of savings, we find that most of the variance in wealth changes over the population is not explained by observed economic variables. This remains true after introduction of health conditions. We find dissaving rates out of liquid wealth, before realization of returns calculated from the S&P 500, that are 5.3 percent for couples, 4.8 percent for singles, and 6.0 percent for survivors whose spouses have died. The dissaving rates from non-liquid wealth, again before realization of returns, are respectively 6.9 percent, 6.3 percent, and 8.0 percent for intact couples, singles, and survivors. The higher dissaving rates from non-liquid assets indicates that the wealth portfolios of the elderly are rebalanced to become more liquid as they age. These dissaving rates can be compared to an average rate of dissaving of 8.3 percent of remaining wealth in an age 70+ population with life table survival probabilities who consume the expected annuitized value of their wealth.[24] Then, observed dissaving rates out of wealth are not grossly lower than would be expected with pure life cycle consumption averaging over retirement and full pooling of mortality risk. We find that low income couples and individuals have significantly higher dissaving rates than their high income counterparts, but the differences are not quantitatively large. Home ownership is associated with significantly less dissaving for intact couples.

The models for both singles and couples show significant departures from normality in the selection equations. The Edgeworth expansion parameters show positive skewness and smaller than normal kurtosis for female singles, negative skewness and insignificantly different from normal kurtosis for male singles. For couples, both males and females have negative skewness and larger tan normal kurtosis. We also find significant selection effects, with ρ_f = -0.49 for couples and -0.21 for singles, and ρ_m = -0.51 for couples and -0.89 for singles. These imply that households that survive despite unfavorable mortality risks have increased dissaving, either because of increased cost of overcoming health problems or because households at elevated risk spend down more rapidly. Equation (18) includes shift parameters that modify the dependence of the wealth change disturbance on the unobserved selection effects in regimes where a spouse dies. These are statistically significant, and sufficiently large to reverse the direction of the intact couple selection effects.

Table 10 summarizes the health conditions and other covariates that are individually significant in explaining changes in wealth. For intact couples, we find some acute conditions *increase* saving, perhaps because they restrict consumption, or perhaps because couples conserve assets for a potential surviving spouse. For the conditions that are associated with increased dissaving (cognitive impairment and stroke for single females), costs of maintenance associated with these conditions may be directly causal to wealth changes.

6.2. Causality Tests. Table 11 summarizes our tests for invariance and absence of direct causal links. We test for common parameters in the wealth change models between waves 1-2 and waves 2-3,

excepting intercepts and age effects to allow for the effects of sample timing. Invariance is convincingly rejected for each demographic group and wealth category, indicating that our model fails to capture the structural determinants of wealth change. As a consequence, our non-causality tests to follow may produce rejections due to model misspecification, confounding the detection of direct causal links. A deconstruction of the invariance failures, detailed in Appendix Table A5, shows that for nonliquid and liquid wealth, invariance passes for demographic and health prevalence and incidence variables, but fails for female SES variables, and for all male variables including SES taken together. Thus, there was an unexplained regime shift before and after wave 2 of AHEAD. Possible explanations for this are an interaction between the criterion of non-institutionalization in the initial panel recruitment and economic behavioral response, a wealth-linked interaction in panel retention, problems in the measurement of wealth in the AHEAD population, which exhibits unexplained mean reversion, or a true behavioral shift with age in a single cohort that is not captured accurately by a model that pools wealth change observations across cohorts.

We expect that terminal medical and burial expenses, estate taxes and other estate settlement costs, insurance payments, and bequests will have a substantial impact on the size of a decedent's estate and surviving spouse's reported wealth. We easily reject the hypothesis that the model coefficients for intact couples and for surviving spouses are the same, but note that measurement problems associated with a change in financially responsible respondent could also produce this rejection.

The hypothesis of no direct causality of health conditions for total wealth changes is rejected at the one-percent level for intact couples and singles. For nonliquid wealth, the hypothesis is rejected for intact couples, and for liquid wealth, the hypothesis is rejected for all demographic groups. The failure of the invariance tests, and the possibility of confounding by persistent common factors and selection suggest that conclusions on the health to wealth link be interpreted with caution. Appendix Table A6 deconstructs the causality tests and identifies the health conditions and genders responsible for rejections. The pattern of results suggest that if there is indeed a direct causal link, then it is most likely to involve liquid wealth and health conditions that require assisted living.

Appendix Table A7 estimates models of income change, given health conditions and other covariates. One would not expect health status to have a significant impact on the incomes of retirees, conditioned on previous wealth, and the empirical results are generally consistent with this expectation. We have not done formal tests of invariance or causality for income. Also included in the state vector Y_t that describes individuals are changes in home ownership status, and dwelling and neighborhood conditions. We estimate incidence models for these components; results are in Appendix Table A.8.

7. Simulation of Life Courses under Counterfactual Conditions

7.1. The Simulation Experiment. For policy analysis of interventions that alter delivery or cost of medical services, or retirement financing, one would like to trace through the direct and indirect causal links between wealth, health, and mortality. If Markov models of the sort developed in Sections 2-6 satisfy the required invariance properties, then they can be used to simulate the impacts of these interventions on the life courses of a synthetic population. In this section, we develop such a simulation analysis, and apply it to illustrative interventions. Because we have not uniformly accepted invariance, and in a number of cases find associations that may be either direct causal links or hidden common factors, this simulation analysis assumes more than our estimates support. It should be interpreted only as an exercise that shows how a model of this general structure might be used in a policy application to

unravel the dynamics of co-morbidities and forecast condition-specific morbidity and mortality, and life expectancy.

