

NBER WORKING PAPER SERIES

SELF-REPORTED VS. OBJECTIVE MEASURES OF  
HEALTH IN RETIREMENT MODELS

John Bound

Working Paper No. 2997

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
June 1989

I am grateful to Robert Barsky, Charlie Brown, Zvi Griliches, Jerry Hausman, Gary Solon and participants at workshops at the NBER and The University of Michigan for helpful comments and to Patrick Halas for excellent research assistance. Data and computer programs used in this research are available on request. This paper is part of NBER's research program in Aging. Any opinions expressed are those of the author not those of the National Bureau of Economic Research.

NBER Working Paper #2997  
June 1989

SELF-REPORTED VS. OBJECTIVE MEASURES OF HEALTH  
IN RETIREMENT MODELS

ABSTRACT

Labor supply estimates are sensitive to the measures of health used. When self reported measures are used health seems to play a larger role and economic factors a smaller one than when more objective measures are used. While most authors have interpreted these results as an indication of the biases inherent in using self-reported measures, there are reasons to be suspicious of estimates based on more objective measures as well. In this paper I construct a statistical model incorporating both self-reported and objective measures of health. I use the model to show the potential biases involved in using either measure of health or in using one to instrument the other. When outside information on the validity of self-reported measures of health are incorporated into the model estimates suggest that the self-reported measures of health perform better than many have believed.

John Bound  
The Population Studies Center  
The University of Michigan  
1225 South University Ave.  
Ann Arbor, MI 48104  
(313)763-7149

## 1. Introduction

The majority of men who retire before the age of 65 report health as the reason they do so.<sup>1</sup> Though health is, no doubt, an important factor in determining the age at which men retire, there are a variety of reasons not to take these self-reports at face value. It seems quite plausible that men often rationalize retirement in terms of health even when it may occur primarily for other reasons. Myers (1982, 1983) has gone so far as to argue that there is no useful information in self-evaluated health. At the same time, for want of alternative measures, econometric analyses of the labor supply decisions of older men have generally used respondents' self-assessment of their health.<sup>2</sup> There remain important questions about the validity of self-reported measures of health and therefore of the inferences that can be drawn from studies that use them. Kathryn Anderson and Richard Burkhauser (1984) have called the question of the appropriateness of the use of self-reported health measures "the major unsettled issue in the empirical literature on the labor supply of older workers."<sup>3</sup>

There are a number of reasons to be suspicious of any survey response to such questions concerning self-evaluated health. First, respondents are being asked for subjective judgments and there is no reason to expect that these judgments will be entirely comparable across respondents. Second, responses may not be independent of the labor market outcomes we may wish to use them to explain. Third, since health may represent one of the few 'legitimate' reasons for working aged man to be out of work, men out of the labor force may mention health limitations to rationalize their behavior. Lastly, since early retirement benefits are often available only for those deemed incapable of work, men will have a financial incentive to identify themselves as disabled, an incentive that will be particularly high for those for whom the relative rewards from continuing to work are low.

Each of these problems has been noted before, but what does not seem to have been fully realized is that each will lead to different kinds of biases. The lack of comparability across individuals represents measurement error that is likely to lead to our underestimating the impact of

---

<sup>1</sup> See Schwab (1974), Reno (1971), Sherman (1985).

<sup>2</sup> For example, Boskin and Hurd (1978), Quinn (1978), Gordon and Blinder (1981), Diamond and Hausman (1983), Hanoch and Honig (1983), Burkhauser and Quinn (1983), and Hogarth (1988) used responses to the question "Does health limit the amount of kind of work you can perform". Burtless and Moffit (1983, 1985). Burtless (1986) and Gustman and Steinmeier (1986) used responses to a question that asks respondents to rate their health in comparison to others their age.

<sup>3</sup> It should be noted that authors cited above were well aware of the potential problems with using self-reported health measures, but used them because they were what was available.

health on labor force participation, while the endogeneity of self-reported health is likely to lead to our exaggerating its impact. Biases in our estimation of health's impact on outcomes will also induce biases on coefficients of any variables correlated with health. Finally the dependence of self-reported health on the economic environment will induce a bias on estimates of the impact of economic variables on participation, regardless of whether we correctly measure the impact of health itself.<sup>4</sup>

Most work on the labor supply of older men has relied on self-reported measures of health. More recently a variety of authors have argued for the use of more objective indicators of health: responses to questions about specific health conditions or limitations, doctors' reports or information on mortality.<sup>5</sup> While such proxies are, presumably, more objective than are self-reported health measures, it is not clear that the use of such proxies will give us a more accurate indication of the impact of health on labor market outcomes. As long as these health proxies are not perfectly correlated with the aspects of health that impact on economic outcomes, they will suffer from errors in variables problems. With self-reported health measures we have biases working in opposite directions and there is some chance they may tend to cancel out. With objective measures there is only one bias, and, as long as the correlation between the proxy and actual health isn't close to perfect, the bias will be quite substantial.

The issues here are important for our understanding not only of the importance of health, but also of the impact of economic variables on early retirement. Both subjective and objective health indicators are correlated with such things as education, race, pre-retirement earnings and pre-retirement occupation. These factors are also important indicators of early labor market withdrawal. On one interpretation of these facts, it is the poor labor market prospects of these men that induces them to leave the labor force. On another it is their health.

The literature that has compared results using a variety of different health measures has tended to find that health seems to play a smaller role and economic variables a greater one when the more objective proxies are used. Most authors have interpreted these results as an indication of the biases inherent in using self-reported measures.<sup>6</sup> These authors have typically either ignored the possible biases inherent in the use of a proxy, or have assumed that these biases are small in comparison to the ones introduced by the use of self-reported measures.

Others have argued in favor of using self-reported information.<sup>7</sup> These authors emphasize the

---

<sup>4</sup> These comments will be made more precise shortly.

<sup>5</sup> Parsons (1980a, 1980b, 1982), and Anderson and Burkhauser (1984, 1985) used subsequent mortality information, Bartlet and Taubman (1979), Bound (1986) and Stern (1988) used information on the presence of specific conditions, Chirikos and Nestel (1981, 1983) used information on activity limitations.

<sup>6</sup> Parsons (1982), Anderson and Burkhauser (1984, 1985), Chirikos and Nestel (1981), Lambrinos (1981).

<sup>7</sup> See, for example, Sickles and Taubman (1987), Lee (1982.)

flaws inherent in most objective measures of health while pointing to the clinically oriented research supporting the reliability of self-reported health measures.<sup>8</sup> These authors ignore the fact that even if self-reported health is a reliable indicator of actual health, this may not be enough to guarantee that it will give sensible results when used as a proxy for health in retirement equations. At issue is whether self-reported health measures are systematically biased, with those out of work being substantially more likely to report health problems than those working. Were this the case, the use of self-reported measures might give misleading information on the reasons why men retire early even if these measures were highly correlated with actual health.

An alternative to using either self-reported or more objective measures of health is to use the objective measures to instrument the subjective ones. Stern (1988) using this methodology has argued that, if anything, the use of self reported health leads to *underestimates* of the actual impact of health on labor force participation.<sup>9</sup> The limitation of Stern's technique is that it leaves unidentified the impact on retirement of any factors that also directly influence men's reports on their health. In particular this implies that in order to use this technique to identify the impact on retirement of financial incentives, one is forced to assume that men's reports on their health are insensitive to these incentives. If, on the other hand, self reports *are*, as many have thought, sensitive to financial incentives, Stern's technique will understate the impact of these financial incentives on participation.

While there has now been a considerable literature comparing the results obtained with different measures of health, there has been no consistent discussion of the statistical matters involved in such comparisons. The purpose of this paper is to clarify the issues involved in these comparisons. I construct a simple statistical model that incorporates most of the issues involved. In the model, a self-reported measure of health is made endogenous, but is also measured with error. A second, objective, health indicator containing information on subsequent mortality, is only imperfectly correlated with current health. It turns out that, without introducing further information, the model is not identified. With only one objective measure of health, I am not able to simultaneously solve the endogeneity and measurement error problems. What the model does do is to show quite clearly the potential biases involved in any of the above strategies together with the assumptions that are implicit in each. What is more, within the context of the model it turns out to be possible to introduce information on the reliability of self-reported measures while allowing for the kinds of simultaneity mentioned above.

---

<sup>8</sup> Studies by Nagi (1979), Maddox and Douglas (1973), LaRue et al. (1979), and Ferraro (1980) all find that self-reported health is highly correlated with medically determined health status. Research by Mossey and Shapiro (1982) even found that self-reported poor health was a slightly better predictor of subsequent mortality than objectively determined health status.

<sup>9</sup> Haveman, Wolfe and Huang (1989), develop a model of health as an unobservable using a MIMIC framework that is essentially quite similar to Stern's and come to similar conclusions.

Recent literature (Krasker and Pratt (1986), Klepper and Leamer (1984)) on the use of proxy variables has emphasized the use of both implicit and explicit priors for bounding potential biases. Simple versions of such results are well known. Still, perhaps because researchers have believed that evidence on error variances is unobtainable, this approach has rarely been used in practice. Instead econometricians have generally sought exogenous instruments to identify models. This paper expositis an example for which evidence on error variances does in fact exist. What is more, the obvious instrumental variable approach to the problem – using the objective measure of health to instrument the subjective one – turns out not to be a solution at all.

The models developed here build on the multiple indicator modeling used by Goldberger (1972a,b), Joreskog and Goldberger (1975), Griliches (1974), and Chamberlain and Griliches (1975, 1977).<sup>10</sup> Unlike the above cited literature, the health indicators I use are all discrete. MIMIC modeling has been extended to the discrete case by Muthen (1976, 1979, 1983), Lee (1981, 1982) and Avery and Hotz (1982) among others.

The rest of this paper is organized as follows: In section 2, I outline a simple statistical model that will guide both our discussion and estimation. In section 2, for expositional convenience, I ignore the fact that we observe only discrete indicators of either retirement or health status and work as if we actually observed their latent counterparts. In section 3, I discuss issues of estimation showing how to take explicit account of the discrete nature of the various indicators. Section 4 illustrates these issues using data drawn from the Retirement History Survey. A brief conclusion follows.

## 2. A Statistical Model

Consider a simple model of the labor supply of older men. The choice of whether to continue to work for individual  $i$  during time period  $t$  depends on the relative rewards of doing so,  $w_{it}$ , unobserved health status  $\eta_{it}$  and other random components  $\epsilon_{it}$ :

$$lf_{it}^* = \beta_1 w_{it} + \lambda_1 \eta_{it} + \epsilon_{it} \quad (1)$$

In the empirical application a measure of permanent income and demographic information will be included as extra explanatory variables. A man continues to work so long as  $lf_{it}^*$  passes some threshold. We expect  $\beta_1$  to be positive. Since  $\eta_{it}$  is unobserved, the sign of  $\lambda_1$  is arbitrary, but if larger values of  $\eta_{it}$  are associated with better health then we would expect that  $\lambda_1$  should be positive as well.

We also have a second indicator of  $\eta_{it}$ , self-reported health,  $h_{it}^*$ .  $h_{it}^*$  depends on health status  $\eta_{it}$ , but also on the economic rewards for continuing to work,  $w_{it}$ , and, again on other random

---

<sup>10</sup> See Aigner, Hsiac, Kapteyn and Wansbeek (1984) for a review.

components  $\epsilon_{it_2}$ . Dropping the *it* subscript for notational convenience,

$$h^* = \beta_2 w + \lambda_2 \eta + \epsilon_2 \quad (2)$$

A man reports himself limited in his ability to work (or in poor health) if  $h^*$  falls below some threshold. We expect both  $\beta_2$  and  $\lambda_2$  to be positive.

I assume that  $\eta$  is orthogonal to both  $\epsilon_1$  and  $\epsilon_2$  but, as long as there are common unobserved components that affect both  $h^*$  and  $lf^*$ , as there will be if the two are definitionally related or if health limitations act as a rationalization for retirement,  $\epsilon_1$  and  $\epsilon_2$  will be positively correlated.

As long as  $\eta$  and  $w$  are positively correlated, ignoring  $\eta$  in estimating equation (1) will lead to overestimates of the importance of economic incentives in determining labor force participation. The obvious alternative would be to use  $h^*$  as a proxy for  $\eta$  but there are a variety of econometric problems with doing so. The correlation between  $\epsilon_1$  and  $\epsilon_2$  introduces a simultaneity bias while variance in  $\epsilon_2$  introduces errors-in-variables biases on  $\hat{\lambda}_1$ . Errors in estimates of  $\lambda_1$  translate into errors in estimates of  $\beta_1$ , while the dependence of  $h^*$  on  $w$  introduces additional bias on  $\hat{\beta}_1$ . In particular, treating  $lf^*$  and  $h^*$  as if they were observable, letting  $r_{\eta,w}$  represent the correlation between  $\eta$  and  $w$ , and  $\rho$  the correlation between  $\epsilon_1$  and  $\epsilon_2$ , and normalizing  $\lambda_2$  to equal 1, I show in appendix 1 that:

$$plim \hat{\lambda}_1 = \frac{\lambda_1 \sigma_\eta^2 (1 - r_{\eta,w}^2) + \sigma_{\epsilon_1} \sigma_{\epsilon_2} \rho}{\sigma_\eta^2 (1 - r_{\eta,w}^2) + \sigma_{\epsilon_2}^2}$$

$$plim \hat{\beta}_1 = \beta_1 + (\lambda_1 - plim \hat{\lambda}_1) \frac{\sigma_{\eta,w}}{\sigma_w^2} - plim \hat{\lambda}_1 \beta_2$$

As long as  $\rho > 0$ , this correlation will impart an upward bias on  $\hat{\lambda}_1$ , while  $\sigma_{\epsilon_2}^2$  will impart the standard errors-in-variables downward bias on  $\hat{\lambda}_1$ . Which one dominates depends on whether  $\frac{\sigma_{\epsilon_1}}{\sigma_{\epsilon_2}} \rho$  is greater or less than  $\lambda_1$ , on whether  $\eta$  or  $\epsilon_2$  has a greater 'impact' on  $lf^*$ . The bias on  $\hat{\beta}_1$  will depend both on the bias on  $\hat{\lambda}_1$  and on  $\beta_2$ . Thus, it is perfectly possible that even were the errors-in-variables and the simultaneity biases on  $\hat{\lambda}_1$  to cancel, we might still tend to under-estimate  $\beta_1$ .

The above expressions make clear that the biases on  $\hat{\lambda}_1$  and  $\hat{\beta}_1$  may be quite substantial even when  $h^*$  is a reliable measure of  $\eta$ <sup>11</sup> (even when  $\sigma_{\epsilon_2}^2$  is quite small). In fact, other things equal, the more reliable is self-reported health (the smaller is  $\sigma_{\epsilon_2}^2$ ), the greater will be the bias.<sup>12</sup> They

<sup>11</sup> By the reliability of a measure I will mean the degree to which the measure reflects the construct that it is meant to. In the testing literature the term would be "construct validity" rather than reliability. As a measure of the reliability of an indicator I will use the correlation between the indicator and the unobserved construct it is meant to measure.

<sup>12</sup> This statement is obviously not true in the limit. In the limit  $\epsilon_2 \equiv 0 \Rightarrow \rho = 0 \Rightarrow plim \hat{\lambda}_1 = \lambda_1$ .

also make clear that the magnitude and even the direction of the bias depends on multiple pieces of information. It is not good enough to know that self-reported health is highly correlated with actual health to know that estimates using it as a proxy for health will give reliable results. Nor, on the other hand, is it enough that self-reported health may be used to justify retirement or that men's reports on their health are responsive to the economic rewards to working to establish that the use of self reports will exaggerate the role of health. Priors about the kind of bias involved using self-reported health as a proxy for actual health implicitly involve simultaneous judgments about all these factors.

Now consider a somewhat more complete model where we have added an equation to make explicit the correlation between  $w$  and  $\eta$  and have one more indicator of disability status, the date of death  $d^*$ .<sup>13</sup>

$$lf^* = \lambda_1 \eta + \beta_1 w + \epsilon_1 \quad (1)$$

$$h^* = \lambda_2 \eta + \beta_2 w + \epsilon_2 \quad (2)$$

$$d^* = \lambda_3 \nu + \epsilon_3 \quad (3)$$

$$w = \lambda_4 \eta + \epsilon_4 \quad (4)$$

$$\eta = \nu + \mu$$

In this model health,  $\eta$ , has two components, one of which,  $\nu$ , influences both longevity and work capacity (e.g., heart problems), while the other,  $\mu$ , influences only work capacity (e.g., arthritis).<sup>14</sup> The  $\epsilon$ 's are assumed to be uncorrelated with either the other right hand side variables or with  $\eta$  or its components  $\nu$  and  $\mu$ .  $\epsilon_3$  is assumed to be uncorrelated with either  $\epsilon_1$ ,  $\epsilon_2$  or  $\epsilon_4$ . These assumptions imply that  $\epsilon_4$  is also uncorrelated with either  $\epsilon_1$  or  $\epsilon_2$ . Lastly,  $\nu$  and  $\mu$  are assumed to be uncorrelated with each other.<sup>15</sup>

$d^*$  is objective in two ways that  $h^*$  is not:  $d^*$  does not depend on  $w$  nor is  $\epsilon_3$  correlated with  $\epsilon_1$ . Still, as long as the date of death is not perfectly correlated with a man's capacity to work, using it as a proxy for health will not adequately control for health. In particular, normalizing  $\lambda_3$  to equal 1, I show in appendix I that:

<sup>13</sup> Since, in the empirical work, I observe the date of death only for those who die during the sample period, I will treat date of death as latent along with the other indicators of health status.