We simulate the life courses of a synthetic population in which heads of household are initially age 70-74. To synthesize this population, we start from the 1612 households in AHEAD Wave 1 whose heads are white and age 70-74. Using the SES and demographic variables for each household in this subsample, we make ten Monte Carlo draws from the prevalence models in Section 4 to create synthetic initial health profiles for the household head, and spouse if present. This gives an initial synthetic sample of 16120 households. We then create life courses for the members of each household by drawing recursively from the Markov incidence models in Sections 5 and 6, adjusted to annual transitions using (9) to approximate the probabilities of moving to new states..

We first consider a base scenario (S0) in which initial prevalence and incidence transitions are given by our models estimated on the AHEAD data. We note that the simulation outcomes can be expected to differ to some degree from the cross-cohort patterns found in AHEAD, because the distribution of conditions at ages 70-74 will differ from the distributions of prevalence that actually prevailed for older individuals in AHEAD when they were ages 70-74. They should also differ to some degree from the actual experience that the age 70-74 cohort in AHEAD will have through the remainder of their lives, because the simulation cannot anticipate the realized future distribution of exogenous variables, and because our models do not allow for drift in disease incidence or condition-specific mortality hazards that will result from changes in medical care. Historically, these drifts have been very significant, reducing morbidities and increasing life expectancies. If these trends continue, then the baseline simulations will underestimate actual survival experience.

We next consider two stylized policy interventions. The first alternative scenario (S1) examines the impact on life courses of the introduction of a hypothetical medical treatment that cures diabetes, for example by stem cell and immune system therapy that rejuvenates the pancreas for both type I and type II diabetics. In this scenario, we assume that prevalence of diabetes at the start of the simulated panel drops to zero, and that there is zero incidence of this condition as the cohort ages. We do not alter the historical prevalence of conditions associated with diabetes. Thus, we assume that the historical impact on individuals of type I diabetes, notably increased prevalence of heart disease and stroke at age 70-74, is not altered. The second alternative scenario (S2) examines the impact on life courses of reducing the entire population from their current socioeconomic status to our definition of a low-SES individual: bottom quartile for wealth and income, less than a high school education, and a poor or fair neighborhood and dwelling. This alternative is obviously hypothetical, and is not even a stylized approximation to any real policy alternative. However, it provides an extreme in which the interactions of health and SES are permitted maximum play, giving an upper bound on the effect that SES could possibly have on health outcomes, and providing a finger exercise that tests the plausibility of our model system.

7.2. Baseline Simulation. Table 12 summarizes the survival probabilities and prevalence of health conditions among survivors in the simulated cohort as it ages, under the baseline scenario (S0), the no-diabetes scenario (S1), and the low-SES scenario (S2). Keeping in mind that we expect the simulation model to differ from the historical cross-cohort record in AHEAD, the success of this model in plausibly mimicking observed conditions in the AHEAD population can be judged by comparing the results for scenario S0 with life table survival probabilities. Life expectancy for a cohort of white females aged 70-74 is 13.15 years from the 1996 Life Tables and 14.36 years in our baseline simulation. The life table probability of survival for 15 years for the aged 70-74 cohort is 0.535, while the corresponding survival

probability in the simulation model is 0.500. For white males, the life expectancy at age 70 is 11.31 years from the Life Tables and 10.81 years from the S0 simulation. The 15-year survival probability is 0.381 from the life tables and 0.279 from the simulation model. Thus, relative to the life tables, the simulation model under-predicts female mortality and over-predicts male mortality. The comparison of annual mortality rates for white females given in Figure 5 indicates that actual AHEAD mortality experience was more favorable than the life tables between waves 1 and 2, presumably due to selection in panel recruitment, and very close to the live tables between waves 2 and 3. Then, the baseline simulation appears to reproduce relatively accurately the cross-cohort survival experience in AHEAD. This provides a reality check for the simulation model, but also suggests that if the survival experience of a current cohort differs from the historical cross-cohort pattern, then the simulation model will miss the drift in mortality hazards that a single cohort will face in the future.

A comparison of prevalence rates for various health conditions among survivors of various ages can be made between AHEAD at wave 1, given in Table 4, and the baseline simulation in Table 12. There are issues of comparability in the definition of prevalence for some conditions, but the pattern that emerges is that the simulated prevalences are systematically higher than the historical prevalences, and increasingly so at older ages. For example, for white females aged 80-84, the historical and simulated prevalence rates are 0.168 versus 0.219 for cancer, 0.361 versus 0.442 for heart disease, and 0.338 versus 0.437 for cognitive impairment. One possible explanation for this is that the links from morbidity to mortality are stronger than the mortality model detects, perhaps because of under-reporting of health conditions that arise prior to death, so that the simulation model underestimates the selection effect of mortality that reduces prevalence among survivors. A second possible explanation is that there is strong unobserved heterogeneity in susceptibility to various health conditions, so that cumulative prevalence is overestimated by our first-order Markov models which describe prevalence for most conditions as the result of one or more positives in a series of Bernoulli trials. It is possible to test for persistent unobserved heterogeneity by asking whether the frequency of a negative for a condition between waves one and three of AHEAD is the product of the frequencies of a negative between successive waves. When we do this, we do not find persistent unobserved heterogeneity. However, the power of the test is modest, and it is possible that even a limited degree of persistent unobserved heterogeneity is enough to explain the differences in AHEAD and the simulation.