<sup>14</sup> The implicit assumption imbedded in the various components formulation ( $\eta = \nu + \mu$ ) is that up to factors of proportionality,  $\frac{\lambda_1}{\lambda_2}$  and  $\frac{\lambda_1}{\lambda_3}$ ,  $\nu$  and  $\mu$  enter the labor force, health and compensation equations with identical coefficients. This assumption would seem a natural one if we are thinking of  $\eta$  as capacity for work, and  $h^*$  as a self-report on this capacity. What is more, as the reader will see, allowing for two kinds of health,  $\nu$  and  $\mu$ , pushes identification as it is. Without more health indicators, relaxing the variance component formulation would destroy all hope of identification.

<sup>15</sup> This assumption is mostly definitional.  $\mu$  is the piece of  $\eta$  that is uncorrelated with  $d^*$ .



$$plim \hat{\lambda}_1 = \lambda_1 \frac{\sigma_v^2(1 - r_{v,w}^2)}{\sigma_v^2(1 - r_{v,w}^2) + \sigma_{\epsilon_3}^2}$$

$$plim \hat{\beta}_1 = \beta_1 + (\lambda_1 - plim \hat{\lambda}_1) \frac{\sigma_{v,w}}{\sigma_w^2} + \lambda_1 \frac{\sigma_{\mu,w}}{\sigma_w^2}$$

As long as there are disabling conditions that are not life threatening (e.g., severe back problems, mental illness) controlling for  $d^*$  will still leave a left-out-variable bias on  $\hat{\beta}_1$ , while as long as current capacity for work does not perfectly predict date of death there will be errors-in-variables biases on both  $\hat{\lambda}_1$  and  $\hat{\beta}_1$ .

With two indicators of  $\eta$  we might be tempted to use one to instrument the other, but in this case this won't work. As long as  $\beta_2 \neq 0$  using  $d^*$  to instrument  $h^*$  will purge  $h^*$  of its dependence on  $\epsilon_2$  and so will correctly estimate  $\lambda_1$ <sup>16</sup> but will tend to underestimate  $\beta_1$  by  $\beta_2\lambda_1$ . The instrumental variable procedure uses the projection of  $h^*$  onto  $w$  and  $d^*$  as a proxy for  $\eta$ . What we need, instead, is the projection of  $\eta$  on  $w$  and  $d^*$ . With  $h^*$  as the dependent variable, the estimated coefficient on  $w$  will reflect not only the errors in  $d^*$  but also  $w$ 's direct effect on  $h^*$ ,  $\beta_2$ . This, in turn, will induce the downward bias on  $\beta_1$  of  $\beta_2\lambda_1$ . We could sort all this out if we were to have a consistent estimate of  $\beta_2$ , but this requires either knowledge of the reliability of  $d^*$  as a proxy for  $\eta$  or another indicator of  $\eta$ .<sup>17</sup>

To summarize, simply using mortality information as a health proxy will tend to underestimate the effects of health and overestimate the effects of economic variables on the labor force participation decision. In contrast, using mortality information to instrument self-reported disability status will tend to correctly estimate the impact of health but underestimate the impact of economic variables on such decisions. Finally, simply using self-reported health status can either over or underestimate the impact of either health or economic variables on such decisions.

Without more information either in terms of variables (another indicator of  $\eta$  that does not depend on either  $w$  or  $\epsilon_1$ ) or prior restrictions on some parameters the model is unidentified. This is most clearly seen if we examine the variance-covariance matrix implied by the model:

---

<sup>16</sup> This result depends on the assumption that  $\lambda_2$  has been normalized to equal one. What will always be consistently estimated is  $\frac{\lambda_1}{\lambda_2}$ .

<sup>17</sup> On the other hand, we cannot use  $h^*$  as an instrument for  $d^*$  since  $h^*$  is correlated with  $\epsilon_1$ .



Though this procedure may be convenient and appropriate for many purposes, it is not for our own. Measurement error for indicator variables is more problematic than it is for continuous ones. The iid normal assumption no longer makes any sense, with multiple indicators (of, for example, the date of death) we would have to account for the interdependence of these errors, and implicit normalizations will change as we move across specifications.

As an alternative we can model  $h^*$  and  $d^*$  as explicitly endogenous, estimating three semi-reduced form specifications with  $lf^*$ ,  $h^*$  and  $d^*$  as dependent variables, and recover the parameters of interest from the estimated reduced form parameters. This strategy is in line with the recent literature on estimating simultaneous limited dependent variable models using two stage methods<sup>19</sup>. Take for example, the case where we want to use  $h^*$  as a proxy for  $\eta$ . Estimating such an equation by OLS is equivalent to first estimating reduced forms for  $lf^*$  and  $h^*$ :

$$lf^* = \pi_1 w + u_1 \quad (1')$$

$$h^* = \pi_2 w + u_2 \quad (2')$$

OLS estimates of  $\lambda_1$  and  $\beta_1$  can be obtained as:

$$\hat{\lambda}_1 = \frac{\hat{\sigma}_{u_1, u_2}}{\hat{\sigma}_{u_2}^2}$$

$$\hat{\beta}_1 = \hat{\pi}_1 - \hat{\pi}_2 \hat{\lambda}_1$$

Even if we observe only discrete indicators of  $lf^*$  or  $h^*$  we can still, subject to two extra normalizations, estimate (1') and (2') and derive the OLS estimates of  $\lambda_1$  and  $\beta_1$  from the reduced form estimates of  $\pi_1$ ,  $\pi_2$  and  $u_1, u_2$ .

To estimate (1) using  $d^*$  rather than  $h^*$  as a proxy for  $\eta$  follows similar lines. We first estimate a reduced form for  $d^*$ :

$$d^* = \pi_3 w + u_3 \quad (3')$$

The OLS estimates of  $\hat{\beta}_1$  and  $\hat{\lambda}_1$  are then:

$$\hat{\lambda}_1 = \frac{\hat{\sigma}_{u_1, u_3}}{\hat{\sigma}_{u_3}^2}$$

$$\hat{\beta}_1 = \hat{\pi}_1 - \hat{\pi}_3 \hat{\lambda}_1$$

In a similar way, to estimate (1) by using  $h^*$  to proxy  $\eta$ , but then instrumenting  $h^*$  with  $d$  amounts to using all three reduced form equations. Here we have:

---

<sup>19</sup> See Newey (1987) for a review.

$$\hat{\lambda}_1 = \frac{\hat{\sigma}_{u_1, u_3}}{\hat{\sigma}_{u_2, u_3}}$$

$$\hat{\beta}_1 = \hat{\pi}_1 - \hat{\pi}_2 \hat{\lambda}_1$$

For the general case it will be helpful to introduce a bit more notation. Write  $\eta$  as  $\gamma w + v$  where  $\gamma = \frac{\lambda_1 \sigma_v^2}{\sigma_\eta^2}$  and  $v$  is orthogonal to  $w$  by construction, and write  $\nu$  as  $\tau \gamma w + \xi$  where  $\tau \equiv \frac{\sigma_\xi^2}{\sigma_\eta^2}$  and  $\xi$  is orthogonal to  $w$  by construction. The reduced form  $\pi$ 's can be rewritten in terms of the  $\lambda$ 's, the  $\beta$ 's,  $\gamma$  and  $\tau$ :

$$\pi_1 = \lambda_1 \gamma + \beta_1$$

$$\pi_2 = \lambda_2 \gamma + \beta_2$$

$$\pi_3 = \lambda_3 \tau \gamma$$

while the reduced form errors, the  $u$ 's, can be written in terms of the  $\lambda$ 's, the  $\epsilon$ 's,  $v$  and  $\xi$ .

$$u_1 = \lambda_1 v + \epsilon_1$$

$$u_2 = \lambda_2 v + \epsilon_2$$

$$u_3 = \lambda_3 \xi + \epsilon_3$$

and a residual covariance matrix:

$$\begin{array}{ccc} \lambda_1^2 \sigma_v^2 + \sigma_{\epsilon_1}^2 & & \\ \lambda_1 \lambda_2 \sigma_v^2 + \rho \sigma_{\epsilon_1} \sigma_{\epsilon_2} & \lambda_2^2 \sigma_v^2 + \sigma_{\epsilon_2}^2 & \\ \lambda_1 \lambda_3 \sigma_\xi^2 & \lambda_1 \lambda_3 \sigma_\xi^2 & \lambda_3^2 \sigma_\xi^2 + \sigma_{\epsilon_3}^2 \end{array}$$

$\sigma_v^2$  equals  $\sigma_\eta^2 [1 - \frac{\lambda_1^2}{\sigma_\eta^2}]$  and  $\sigma_\xi^2$  equals  $\tau \sigma_\eta^2 [1 - \frac{\lambda_3^2 \tau}{\sigma_\eta^2}]$ .

This "reduced form" representation of the model can be estimated even if we observe only discrete indicators of  $lf^*$ ,  $h^*$  and  $d^*$ . Assuming the  $u$ 's are jointly normally distributed, the model is a trivariate probit. As is standard, we can estimate the  $\pi$ 's only up to a scale factor, but otherwise everything else goes through. As long as the model in question is exactly identified one can then solve for the structural parameters in terms of the reduced form estimates. Standard errors can be obtained using the so-called delta method.<sup>20</sup> The various estimation strategies outlined above

<sup>20</sup> For overidentified models, the structural parameters can be obtained from the reduced forms using minimum  $\chi^2$  procedures (see Ferguson, 1958). In particular let  $\theta$  be the vector of reduced form parameters and  $\delta$  be the vector of structural parameters. Estimation of the reduced form gives us estimates of both  $\theta$  and the variance covariance matrix of  $\hat{\theta}$ ,  $\hat{V}(\hat{\theta})$ . Then each of our models implies that  $f(\delta) = \theta$ , where  $f$  is a known function (that depends on the model). Minimum  $\chi^2$  methods estimate  $\delta$  by minimizing the quadratic form:

$$[\hat{\theta} - f(\delta)]^T \hat{V}(\hat{\theta})^{-1} [\hat{\theta} - f(\delta)]$$

correspond to different restrictions on the various parameters of the model. If  $h^*$  is used as a proxy for  $\eta$  the parameters of interest are exactly identified within the first two equations by setting both  $\beta_2$  and  $\sigma_{\epsilon_2}^2$  to 0. If  $d^*$  is used as a proxy for  $\eta$  the parameters are exactly identified within the first and third equation by setting  $\sigma_{\epsilon_3}^2$  to 0. When we use  $d^*$  as an instrument for  $h^*$ , the model is exactly identified only if we use all three equations. Here estimates are derived by setting  $\beta_2 = 0$ .

Alternatively, we can explicitly introduce outside information to identify the model. For example, we might introduce information on the reliability with which either  $d^*$  or  $h^*$  reflect the capacity to work to fix either  $\lambda_2$  or  $\tau\lambda_3$ . Alternatively, if we had some information about the relative reliability of  $h^*$  and  $d^*$  we could also use this information.

#### 4. Implementation

Data to estimate the above model is drawn from the Retirement History Survey (RHS). The RHS followed a nationally representative sample of 8131 men<sup>21</sup> aged 58-63 in the base year for 10 years 1969-1979. Information was collected on, among other things, respondents' work lives, their pension eligibility and their health. Respondents were matched to their social security records and, for those who died during the interview years, information is available on their date of death.<sup>22</sup> As of 1969, 14.5% of the men reported that they were fully retired while another 8.1% reported that they were partially retired. 75% of those fully retired and 53% of those partially retired identified health as the reason for their retirement.

I will focus my attention on the labor force participation decision as of the 1969 survey week and will restrict my attention to those who were either currently working in or whose primary affiliation had been the private sector.<sup>23</sup> The RHS survey included two different questions asking respondents to evaluate their overall health: "Does health limit the amount or kind of work you can do?" and "Is your health better, worse, or the same as that of other people your age?". I report results using each.<sup>24</sup> For a more objective measure of health, I use the information on the date

---

The variance covariance matrix for these estimates is just:

$$[f'X'V(\theta)^{-1}f']^{-1}$$

With exactly identified models it may be simpler to just solve for  $\hat{\delta}$  ( $\hat{\delta} = f^{-1}(\hat{\theta})$ ) and use the so called delta method to derive standard errors. In practice I used both methods, one as a check on the other.

<sup>21</sup> Unmarried women were also followed, but we will restrict attention to the men.

<sup>22</sup> The mortality information comes from the Social Security's records. Available evidence suggests that these records are highly reliable (see Sickles and Taubman for a discussion of this issue. The mortality data was kindly provided by Kathryn Anderson and Richard Burkhauser.

<sup>23</sup> While I have information on men's social security earnings and can therefore calculate potential social security benefits, I do not have comparable information for those working in the public sector.

<sup>24</sup> As long as one allows the two self-reported health measures to be freely correlated, the extra self-reported health measure does not aid in identification. The extra measure of health adds one reduced form  $\pi$  and three reduced from cross equation correlations, but adds an equal number of parameters, a  $\lambda$ , a  $\beta$ , and two cross equation correlations.

of death. For economic variables I use the individuals' social security earning history to calculate measures of permanent income together with measures of the total compensation from continuing to work for another year. Following Burkhauser (1979, 1980) I calculate three components to compensation: 1) annual earnings, 2) the change in the present discounted value of expected social security benefits and 3) the change in the present discounted value of pension benefits. The details of how I constructed both the lifetime income and total compensation variables are contained in appendix 2. For demographic variables I used age as of 1969, marital status, race and educational attainment.

For estimation I restricted the sample to men who were or had been employed in the private sector and who had complete data on items used in estimation. The final sample size was 6022.<sup>25</sup> Sample statistics together with variable definitions are reported in Table 1. 82% of the sample were in the labor force as of the survey, 35% reported health limitations on their ability to work while 20% reported being in worse health than other men their age.

As shown in Table 2, 27.7% of the sample had died by 1979. Of those who died within two years of the survey, more than two-thirds report health limitations on their ability to work, while more than half report being in poor health. Over 45% were out of the labor force. This contrasts with those who were still alive at the end of the survey. Of these, roughly one quarter report health limitations on their ability to work, 15% report being in poor health while 14% are out of the labor force. There is a clear pattern of association between mortality and our other indicators of health status, but the 'fit' is far from perfect.

Table 3 reports estimates of the effects of compensation, lifetime income and the demographic factors on the probability that a respondent will be in the labor force as of the 1969 survey week, will report health limitations on his ability to work, will report himself in poor health or will die during the sample period. Date of death ( $d$ ) is polychotomous, with seven categories ordered so that higher values are associated with a later death (e.g., 7 denotes alive in 1979, 6 died between 1974 and 1979, 5 died in 1973, etc.). Limit ( $l$ ), health ( $h$ ) and labor force participation ( $lf$ ) are dichotomies, 0-1 variables with 1's representing no health limitations, health as good or better than average and participation respectively, while  $\log(wage)$  and  $\log(income)$  represent the natural log of annual compensation and permanent income respectively. Other exogenous variables included in the specification are the age of the respondent, his educational attainment, his race and his marital status. The coefficients on  $\log(wage)$  imply that those men who would be well compensated for staying in the work force another year are more likely to do so but are also less likely to report health limitations on their ability to work, are less likely to report being in poor health and more

---

<sup>25</sup> Eliminating those not employed in the private sector eliminated 1681 men, insisting that a man's discounted lifetime earnings be above \$100 and that his annual salary be above \$100 eliminated 12 men, while 416 had bad data of some kind or another.

likely to survive the sample period. Age lowers the probability that a man will be working while raising the probability that he will suffer health limitations or will die during the sample period. The other demographic variables have their expected effects. More educated and married men are more likely to work, but are less likely to report health limitations or to die during the sample period.