A comparison is given in Table 13 between median wealth and income by age in the AHEAD panel and in the baseline simulation. The historical cross-cohort data shows sharply declining wealth and income with age, and a less liquid portfolio with age, reflecting strong cohort effects as well as life cycle and selection effects. The simulation results, which exclude cohort effects, nevertheless show even more sharply declining wealth with age, and a strong shift toward a more liquid portfolio mix. If the simulation is correctly describing portfolio balance of a single cohort over its life course, then there is a strong cross-chort effect, with older cohorts starting from retirement portfolios that are more heavily invested in housing equity and other nonliquid forms. The simulated semi-interquartile range is narrower than its historical counterpart, particularly for older households. In the simulation, variability (defined as the ratio of the semi-interquartile rannge divided by the median) falls with age, whereas in the historical cross-cohort data variability rises with age. This suggests that in addition to cross-cohort effects, that there may be persistent heterogeneity in savings behavior that is not captured in our model.

7.3. Alternative Scenarios. Table 12 gives the survival probabilities and prevalences of various health conditions under our alternative no-diabetes scenario (S1) and the low-SES scenario (S2). In the no-diabetes simulation, the direct mortality risk from diabetes, and the incidence of co-morbidities with

diabetes are eliminated, although our simulated population will display elevated prevalence of heart disease and stroke at age 70 among former diabetics. Life expectancy at age 70 under this scenario increases from 14.29 years to 14.69 years for females, and from 10.78 years to 11.26 years for males. These rates imply in turn that a former diabetic's life expectancy increases by 2.72 years for females and 3.32 years for males. Other health condition prevalences that fall in the absence of diabetes are heart disease, stroke, cognitive impairment, ADL and IADL impairment, and self-rated health. While reduction in mortality risk from one source must as a matter of accounting eventually lead to more deaths from competing risks, there are no substantial movements in prevalences of the remaining health conditions.

The alternative low-SES scenario reduces our entire age 70 simulated population to the bottom quartile for wealth and income, gives them less than a high school education, and places them in a dwelling in poor condition in an unsafe neighborhood. They are kept in this low-SES status for the remainder of their lives; i.e., there is no opportunity in this simulation for households to escape low SES by a lucky change in wealth or income. However, our population displays the patterns of prevalence of health conditions established in their earlier lives with their historical SES status. This scenario is quite artificial, but it demonstrates the holistic effect on the broad spectrum of health conditions of low SES. Life expectancies at age 70 in this scenario drop dramatically, from 14.36 years to 12.27 years for females, and from 10.81 to 9.56 years for males. Prevalences of cancer, heart disease, lung disease, diabetes, arthritis, incontinence, hip fractures, cognitive impairment, psychiatric disease, and depression all increase sharply, as do the number of ADL and IADL impairments. Conditions whose prevalence is not affected substantially by low SES are stroke, high blood pressure, and falls. These results indicate that *if* the associations of SES and incidence of health conditions that we find in AHEAD *were entirely* the result of direct causal links from wealth to health, then the protective effect of the prevailing pattern of higher SES is 1.26 to 2.08 years of added life expectancy. Thus, our findings that for most health conditions the evidence is against direct causal links from SES to incidence do not appear to rule out a substantial cumulative effect of SES over conditions and time that induce a noticeable SES gradient in mortality. Given our specific findings against direct causal links from SES to incidence of acute conditions and mortality, the most obvious possible source for this gradient are SES-linked differences in genetic susceptibility and behavior.

We have emphasized that our stylized, hypothetical policy intervention and the changes in health they produce over the life course are strictly illustrative, and should be interpreted with great caution. These finger exercises *cannot* be used to draw conclusions about any real policy initiatives. This is particularly true since we have included within our model system components that fail the invariance tests that we have emphasized must be met by a valid policy model, and because in many cases our models display *some* wealth or income gradients for incidence that we cannot with our statistical methods identify as the sole result of direct causal links. While most of these effects are not statistically significant, it is possible that in a larger or longer panel with greater statistical power, they will prove to be significant. Because the simulation model accumulates and amplifies the direct and indirect causal links between health and SES, it potentially could be a vehicle for more powerful statistical tests for causal links than one can attain in a relatively short panel. When causal links cannot be rules out, it is essential to turn to the more advanced statistical methods of Heckman (2001) and others to identify the direct causal components in these incidence associations, and improve the models to achieve invariance. Only after this is done, and realistically detailed policy scenarios are considered, could policy makers take our model system seriously as a policy tool. However, we believe that our results do demonstrate that it would be useful for health policy analysis to utilize a system of invariant models with a causal

chain structure to simulate policy impacts, in a framework that takes into account indirect impacts, competing hazards, and direct causal links between SES and health. We believe that analysis of the broad sweep of co-morbidities and wealth effects over the life course is an important complement to the disease-centric orientation of many medical and epidemiological studies of health outcomes.

8. Summary and Speculations for Further Research

This paper has used innovations in health conditions and in wealth in the AHEAD panel to carry out tests for the absence of direct causal links from SES to health, and from health conditions to wealth. By advancing beyond the detection of association to a framework in which there is some possibility of detecting the absence of causal links, this paper provides a methodology that may be useful in winnowing the list of possible direct causal mechanisms, or delimiting their domain of action. For the AHEAD sample, a panel of U.S. elderly age 70 and older in 1993, we conclude for females that for mortality and for acute, sudden-onset diseases, the hypothesis of *no* causal link from SES is accepted, and for incidence of mental problems the hypothesis is rejected. For both makes and females, we conclude for chronic diseases and accidents that the hypothesis is accepted. The results for degenerative diseases are mixed. We generally reject the hypothesis of no direct causal link from health conditions to total wealth changes, but cannot rule out confounding of the test by invariance failures.