Estimated correlations across the equations are also reported. All correlations are positive, as expected. The correlation between either self-reported health limitations or poor health and labor force withdrawal is very strong (.707 and .717), with the correlation between date of death and the other indicators being relatively weak. In particular, the relatively weak correlations between date of death and either self-reported health limitations or labor force status suggest that mortality itself may not be that highly correlated with current disability status. This should not be very surprising. The most common kinds of health conditions associated with self-reported disability are musculoskeletal (e.g., arthritis) and these will, in general, not be life threatening. Moreover, the leading causes of death among men in this age range, heart disease and cancer, often manifest themselves only shortly before death. Lastly we note that, contrary to some researchers impressions, there seems to be very little difference between the two self-reported health measures. *Log(wage)* does seem to have somewhat stronger impact on self-reported health limitations than on self-reported poor health, suggesting that the endogeneity problem may be greater for the former variable, but the differences are not large and the cross equation correlation patterns are almost identical.

The estimates reported in Table 3 will be the basis for our estimation of the factor model, but it may be worthwhile to first consider some more standard estimates of the impact of health and the replacement rate on labor force participation. Table 4 reports various such estimates. Column 1 reports estimates with no controls for health status. The coefficient of .193 on annual compensation corresponds to an elasticity of non-participation with respect to compensation (evaluated at the sample proportion) of a substantial .28. The specifications reported in columns 2 and 3 include measures of self-reported health. These two indicator variables both have a very strong negative impact on participation, while the coefficient on *log(wage)* drops by between 42% and 61%. Column 4 replaces self-reported health with six indicators of the date of death. These pick up significant coefficients but, judging by the values of the log-likelihood statistics for the various models, have nothing like the impact on participation that the self-reported measures do. The impact of including the mortality indicators on the estimated effect compensation is negligible.

There are two problems with the kinds of specifications reported in Table 4. First, since the units in which the health proxies are measured change, it is hard to make comparisons across specifications of the implied impact of health on participation. Furthermore each specification implies something different about the measurement errors involved in the proxies. To compare the impact of health across specifications we would need to take this unmeasured component of health

into account. This is more easily done in the context of the explicit factor model outlined in the preceding section.

Table 5 reports six sets of estimates of the labor force participation equation (1), but with the estimates now based on the estimates reported in Table 3. We normalize  $\sigma_{lf}^2$ ,  $\sigma_{\eta}^2$ ,  $\sigma_{h^*}^2$ ,  $\sigma_{d^*}^2$  and  $\sigma_{\eta}^2$  to be 1. With these normalizations  $\lambda_1$  and  $\lambda_2$  can be interpreted as the correlation between  $\eta$  and the two dependent variables while  $\lambda_3\tau$  can be interpreted as the correlation between  $\eta$  and  $d^*$ . Choosing a normalization that remains constant across specifications facilitates the comparison of the impact of health on participation.<sup>26</sup>

Column 1 reports the ‘OLS’ estimates of  $\beta_1$ : the effect of  $\log(wage)$  on participation that ignores the effect of health. The estimate simply rescales to take account of the change in the normalization (from  $\sigma_{u_1}^2 = 1$  to  $\sigma_{lf}^2 = 1$ .) the comparable one in table 4 and is reported as a baseline. Columns 2 through 4 report estimates that use  $d^*$  and then either  $l^*$  or  $h^*$  as proxies for  $\eta$ , while columns 5 and 6 report estimates where  $d^*$  is used to instrument  $l^*$  or  $h^*$ . Using  $d^*$  as a proxy for  $\eta$  lowers the estimate of  $\beta_1$  by just 15%. Using  $l^*$  or  $h^*$  as the proxy for  $\eta$  dramatically lowers the estimates of  $\beta_1$  to essentially 0, and more than doubles the estimates of  $\lambda_1$ . Whether  $l^*$  or  $h^*$  is used makes very little difference for the estimates of  $\lambda_1$ , but using  $l^*$  does produce somewhat lower estimates of  $\beta_1$ . Again we see the suggestion that endogeneity is more of a problem with  $l^*$  than  $h^*$ , but the differences aren’t large. The estimates of  $\lambda_1$  suggest that, among men this age, the variation in health across individuals can explain 50% of the variation in labor force behavior.<sup>27</sup> Finally, using  $d^*$  to instrument  $l^*$  or  $h^*$  lowers the estimates of  $\beta_1$  and raises the estimates of  $\lambda_1$  still further.

The disparities in the three sets of estimates reported in Table 5 are enormous. The results using the mortality information provide a lower bound on the impact of health on retirement while providing an upper bound on the impact of compensation on retirement. On the other hand, using the mortality information to instrument the self-reported measures gives us upper bounds on the impact of health and lower bounds on the impact of compensation.

In Tables 6, I take a different tack. Estimates for the full model are presented under various assumptions about the magnitude of  $\lambda_2$ . At one extreme we have the case where  $\lambda_3 = 1$  which amounts to simply proxying health with mortality. At the other extreme we have the case where  $\beta_2 = 0$  which amounts to instrumenting self reported health with the mortality information. When  $l^*$  is used, estimates with  $\beta_2 = 0$  imply estimates of  $\rho_{\epsilon_1, \epsilon_2}$  that are outside the unit circle. Alternative estimates with  $\rho_{\epsilon_1, \epsilon_2}$  set equal to -1 are also reported. Lowering  $\lambda_2$  raises the estimates of  $\beta_1$  and

<sup>26</sup> It is worth noting that these normalizations are not the standard ones. Using  $l^*$ ,  $h^*$  or  $d^*$  as proxies for  $\eta$  would typically involve normalizing either  $\lambda_2$  or  $\lambda_3$  to equal one, while using  $d^*$  to instrument either  $l^*$  or  $h^*$  would typically involve normalizing  $\lambda_2$  to equal one.

<sup>27</sup> Recall that  $\lambda_1$  can be interpreted as the correlation between  $\eta$  and  $lf^*$ . Thus  $\lambda_1^2 = R_{\eta, lf^*}^2 = .71^2 = .50$  and  $.72^2 = .52$ .



$\beta_2$ , while lowering the estimates of  $\lambda_1$  and  $\lambda_3$ . Varying  $\lambda_3$  would produce a similar pattern. Here, lowering  $\lambda_3$  would lower the estimates of  $\beta_1$  and  $\beta_2$  while raising the estimates of  $\lambda_1$  and  $\lambda_2$ .

To narrow the range of estimates we need to introduce more information. First, it would seem plausible that  $l^*$  would be at least as correlated with  $\eta$  as  $d^*$  would be. After all, the primary objection to the  $l^*$  was not that it was poorly correlated with  $\eta$ , but that  $\epsilon_2$  was correlated with  $\epsilon_1$ . It would also seem plausible that compensation should have a positive impact on participation. Imposing the restriction that  $\lambda_2 \geq \lambda_3\tau$  and that  $\beta_1 \geq 0$  tightens the feasible range over the parameter space to the middle rectangles in Table 6.

To get still further we can explicitly introduce information from some of the reliability studies mentioned in the introduction. Most of the studies cited above were done on older individuals and it is unclear how well results generalize to a working aged population. Perhaps the study that is the most appropriate for our purpose is the one by Nagi (1979). Nagi, using a sample of older working aged men, compared their reports of physical limitations to those of attending physicians. He compared doctors' reports to self reports along a number of dimensions including the capacity for work. The polychoric correlation between respondents' and doctors' reports implied by this study is .53 (.00).<sup>28</sup> Thus, even if we took the doctors' reports of capacity to work as completely accurate, the correlation between self-assessments and actual health would be .53. Should we assume that the doctors are no more accurate than the individuals themselves, but that the errors in the two reports are independent, we get a correlation of about .7.<sup>29</sup> Thus it is unclear how well the Nagi results can be generalized to the RHS sample. What is more, it is unclear whether our results can be generalized either to other age groups or other time periods. The reliability with which with self-reported health or mortality information indicates capacity for work could easily be a function of both age and cohort.

Given the normalization  $\sigma_l^2 \equiv 1$ ,  $\lambda_2$  is equal to the correlation between  $\eta$  and  $l^*$ . Thus, estimates of the correlation between  $l^*$  and  $\eta$  translate directly into estimates of  $\lambda_2$  that can then be imposed to resolve the identification issue. Table 6 shows that a correlation of .53 comes close to reproducing the results where  $\lambda_2 = \lambda_3\tau$ . A correlation near .7 implies an estimate of  $\beta_1$  of .06 and an estimate of  $\lambda_1$  of .6.