The pattern of results suggests that incidence of accidents, chronic diseases, and (for females) acute, sudden onset health conditions, conditioned on existing health conditions, does not exhibit a significant SES gradient, while incidence of mental and degenerative conditions appear to have an association to SES due to some combination of direct causal links and common unobserved behavioral or genetic factors. The results suggest that there may be an SES gradient in seeking treatment for the second class of conditions, which may influence detection, or for maintaining preventative regimens that may maintain some conditions below the reporting thresholds. Our findings are not inconsistent with the possibility that for mental and chronic illnesses where the acute care procedures covered by Medicare are often inapplicable, ability to pay may be a causal factor in seeking and receiving treatment. We do not find systematic persuasive associations of health conditions and changes in total wealth, except for surviving spouses. Problems in measuring and modeling wealth changes suggest caution in concluding from these results that there is generally no direct causal link from health conditions to wealth changes.

We emphasize that our results apply only to elderly individuals in the U.S., where Medicare and Medicaid programs limit out-of-pocket medical costs, particularly for acute care, and where retired status eliminates a possible direct causal link from health status to ability to work. Further, in an elderly population, common factors may be manifest in prior health conditions and economic status, so they have little impact once incidence is conditioned on prior state. Our results provide no evidence on the nature of the causal links at younger ages, during the stages of life where association of health and wealth is emerging as a consequence of some causal structure.

Future waves of the AHEAD (HRS) panel will allow the hypotheses of invariance and non-causality to be tested with greater power. This will particularly be the case when full tracking of decedents, and determination of cause of death from medical records, become part of the data. It seems likely that some of the associations we have found between changes in health and wealth will survive more detailed analysis, and that suitably defined natural or designed experiments are likely to be needed to fully unravel the causal structure underlying these associations.

The modeling structure used in this paper is parametric, and the high dimensionality of the vector of possible explanatory variables and the relatively limited information contained in binomial outcomes in the AHEAD panel make it difficult to move to a more robust non-parametric analysis. However, we have been flexible in specifying the variable transformations that appear in our models, and we interpret our analysis as conforming in spirit, if hardly in fact, to a method of sieves approach to non-parametric analysis. One of the major limitations of our models, which would be likely to lead them to fail invariance tests in situations where a sharp test is possible, is that they do not account adequately for the

multiple risk structure of health conditions and its implications for the duration patterns that can emerge, particularly over the relatively long intervals between waves. Some outcomes, such as mortality and non-fatal heart disease, are competing risks, while others, like diabetes and heart conditions, are complementary risks. For future research, we are investigating hidden Markov models in which a latent vector of propensities for all health and SES conditions follows a first-order Markov process, conditioned on demographic state, and all possible causal links across the components of this latent vector appear in the model. Given thresholds that trigger observed states, this model provides a consistent but computationally demanding data generation process for the vector of Markov states month-by-month. Within this model, it is possible to carry out joint tests for the absence of classes of causal links. The next wave of this research, incorporating Wave 4 of AHEAD, will include full development of flexible multiple-risk duration models.

References

Adler, N.; Ostrove, J. (1999) "Socioeconomic status and health: what we know and what we don't," in N. Adler et al (eds) *Socioeconomic Status and Health in Industrialized Nations*, Annals of the New York Academy of Sciences, 896, 3-15.

Alessie, Rob; Lusardi, Annamaria; Kapteyn, Arie (2000) Saving after Retirement: Evidence from Three Different Surveys *Labour Economics*; 6(2), 277-310.

Angrist, J.; Imbens, G.; Rubin, D. (1996) "Identification of Causal Effects using Instrumental Variables," *Journal of the American Statistical Association*, 91, 444-472.

Attanasio, Orazio; Hoynes, Hilary (1995) Differential Mortality and Wealth Accumulation National Bureau of Economic Research, Working Paper: 5126.

Backlund, E., P. Sorlie; N. Johnson (1999) "A Comparison Of The Relationships Of Education And Income With Mortality: The National Longitudinal Mortality Study," *Social Science & Medicine*, 49, 1373-1384.

Baker, M.; Stabile, M.; Deri, C. (2001) "What Do Self-Reported Objective Measures of Health Measure?" National Bureau of Economic Research Working Paper 8419.

Barsky, R.; Juster, T.; Kimball, M.; Shapiro, M. (1997) "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Survey," *Quarterly Journal of Economics*, 112, 537-579.

Bosma, H., Marmot, M.G., Hemingway, H., Nicholson, A.C., Brunner, E., Stansfeld, S.A. (1997) Low job control and risk of coronary heart disease in Whitehall II (prospective cohort) study *BMJ* 314, 558-565.

Chandola, T., (1998) Social inequality in coronary heart disease: a comparison of occupational classifications. *Social Science & Medicine* 47, 525-533.

Chandola, T. (2000) "Social Class Differences In Mortality Using The New UK National Statistics Socio-Economic Classification," *Social Science & Medicine*, 50, 641-649.

Chapman, K. and G. Hariharran (1994) "Controlling for Causality in the Link from Income to Mortality," *Journal of Risk and Uncertainty*; 8(1), 85-93.

Chay, K.; Greenstone, M. (2000) "Does Air Quality Matter? Evidence from the Housing Market," *Journal of Political Economy*, forthcoming.

Davey-Smith, G.; Blane, D.; Bartley, M. (1994) "Explanations for socio-economic differentials in mortality: evidence from Britain and elsewhere," *European Journal of Public Health*, 4, 131-144.

Dawid, A. (2000) "Causal Inference without Counterfactuals," *Journal of the American Statistical Association*, 95, 407-48.

Dohrenwend, B. et al (1992) "Socioeconomic Status and Psychiatric Disorders: the Causation-Selection issue," *Science*, 255, 946-952.

Drever, F., Whitehead, M. (1997) *Health Inequalities* ONS, London.

Ecob, R., Smith, G. (1999) "Income and health: what is the nature of the relationship?" *Social Science & Medicine* 48, 693-705.

Elo, I.; Preston, S. (1996) "Educational Differentials In Mortality: United States, 1979-85," *Social Science & Medicine*, Volume 42, Issue 1, Pages 47-57

Engle, R., D. Hendry, and J. Richard (1983) "Exogeneity," *Econometrica*, 51, 277-304.

Ettner, S. (1996) "New Evidence on the Relationship Between Income and Wealth," *Journal of Health Economics*, 15, 67-85.

Evans, A. (1978) "Causation and Disease: A Chronological Journey," *American Journal of Epidemiology*, 108, 249-258.

Feigl, H. (1953) "Notes on Causality," in H. Feigl and M. Brodbeck (eds) *Readings in the Philosophy of Science*, Appleton-Century-Crofts: New York.

Feinstein, J. (1992) "The Relationship between Socioeconomic Status and Health: A Review of the Literature," *The Millbank Quarterly*, 71 (2) pp 279-322.

Felitti, V. *et al* (1998) "Relationship of Childhood Abuse and Household Dysfunction to Many of the Leading Causes of Death in Adults," *American Journal of Preventative Medicine*, 14, 145-158.

Fitzpatrick, R., Bartley, M., Dodgeon, B., Firth, D., Lynch, K. (1997) Social variations in health: relationship of mortality to the interim revised social classification. In: Rose, D., O'Reily, K. (Eds.),*Constructing Classes*. ESRC/ONS, Swindon.

Fitzpatrick, J., Dollamore, G. (1999) Examining adult mortality rates using the National Statistics Socio-Economic Classification. *Health Statistics Quarterly* 2, 33-40.
Fox, A., Goldblatt, P.; Jones, D. (1985) Social Class Mortality Differentials: Artifact, Selection, or Life Circumstances?" *Journal of Epidemiology and Community Health*, 39, 1-8.
Freedman, D. (1985) "Statistics and the Scientific Method," in W. Mason and S. Fienberg, eds, *Cohort Analysis in Social Research: Beyond the Identification Problem*, Springer-Verlag: New York, 343-90.
Freedman, D. (2001) "On Specifying Graphical Models for Causation and the Identification Problem," Working Paper, Department of Statistics, University of California, Berkeley.
Fuchs, V. (1993) "Poverty and health: Asking the Right Questions," in D. Rogers and E. Ginzberg, eds, *Medical Care and the Health of the Poor*, Boulder: Westview Press, 9-20.
Geweke, J. (1984) "Inference and Causality in Economic Time Series Models," in Z. Griliches and M. Intriligator (eds) *Handbook of Econometrics*, North Holland: Amsterdam, Vol. 2, 1101-1144.
Gill, R.; Robins, J. (2001) "Causal Inference for Complex Longitudinal Data: The Continuous Case," *Annals of Statistics*, forthcoming.
Goldblatt, P.O. (1990) *Longitudinal Study, Mortality and Social Organization.* HMSO, London.
Goldman, N. (1994) "Social Factors and Health: The Causation-Selection Issue Revisited," *Proceedings of the National Academy of Science*, 91, 1251-1255.
Goldman, N. (2001) "Social Inequalities in Health: Disentangling the Underlying Mechanisms," in *Demography and Epidemiology: Frontiers in Population Health and Aging*, forthcoming.
Granger, C (1969) "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*; 37, 424-38.
Haynes, R. (1991) Inequalities in health and health service use: evidence from the General Household Survey. *Social Science & Medicine* 33, 361-368.
Heckman, J. (2000) "Causal Parameters and Policy Analysis in Econometrics: A Twentieth Century Retrospective," *The Quarterly Journal of Economics*, 105, 45-97.
Heckman, J. (2001) "Econometrics Counterfactuals and Causal Models," International Statistical Institute, Working Paper, University of Chicago.
Hendry, D. and G. Mizon (1999) "The Pervasiveness of Granger Causality in Econometrics," in R. Engle and H. White (eds) *Cointegration, Causality, and Forecasting*, Oxford University Press: Oxford.
Hertzman, C. (1999) "Population Health and Human Development," in D. Keating and C. Hertzman, eds, *Developmental Health and the Wealth of Nations*, New York: Guilford Press.
Herzog, R. and R. Wallace (1997) "Measures of Cognitive Functioning in the AHEAD Study," *Journals of Gerontology*, Series B, 52B, 37-48
Holland, P. (1986) "Statistics and Causal Inference," *Journal of the American Statistical Association*, 81, 945-60.
Holland, P. (1988) ""Causal Inference, Path Analysis, and Recursive Structural Equation Models," in C. Clogg, ed, *Sociological Methodology*, Washington: American Sociological Association, Chapter 14.
Hoover, K. (2001) *Causality in Macroeconomics*, Cambridge: Cambridge University Press.
Humphries, K. and van Doorslaer, E. (2000) Income-Related Health Inequality In Canada *Social Science & Medicine*, 50 (5), 663-671
Hurd, M. (1987) "Savings of the Elderly and Desired Bequests," *American Economic Review* 77, 298-312.
Hurd, M. and D. Wise (1989) "Wealth Depletion and Life Cycle Consumption by the Elderly," in D. Wise, ed., *Topics in the Economics of Aging*, University of Chicago Press.
Hurd, M.; McFadden, D.; Merrill, A. (1999) Healthy, Wealthy, and Wise? The Causal Links between Socioeconomic Status and Health, NBER Working Paper.
Hurd, M., D. McFadden and L. Gan (1998) "Subjective Survival Curves and Life Cycle Behavior," in D. Wise, ed., *Inquiries in the Economics of Aging*, University of Chicago Press.
Hurd, M., D. McFadden, L. Gan, A. Merrill, and M. Roberts (1998) "Consumption and Savings Balances of the Elderly: Experimental Evidence on Survey Response Bias," in D. Wise, ed., *Frontiers in the Economics of Aging*, University of Chicago Press, 353-387.
Hurd, M. and K. McGarry (1997) "Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Survey," *Journal of Human Resources*, 30, S268-292.