It may be worthwhile to turn back for a while to the issue of whether self-reported or objective measures of health give more accurate estimates of the impact of either health of compensation

<sup>28</sup> Nagi had both doctors and the older men themselves rate their capacity to work on a six point scale and reports the resulting 6 by 6 table. To calculate the correlation between the two reports I assumed that the scales represented discrete indicators of underlying latent variables and estimated the implied correlation between these latent factors by maximum-likelihood. See Olsson (1979).

<sup>29</sup> A few cautions are probably in order. Nagi's study was based not on a random sample of the population but on referrals from physicians and clinics. 90% reported some kind of work limitations. While we would normally think that truncating on a variable should reduce the correlation between this variable and others, the fact that all the men in Nagi's sample had recently been involved with some clinic or rehabilitation service might have an impact on the accuracy with which these men reported on their capacity for work.

of retirement behavior. Whether we impose the restrictions that  $\lambda_2 \geq \lambda_3\tau$  and that  $\beta_1 \geq 0$  or take the Nagi results as some kind of indication about plausible magnitudes for  $\lambda_2$  we come to the conclusion that the use of the mortality information as a proxy for health exaggerates the impact of economic factors and underestimates the impact of health by substantial margins. We also see pretty clear signs that the use of self reported health limitations exaggerates the impact of health on participation and biases the estimates of economic factors downward. This is the conclusion we reach comparing estimates using health limits to those using bad health or by simply noting that when health limits is used  $\hat{\beta}_1$  is of the 'wrong' sign. The Nagi results suggest that self-reported health measures lead to estimates that exaggerate the impact of health and underestimates the impact of economic factors on retirement by substantial margins. What these results would seem to do is to question any presumption in favor of either the objective or the self reported measures of health. They also suggest that procedures that amount to instrumenting self-reported health with objective measures are likely to give results that are even more misleading than the ones that 'naively' use the self-reported measures themselves.

## Conclusion

Various authors have used 'objective' rather than self-reported indicators of health status in the estimation of labor force participation equations. Typically such proxies yield results that suggest that health plays a smaller role and economic variables a larger one than estimates based on self-reported measures would suggest. We have seen that there is another interpretation of this pattern. Proxies commonly used, such as the date of subsequent mortality, are bound to be imperfectly correlated with health status. Even a moderate amount of measurement error in such proxies can easily lead to the conclusion that the self-reported measure will give a more accurate picture of the impact of health and financial incentives on labor supply.

The search for 'objective' or exogenous indicators of health status may have been a bit misplaced. For example, even if we were to have information from physical exams available, using these alone in retirement or labor force equations would not eliminate the biases involved. We could always imagine that there were conditions or circumstances not included in the examination that would in some circumstances imply that an individual was unable to work. Perhaps more importantly, results from physical exams would not control for the specific demands of a person's job environment or the interaction between a person's skills and his health limitations. Moreover, the severity of certain medical conditions (e.g., pain) are impossible to measure through objective tests.<sup>30</sup> Part of the problem with 'objective' measures of health is that they measure health rather than capacity for work. As a result, information on the reliability of these measures can ultimately never be informative about the validity of these measures as proxies for work capacity.

---

<sup>30</sup> Were reliable objective measures of work capacity available, we wouldn't see the kind of controversy that we do surrounding the medical screening that individuals go through to qualify for disability benefits.

On the other hand, self-reports of health limitations, while being direct measures of the capacity to work, suffer from a host of other problems. Most importantly, if men rationalize retirement decisions made for other reasons by identifying themselves as incapable of work or in poor health, self reported measures of health are likely to exaggerate the importance of health in retirement decisions. What is more, the significance attributable to health may mask the effect of economic variables.

While these issues have been widely recognized, most researchers have approached the issue by presuming that the errors-in-variables problems inherent in the use of objective measures of health is small in comparison to the endogeneity problems inherent in the use of self-reported measures. Yet, what informs this presumption is does not get spelled out. More recently, several researchers have combined information on self-reported and more objective measures of health by using the objective measures to instrument the self-reported ones. While this procedure would seem to have much appeal, we have seen that it will tend to underestimate the impact of economic variables on retirement. In fact, the empirical results reported in table 5 suggest that using objective measures to instrument subjective ones may actually exacerbate the biases that occur when self-reported measures are used alone.

In this paper I have constructed a simple statistical model that incorporates information from both self-reported and objective measures of health. What this model makes clear is that even with these multiple indicators of health status, the impact of health on labor force participation is not identified. Without further information it is not possible to determine whether objective or self reported measures of health give more accurate indications of the importance of health in determining retirement behavior. In the final section of the paper I suggest ways of introducing outside information into the model to help resolve the identification problem. Results using this auxiliary information suggests that both the errors-in-variables problem associated with the objective measures of health and the endogeneity problem associated with the self-reported measures are substantial. What is more, combining information from both sources by instrumenting self-reported measures of health with the objective measures can make things worse.

At the pragmatic level, these results suggest that neither objective nor self-reported health measures can be counted upon to provide reliable estimates of the impact of health or other variables on the labor force attachment of older men. On a more positive note, the results do suggest that objective and self reported measures lead to biases in opposite directions. Thus, results using both can be used to bound on the actual effect of health and other factors on labor force behavior.

## Appendix 1: Derivation of Potential Bias's Using Different Proxies for Health

In this appendix I will derive the bias formulas asserted in section 2 of the text.

Recall the equation system:

$$lf^* = \lambda_1 \eta + \beta_1 w + \epsilon_1 \quad (1)$$

$$h^* = \lambda_2 \eta + \beta_2 w + \epsilon_2 \quad (2)$$

$$d^* = \lambda_3 \nu + \epsilon_3 \quad (3)$$

$$w = \lambda_4 \eta + \epsilon_4 \quad (4)$$

$$\eta = \nu + \mu$$

I assume that both  $\eta$  and  $w$  are uncorrelated with the  $\epsilon$ 's, that both  $\epsilon_3$  and  $\epsilon_4$  are uncorrelated with either  $\epsilon_1$ ,  $\epsilon_2$  or each other, but allow  $\epsilon_1$  and  $\epsilon_2$  to be correlated with each other and denote this correlation  $\rho$ . Finally, I assume that  $\mu$  and  $\nu$  are uncorrelated with each other and with the  $\epsilon$ 's.

Before proceeding it will help to introduce some additional notation: Write  $\eta$  as  $\gamma w + \nu$ ,  $\nu$  as  $\gamma_\nu w + \xi$  and  $\mu$  as  $\gamma_\mu w + \zeta$ .  $\gamma \equiv \frac{\sigma_{\eta w}}{\sigma_w^2}$ ,  $\gamma_\nu \equiv \frac{\sigma_{\nu w}}{\sigma_w^2}$  and  $\gamma_\mu \equiv \frac{\sigma_{\mu w}}{\sigma_w^2}$ , while  $\nu$ ,  $\xi$  and  $\zeta$  are all orthogonal to  $w$ . Note that the 0 correlation between  $\nu$  and  $\mu$  implies that  $\gamma = \gamma_\nu + \gamma_\mu$ , while  $\sigma_\nu^2 = \sigma_\eta^2(1 - r_{\eta,w}^2)$ ,  $\sigma_\xi^2 = \sigma_\nu^2(1 - r_{\nu,w}^2)$  and  $\sigma_\zeta^2 = \sigma_\mu^2(1 - r_{\mu,w}^2)$ .

Substituting for  $\eta$  and  $\nu$ :

$$lf^* = [\lambda_1 \gamma + \beta_1]w + [\lambda_1 \nu + \epsilon_1] \quad (1')$$

$$= \pi_1 w + u_1$$

$$h^* = [\lambda_2 \gamma + \beta_2]w + [\lambda_2 \nu + \epsilon_2] \quad (2')$$

$$= \pi_2 w + u_2$$

$$d^* = [\lambda_3 \gamma_\nu]w + [\lambda_3 \xi + \epsilon_3] \quad (3')$$

$$= \pi_3 w + u_3$$

The various estimates of  $\beta_1$  and  $\lambda_1$  can be written in terms of the  $\pi$ 's and the covariances between the  $u$ 's. In particular, using  $h^*$  as a proxy for  $\eta$  gives:

$$\hat{\lambda}_1 = \frac{\hat{\sigma}_{u_1, u_2}}{\hat{\sigma}_{u_2}^2}$$

$$\hat{\beta}_1 = \hat{\pi}_1 - \hat{\pi}_2 \hat{\lambda}_1$$

Using  $d^*$  as a proxy for  $\eta$  gives:

$$\tilde{\lambda}_1 = \frac{\hat{\sigma}_{u_1, u_3}}{\hat{\sigma}_{u_3}^2}$$

$$\tilde{\beta}_1 = \hat{\pi}_1 - \hat{\pi}_3 \tilde{\lambda}_1$$

Using  $h^*$  as a proxy for  $\eta$ , but instrumenting with  $d^*$  gives:

$$\hat{\lambda}_1 = \frac{\hat{\sigma}_{u_1, u_3}}{\hat{\sigma}_{u_2, u_3}}$$

$$\hat{\beta}_1 = \hat{\pi}_1 - \hat{\pi}_2 \hat{\lambda}_1$$

Writing out the probability limits of these estimates we have:

Using  $h^*$  as a proxy for  $\eta$ :

For  $\hat{\lambda}_1$ :

$$\begin{aligned} plim \hat{\lambda}_1 &= \frac{plim \hat{\sigma}_{u_1, u_2}}{plim \hat{\sigma}_{u_2}^2} \\ &= \frac{\lambda_1 \lambda_2 \sigma_v^2 + \sigma_{\epsilon_1, \epsilon_2}}{\lambda_2^2 \sigma_v^2 + \sigma_{\epsilon_2}^2} \end{aligned}$$

Normalizing  $\lambda_1$  to equal 1 and substituting for  $\sigma_v^2$  we get:

$$plim \hat{\lambda}_1 = \frac{\lambda_1 \sigma_\eta^2 (1 - r_{\eta, w}^2) + \sigma_{\epsilon_1} \sigma_{\epsilon_2} \rho}{\sigma_\eta^2 (1 - r_{\eta, w}^2) + \sigma_{\epsilon_2}^2}$$

For  $\hat{\beta}_1$ :

$$\begin{aligned} plim \hat{\beta}_1 &= plim \hat{\pi}_1 - plim \hat{\pi}_2 plim \hat{\lambda}_1 \\ &= \beta_1 + (\lambda_1 - plim \hat{\lambda}_1) \gamma - plim \hat{\lambda}_1 \beta_2 \\ &= \beta_1 + (\lambda_1 - plim \hat{\lambda}_1) \frac{\sigma_{\eta, w}}{\sigma_w^2} - plim \hat{\lambda}_1 \beta_2 \end{aligned}$$

Using  $d^*$  as a proxy for  $\eta$ :

For  $\hat{\lambda}_1$ :

$$\begin{aligned} plim \hat{\lambda}_1 &= \frac{plim \hat{\sigma}_{u_1, u_3}}{plim \hat{\sigma}_{u_3}^2} \\ &= \frac{\lambda_1 \lambda_3 \sigma_\xi^2}{\lambda_3^2 \sigma_\xi^2 + \sigma_{\epsilon_2}^2} \end{aligned}$$

Normalizing  $\lambda_3$  to equal 1 and substituting for  $\sigma_\xi^2$  we get:

$$plim \hat{\lambda}_1 = \lambda_1 \frac{\sigma_v^2(1 - r_{v,w}^2)}{\sigma_v^2(1 - r_{v,w}^2) + \sigma_{\epsilon_3}^2}$$

For  $\hat{\beta}_1$ :

$$\begin{aligned} plim \hat{\beta}_1 &= plim \hat{\pi}_1 - plim \hat{\pi}_3 plim \hat{\lambda}_1 \\ &= \beta_1 + (\lambda_1 - plim \hat{\lambda}_1)\gamma_\nu + \lambda_1 \gamma_\mu \\ &= \beta_1 + (\lambda_1 - plim \hat{\lambda}_1) \frac{\sigma_{\nu,w}}{\sigma_w^2} + \lambda_1 \frac{\sigma_{\mu,w}}{\sigma_w^2} \end{aligned}$$

IV using  $d^*$  to instrument  $h^*$ :

For  $\hat{\lambda}_1$ :

$$\begin{aligned} plim \hat{\lambda}_1 &= \frac{plim \hat{\sigma}_{u_1, u_3}}{plim \hat{\sigma}_{u_2, u_3}} \\ &= \frac{\lambda_1 \lambda_3 \sigma_\xi^2}{\lambda_2 \lambda_3 \sigma_\xi^2} \\ &= \frac{\lambda_1}{\lambda_2} \end{aligned}$$

Normalizing  $\lambda_1$  to equal one we have:

$$plim \hat{\lambda}_1 = \lambda_1$$

For  $\hat{\beta}_1$ :

$$\begin{aligned} \text{plim } \hat{\beta}_1 &= \text{plim } \hat{\pi}_1 - \text{plim } \hat{\pi}_2 \text{plim } \hat{\lambda}_1 \\ &= \beta_1 - \frac{\lambda_1}{\lambda_2} \beta_2 \\ &= \beta_1 - \lambda_1 \beta_2 \end{aligned}$$

## Appendix 2: Specification of the Retirement Equation

In specifying the retirement equation in this paper I followed the existing literature. I imagine that a man continues to work so long as the compensation for doing so,  $w$ , is above a reservation wage  $w^*$ . I also assume that the log of the reservation wage is a linear function of the log of lifetime income  $y$ , health,  $\eta$ , and demographics,  $X$ . This specification results in a very simple retirement or participation equation: a man continues to work if:

$$\ln(w) > \ln(y) + X'\beta + \lambda\eta + \epsilon$$

To calculate the presented discounted value of lifetime earnings I include three components: 1) the present discounted value of past earnings as of time  $t$ ; 2) the present discounted value of social security benefits were the man to retire as of time  $t$ ; and, 3) for those who indicated that they were eligible for a private pension, the present discounted value of this pension as of time  $t$ . Total compensation is calculated as the change in the present discounted value of life time income were the man to work one more year.

I focus on the participation decision in 1969, calculating both lifetime income and compensation as if the man had continued working up until that point. For my measure of past earnings I used the social security earnings history. For men who were not working as of 1969 I projected forward their earnings from the year before they reported having ceased working assuming that these earnings would have remained constant. For men still in the work force in 1969 I used their 1968 earnings. The present discounted value of earnings was then calculated using the earnings history back to 1951 together with the prime interest rate for each year between 1951 and 1969. Social security earnings are truncated at the taxable maximum, but the RHS contains information on which quarter the maximum was attained. This allowed me to impute earnings for those above the maximum. For men who were currently married their wives' earnings were included in the calculation.

Social security benefits were calculated using the 1969 law and assuming the man worked through 1968. To calculate the present discounted value of benefits I assumed that benefits would continue to grow with inflation (5%) and discounting assuming a nominal interest rate of 7% and model life tables. For married men, I included both dependent and survivor benefits in the calculation. Thus, for married men the calculation is:

$$PDV_{69} = PIA_{69} \times \sum_{t=0}^{t=\infty} B^t [1 + .5\pi_{t,h}\pi_{t,w} - (1 - \pi_{t,h})(1 - \pi_{t,w})]$$

Where  $B$  represents  $\frac{1.05}{1.07}$ , and  $\pi_{t,h}$  and  $\pi_{t,w}$  represent the probabilities that the man or his wife survive up through time  $t$ .



All respondents were asked if they would be eligible for pensions benefits from their present or previous jobs. Those who identified themselves as eligible for a pension were also asked at what age they would be eligible for full retirement benefits. No information was obtained on the kind of pension plan that individuals were eligible for but tabulations from the BLS's Level of Benefits Survey show that over 80% of those eligible for pensions are subject to some kind of defined benefit plan. Work by Lazear and Kotlikoff and Wise suggest that a salient feature of such plans are that increments to pension wealth decrease at early and normal retirement ages. This feature of such plans can be captured using the formula:

$$P = (1 - r n) b s e$$

where  $e$  represents the final salary,  $s$  years of service,  $b$  benefits as a fraction of final salary per year of service,  $r$  the reduction rate for early retirement, and  $n$  the number of years prior to normal retirement age that a man retires.  $e$ ,  $s$  and  $n$  are derived from survey responses assuming that a man worked through 1968.  $b$  and  $r$  were imputed using information on the industry of employment and tables 4.5.31 and 4.8.9 In Kotlikoff and Smith (1983). The present discounted value of pension benefits is then calculated assuming that benefits would remain constant in nominal terms from the date of retirement using a 7% discount rate.

To calculate total compensation for continuing to work, if a man were to continue working through 1969 he would continue to earn what he did in 1968. For those in the work force as of 1969 this number represents their actual 1968 earnings. For men who were out of the labor force as of 1969 this represents earnings during the last year for which they worked. Using the additional year of earnings I redid the present value calculation. Compensation is the difference between the present discounted value of earnings, social security benefits and pensions calculated assuming first that the man works through 1968 and then through 1969.