Kaplan, J.; Manuck, S. (1999) "Status, stress, and arteriosclerosis: the role of environment and individual behavior," in N. Adler et al, eds, *Socioeconomic Status and Health in Industrialized Nations*, Annals of the New York Academy of Sciences, 896, 145-161.

Karasek, R.A., Theorell, T., Schwartz, J., Schnell, P., Peiper, C., Michela, J. (1988) Job characteristics in relation to the prevalence of myocardial infarction in the US Health Examination Survey (HES) and the Health and Nutrition Examination Survey (NHANES). *American Journal of Public Health* 78, 910-918.

Kelley, S.; Hertzman, C.; Daniels, M. (1997) "Searching for the biological pathways between stress and health," *Annual Review of Public Health*, 18, 437-462.

Kitagawa, E. and P. Hauser (1973). Differential Mortality in the United States: A Study in *Socioeconomic Epidemiology*, Cambridge, MA: Harvard University Press.

Kunsch, H.; Geman, S; Kehagias, A. (1995) "Hidden Markov Random Fields," *Annals of Applied Probability*, 5, 577-602.

Lewis, G., Bebbington, P., Brugha, T., Farell, M., Gill, B., Jenkins, R., Meltzer, H. (1998) Socioeconomic status, standard of living and neurotic disorder. *Lancet* 352, 605-609.

Marmot, M. et al (1991) "Health Inequalities Among British Civil Servants: the Whitehall II Study," *Lancet*, 337, 1387-93.

Marmot, M., Bosma, H., Hemingway, H., Brunner, E., Stansfeld, S. (1997) Contribution of job control and other risk factors to social variations in coronary heart disease incidence. *Lancet* 350, 235-239.

Marmot, M.; Bobak, M.; Davey-Smith, G. (1995) ""Explanations for Social Inequalities in Health," in B. Amick et al, eds, *Society and Health*, Oxford: Oxford University Press.

McDonough, P.; Duncan, G.; Williams, D.; House, J. (1997) "Income Dynamics and Adult Mortality in the United States," *American Journal of Public Health*, 87, 1476-1483.

Meer, J.; Miller, D.; Rosen, H. (2001) "Exploring the Health-Wealth Nexus," Princeton University Working Paper.

Leigh, J.; Dhir, R. (1997) "Schooling and Frailty among Seniors," *Economics of Education Review*, 16(1), 45-57.

Luft, H. (1978) *Poverty and Health: Economic Causes and Consequences of Health Problems* Cambridge, Mass., Ballinger Publishing Co.

Martin, L.; Preston, S. (1994) *Demography of Aging*, Washington, D.C.: National Academy Press.

Martin, L.; Soldo, B. (1997) *Racial and ethnic differences in the health of older Americans* Washington, D.C.: National Academy Press.

McEwen, B.; Stellar, E. (1993) "Stress and the Individual: Mechanisms Leading to Disease," *Archives of Internal Medicine*, 153, 2093-2101.

Murray, C.; Yang, G.; Qiao, X. (1992) "Adult Mortality Levels: Patterns and Causes," in R. Feachem *et al*, eds, *The Health of Adults in the Developing World*, New York: Oxford University Press.

Newey, W.; Powell, J.; Walker, J. (1990) "Semiparametric Estimation of Selection Models: Some Empirical Results," *American Economic Review*, 80, 324-328.

Nozick, R. (2001) *Invariances: The Structure of the Objective World*, Harvard University Press: Cambridge.

Pearl, J. (2000) *Causality: Models, Reasoning, and Inference*, Cambridge: Cambridge University Press.

Power, C.; Matthews, S.; Manor, O. (1996) "Inequalities in Self-Rated Health in the 1958 Birth Cohort: Life Time Social Circumstances or Social Mobility?" *British Medical Journal*, 313, 449-453.

Power, C.; Matthews, S. (1998) "Accumulation of health risks across social groups in a national longitudinal study," in S. Strickland and P. Shetty, eds, *Human Biology and Social Inequality*, Cambridge: Cambridge University Press.

Preston, S. and Taubman, P. (1994) "Socioeconomic Differences in Adult Mortality and Health Status," in L. Martin and S. Preston (eds) *Demography of Aging*, National Academy Press: Washington, p. 279-318.

Riberio, T. (2002) "Selection Model Used for HWW Wealth Regressions," discussion paper.

Robins, J. (1999) "Association, Causation, and Marginal Structural Models," *Synthese*, 121, 151-179.

Rodgers, B., (1991). Socio-economic status, employment and neurosis. Social Psychiatry and *Psychiatric Epidmiology* 26, 104-114.

Ross, C.; Mirowsky, J. (2000) "Does Medical Insurance Contribute to Socioeconomic Differentials in Health?" *The Millbank Quarterly*, 78, 291-321.

Schnall, P.; Landsbergis, P.; Baker, D. (1994) Job strain and cardiovascular disease. *Annual Review of Public Health* 15, 381-411.

Seeman, T.; Singer, B.; Love, G.; Levy-Storms, L. (2001) “Social Relationships, Gender, and Allostatic Load Across Two Age Cohorts,” *Psychosomatic Medicine*, forthcoming.

Shorrocks, A. (1975) "The Age-Wealth Relationship: A Cross-Section and Cohort Analysis," *Review of Economics and Statistics*, 57, 155-163.

Sims, C. (1972) "Money, Income, and Causality," *American Economic Review*; 62, 540-52..

Smith, J. and R. Kington (1997) "Race, Socioeconomic Status, and Health in Late Life," in L, Martin and B. Soldo (eds) *Racial and Ethnic Differences in the Health of Older Americans*, National Academy Press: Washington, 106-162.

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

Smith, J. (1998) "Socioeconomic Status and Health," *American Economic Review*, 88, 192-196.

Smith, J.; Kington, R. (1997) Demographic and Economic Correlates of Health in Old Age *Demography*; 34(1), 159-70..

Sobel, M. (1997) "Measurement, Causation, and Local Independence in Latent Variable Models," in M. Berkane (ed) *Latent Variable Modeling and Applications to Causality*, Springer: Berlin.

Sobel, M. (2000) "Causal Inference in the Social Sciences," Journal of the American Statistical Association, 95, 647-51.

Soldo, B., M. Hurd, W. Rodgers and R. Wallace (1997) "Asset and Health Dynamics among the Oldest Old: An Overview of the AHEAD Study," *The Journals of Gerontology*, Series B, 52B, 1-20.

Staiger, D.; Stock, J. (1997) "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 64, 556-586.

Swert, W. (1979) "Tests of Causality," in K. Brunner and A. Meltzer (eds) *Three Aspects of Policy and Policymaking* North Holland: Amsterdam, 55-96.

Stern, J. (1983) Social mobility and the interpretation of social class mortality differentials. *Journal of Social Policy* 12, 27-49.

Wadsworth, M. (1991) *The Imprint of Time: Childhood, History, and Adult Life*, Oxford: Clarendon Press.

Whitehead, M. (1988) *The Health Divide*. Penguin, London

Wilkinson, R. (1998) "Equity, Social Cohesion, and Health," in S. Strickland and P. Shetty, eds, *Human Biology ans Social Inequality*, Cambridge: Cambridge University Press.

Woodward, J. (1999) Causal Interpretation in Systems of Equations," *Synthese*, 212, 199-247.

Woodward, M., Shewry, M., Smith, C., Tunstall-Pedoe, H. (1992) Social status and coronary heart disease: results from the Scottish Heart Health Study. *Preventive Medicine* 21, 136-148.

Zellner, A. (1979) "Causality and Econometrics," in K. Brunner and A. Meltzer (eds) *Three Aspects of Policy and Policymaking*, North Holland: Amsterdam, 9-54.

Footnotes

1.Backlund, Sorlie, and Johnson, 1999; Barsky, Juster, Kimball, and Shapiro, 1997; Bosma et al,, 1997; Chandola, 1998,2000; Davey-Smith, Blane, and Bartley, 1990; Drever and Whitehead, 1997; Ecob and Smith 1999; Elo and Preston, 1996; Ettner, 1996; Feinstein, 1992; Fitzpatrick et al, 1997; Fitzpatrick and Dollamore, 1999; Fox and Goldblatt, 1982; Goldblatt, 1990; Hayes, 1991; Hertzman, 1999; Hurd, 1987; Hurd and Wise, 1989; Karasek et al, 1988; Kitagawa and Hauser, 1973; Lewis et all, 1998; Leigh and Dhir, 1997; Luft, 1978; Marmot, Bobak, and Davey-Smith, 1995; Marmot *et al*, 1997; Martin and Preston, 1994; Martin and Soldo, 1997; McDonough, Duncan, Williams, and House, 1997; Murray, Yanf, and Qiao, 1992; Power, Matthews, and Manor, 1996; Powers and Matthews, 1998; Rogers, 1991; Ross and Mirowsky, 2000; Schnall, Landsbergis, and Baker, 1994; Seeman, McEwan, Rowe, and Singer, 2001; Shorrocks, 1975; Stern, 1983; Whitehead, 1988; Wilkinson, 1998; Woodward et al, 1992.

2.The associations can become more complex when multiple health conditions and multiple SES measures are studied. Competing risks may mask the hazard for late-onset diseases; e.g., elevated mortality risk from cardiovascular diseases in low SES groups may induce an apparent reverse relationship between SES and later-onset cancer in the surviving population (Adler and Ostrove, 1999). Longer-run measures of SES such as education, occupation, and wealth appear to have a stronger association with health status than current income (Fuchs, 1993). Using carefully measured wealth, we find that it explains most of the association with health, and education conditioned on wealth is not systematically correlated with health.

3. Papers examining explicit causal mechanisms include Dohrenwend et al, 1992; Evans, 1978; Felitti et al, 1998; Fox, Goldblat, and Jones, 1985; Goldman, 1994; and McEwen and Stellar, 1993.

4.There may be an important distinction between direct causal mechanisms influencing mortality, conditioned on health status, and direct causal mechanisms influencing onset of health conditions. For mortality, an SES gradient could be due to differentially effective treatment of acute health conditions. For morbidity, an SES gradient could reflect differentials in prevention and detection of health conditions. These involve different parts of the health care delivery system, and differ substantially in the importance of individual awareness and discretion, and allocation of costs between Medicare and the individual.

5.For example, industrial and traffic pollution, and poor dwelling ventilation, are risk factors for lung disease, and housing prices and household income are negatively correlated with air pollution levels in census data (Chay and Greenstone, 2000).

6.Consider phenomena such as under-estimation of the hazard of a disease due to competing risks from other illnesses or death. In an unfortunate discrepancy in terminology, economists would call this a selection effect while epidemiologists would classify it as an artifactual mechanism rather than a selection mechanism.

7.Usually, one can argue that observed association must originate from some initial causal action, so that common factors originate from some initial direction of causation. However, there is no apparent initial causal action for genetically-linked conditions such as Downs syndrome, which increase mortality risk and preclude work. Further, as a practical matter, it is often impossible to make observations at the high frequencies that would be required to identify causal chains when feedbacks are nearly instantaneous. Then common factors will appear at feasible levels of detection to operate simultaneously, and their true causal structure will not be identified. For these reasons, there would be considerable merit in adding *common mechanisms* to the epidemiologist's classifications.

8.Dawid, 2000; Freedman. 1985, 2001; Granger, 1969; Sims, 1972; Zellner, 1979; Schwert, 1979; Engle, Hendry, and Richard, 1983; Geweke, 1984; Gill and Robins, 2001; Heckman, 2000,2001; Holland, 1986, 1988; Pearl, 2000; Robins, 1999; Sobel, 1997, 2000; Hendry and Mizon, 1999; Woodward, 1999

9.Any discrete-time stationary stochastic process can be approximated (in distribution for restrictions to a finite number of periods) by a first-order hidden Markov model, so there is no loss of generality in considering only

models of this form; see Kunsch, Geman, and Kehagias (1995).

10. Econometricians have traditionally used the term strictly exogenous to refer to properties of variables in the true data generation process, a stronger non-positivistic version of this condition.

11. For the CDF F of a random variable Y, define $F(y\_) = \sup_{y'<y} F(y')$ and $F^{-1}(p) = \min\{x' | F(x') \geq p\}$. Define the random variable $Z \equiv h(Y) = \Phi^{-1}(F(Y\_) + U[F(Y) - F(Y\_)])$, where U is a uniform (0,1) random variable. Then Z is a.s. standard normal. Define $Y^* = F^{-1}(\Phi(Z))$. Then, $Y^* = Y$ a.s., so that Y is given a.s. by a non-decreasing transformation of a standard normal random variable.

12. It does not matter for the example if the previous wave interview month is fixed or has a distribution, provided the *relative* inter-wave interval c is fixed, and current wave outcomes are independent of the timing of the previous wave interview.

13. A Box-Cox transformation of time at risk, $z = 4(t^{1/4} - 1)$, gives a somewhat better approximation in the probit model than log(t),but has no appreciable effect on the accuracy of estimated monthly transition probabilities.

14. The model estimated with interview delay and without imputation, is $\Phi(-4.6639+0.5349 \cdot x+1.1178 \cdot \log(t))$.

15. The AHEAD survey is conducted by the University of Michigan Survey Research Center for the National Institute on Aging; see Soldo, Hurd, Rodgers, and Wallace, 1997.

16. The AMR for the AHEAD sample is the actual death rate between waves for each five-year segment of ages in the initial wave, annualized using the median 25.5 month interval between the waves. The AMR from the life tables is obtained by applying life table death rates by month to the actual months at risk for each individual in the five-year segment of ages in the initial wave to calculate expected deaths between waves. This is annualized. For these calculations, the distribution of months at risk for decedents is assumed to be the same as that for survivors.

17. The construction mimics Figure 5, with life table rates applied using the age, sex, and race of each subject.

18. All Appendices can be found at http://elsa.berkeley.edu/wp/hww/ .

19. Serial 7's asks the subject to subtract 7 from 100, and then to continue subtracting from each successive difference for a total of five subtractions.

20. When an interview was done with a proxy, the cognitive battery was not given, but the interviewee was asked if the respondent was cognitively impaired. In our analysis, we treat proxy interview status as a component of the state that appears as a contemporaneous explanation of cognitive impairment; the coefficient on this variable compensates for differences in the definitions of cognitive impairment.

21. Heteroskedasticity in (10) will arise from selection effects, described later, as well as possibly from a failure of the transformation to fully control the effects of gross reporting errors. When working with this model, we use standard error estimates that are robust with respect to heteroskedasticity of unknown form, and do not attempt direct tests of the implicit error specification underlying the transformation (10).

22. We find that the coefficients of the index in the mortality model are not sensitive to the Edgeworth generalization, and in estimation of the model for wealth change use the first-stage probit models for mortality estimated earlier, with invariance imposed, to obtain the indices that appear in the selection effects.

23. The incidence rates in Table.2 can be converted to crude annual rates via the formula 0.4706*log(1+rate). These rates are uncorrected for population composition effects.

24. This calculation is made from the 1996 life tables and assumes the historical S&P rate of return from 1993 to 1997, and a 7 percent real rate of return on assets after 1997.

Table 1. Approximation Accuracy with Interview Delay

x	Exact Annual Mortality Rate	Approximate Annual Mortality Rate without Interview Delay	Error	Approximate Annual Mortality Rate with Interview Delay and Imputed Contact Time	Error
0	7.71%	7.90%	2.41%	7.93%	2.79%
-1	3.27%	3.06%	-6.55%	3.07%	-6.03%
+1	16.41%	16.71%	1.80%	16.75%	2.05%
Avg	9.13%	9.22%	0.98%	9.25%	1.34%

Note: The approximate annual mortality rate is given by $1 - (1 - \Phi(\theta + \gamma x + \lambda \log(t)))^{12/t}$, where t is the exact or imputed initial contact time and the model is estimated from the data generated by the Monte Carlo experiment..