## References

- Aigner, D. C. Hsiao, A. Kapteyn and T. Wansbeek, (1984) "Latent Variable Models in Econometrics", ed. Griliches and Intriligator, *Handbook of Econometrics*, Amsterdam: North-Holland Publishing Co.
- Amemiya, T., (1976) "The Maximum Likelihood, the Minimum Chi-square, and the Non-linear Weighted Least Squares Estimator in the General Qualitative Response Model," *Journal of the American Statistical Association*, 71, 347-351.
- Amemiya, T., (1978) "The Estimation of a Simultaneous Equation Generalized Probit Model," *Econometrica*, 8(1), 13-21.
- Anderson, Kathryn H., and Richard V. Burkhauser, (1984) "The Importance of the Measure of Health in Empirical Estimates of the Labor Supply of Older Men," *Economic Letters*, 16, 375-380 .
- Anderson, Kathryn H., and Richard V. Burkhauser, (1985) "The Retirement-Health Nexus: A New Measure of an Old Puzzle," *Journal of Human Resources*, Summer, 20 (3), 315-330 .
- Andrisani, Paul J., (1977) "Health Limitations and Labor Market Experience During the Pre-retirement Years," in Seymour L. Wolfbein (ed.), *Men in Their Pre-retirement Years* (pp. 9-38). Philadelphia: Temple University.
- Avery, R. B., and V. J. Hotz, (1981) "Estimation of Multiple Indicator Multiple Cause (MIMIC) Models with Dichotomous Indicators." Carnegie- Mellon University, Working Paper No. 60-80-81.
- Baker, Diane, et al., (1982) "Mortality and Early Retirement," *Social Security Bulletin*, December, 45(12), 3-10 .
- Bartel, Ann, and Paul Taubman, (1979) "Health and Labor Market Success: The Role of Various Diseases," *Review of Economics and Statistics*, February, 61(1), 1-9.
- Bazzoli, Gloria J., (1985) "The Early Retirement Decision: New Empirical Evidence on the Influence of Health," *Journal of Human Resources*, Spring, 20 (2), 214-234.
- Berkowitz, Monroe, (1979) "The Search for a Health Variable," in *Proceedings of the Workshop on Policy Analysis with Social Security Research Files*. Washington, D.C.: U.S. Social Security Administration.
- Berkowitz, Monroe, and William G. Johnson, (1974) "Health and Labor Force Participation," *Journal of Human Resources*, 9, 115-128.
- Boskin, Michael J., (1977) "Social Security and Retirement Decisions," *Economic Inquiry*, January, 15, 1-15.
- Boskin, Michael J., and M. D. Hurd, (1978) "The Effect of Social Security on Early Retirement," *Journal of Public Economics*, December, 10, 361-377.
- Burkhauser, Richard V., (1979) "The Pension Acceptance Decision of Older Workers," *Journal of Human Resources*, Winter, 14, 63-75.

- Burkhauser, Richard V., (1980) "The Early Acceptance of Social Security: An Asset Maximization Approach," in *Industrial and Labor Relations Review*. Vol. 33, pp. 484-92.
- Burkhauser, Richard V., and Joseph F. Quinn, (1983) "Is Mandatory Retirement Overrated? Evidence from the 1970's," *Journal of Human Resources*, Summer, 18(3).
- Burtless, Gary, (1986) "Social Security, Unanticipated Benefit Increases, and the Timing of Retirement," in *Review of Economic Studies*, 53 (October), pp. 781-805.
- Burtless, Gary, and Robert A. Moffitt, (1984) "The Effect of Social Security Benefits on the Labor Supply of the Aged," in Henry J. Aaron and Gary Burtless (eds.), *Retirement and Economic Behavior*. Washington, DC: Brookings Institution.
- Burtless, Gary, and Robert A. Moffitt, (1985) "The Joint Choice of Retirement Age and Post-retirement Hours of Work," *Journal of Labor Economics*, April, 3(2), 209-236.
- Butler, J.S., Richard Burkhauser, Jean M. Mitchell and Theodore P. Pincus. (1987) "Measurement Error in Self-Reported Health Variables", *Review of Economics and Statistics*, 69:644-650.
- Butler, Robert N., (1983) "Current Data Inconclusive about Aged's Health and Work," *Aging and Work*, 6(3).
- Campbell, C. D., and R. G. Campbell, (1976) "Conflicting Views on the Effect of Old-age and Survivors Insurance on Retirement," *Economic Inquiry*, 17, 369- 388.
- Chamberlin, Gary and Zvi Griliches (1975) "Unobservables with a Variance Components Structure: Ability, Schooling and the Economic Success of Brothers", *International Economic Review*. 16:422-449.
- Chamberlin, Gary and Zvi Griliches (1977) "More on Brothers", Chp. 4 in Paul Taubman, ed., *Kinometrics: the Determinants of Socio-Economic Success Within and Between Families*, Amsterdam: North Holland Publishing Co.
- Chirikos, Thomas N., and Gilbert Nestel, (1981) "Impairment and Labor Market Outcomes: A Cross-sectional and Longitudinal Analysis," in Herbert S. Parnes (ed.), *Work and Retirement: A Longitudinal Study of Men* (pp. 93-131). Cambridge, MA: MIT Press.
- Chirikos, Thomas N., and Gilbert Nestel, (1982) "Sex and Race Differentials in the Economic Consequences of Poor Health." Research Report. Columbus: Center for Human Resource Research, Ohio State University.
- Chirikos, Thomas N., and Gilbert Nestel, (1983) "Economic Aspects of Self-reported Work Disability." Discussion paper. Columbus: Center for Human Resource Research, Ohio State University, January.
- Daymont, T. N., and Paul J. Andrisani, (1983a) "The Health and Economic Status of Very Early Retirees," *Aging and Work*, 6(2), 117-135.
- Diamond, P. A., and J. H. Hausman, (1984) "The Retirement and Unemployment Behavior of Older Men," in H. Aaron and G. Burtless (eds.), *Retirement and Economic Behavior*. Washington, DC: Brookings Institution.
- Ferguson, T.S., (1953) "A Method of Generating Best Asymptotically Normal Estimates with an Application to the Estimation of Bacterial Densities," *Annals of Mathematical Statistics* 29.

- Ferraro, Kenneth F., (198) "Self-ratings of Health among the Old and Old-old," *Journal of Health and Social Behavior*, December, 21, 377-383.
- Fields, Gary S., and Olivia S. Mitchell, (1984d) *Retirement Pensions and Social Security*. Cambridge, MA: MIT Press.
- Finnenbaum, G. G., and George Maddox, (1977) "Assessing the Functional Status of LRHS Participants." Center for Study of Aging and Human Development, Duke University, Technical Report No. 2, September.
- Fox, Alan (1976). *Alternative Measures of Earnings Replacement for Social Security Benefits, from Reaching Retirement Age*, Research Report no. 47, Office of Research and Statistics, Social Security Administration, Government Printing Office. (Appears as part of Appendix F, Methodology for Chp.14.)
- Garber, Steven and Steven Klepper, (1980) "Extending the Classical Normal Errors-in-Variables Model," *Econometrica* 48(6), September, 1541-1545.
- Garrad, J., (1974) "Impairment and Disability: Their Measurement, Prevalence and Psychological Cost," in Dennis Lees and Stella Shaw (eds.), *Impairment, Disability, and Handicap*. London: Heineman.
- Geraci, Vincent J., (1976) "Identification of Simultaneous Equation Models with Measurement Error," *Journal of Econometrics*, 4, 263-283.
- Gordon, Roger H., and Allan S. Blinder, (1987) "Market Wages, Reservation Wages, and Retirement Decisions," *Journal of Public Economics*, October, 14.
- Goldberger, A. (1972a), "Structural Equations Methods in the Social Sciences", *Econometrica*, 40, 979-1001.
- Goldberger, A. (1972b), "Maximum Likelihood of Regressions Containing Unobservable Independent Variables", *International Economic Review*, 13:1-15. *Econometrica*, 40, 979-1001.
- Griliches, Zvi, (1974), "Errors in Variables and Other Unobservables," *Econometrica*, 42.
- Gustman, Alan L., and Thomas Steinmeier, (1986) "A Structural Retirement Model," *Econometrica*, May, 54(3).
- Hanoch, Giora, and Marjorie Honig, (1983) "Retirement, Wages, and Labor Supply of the Elderly," *Journal of Labor Economics*, April, 1, 131-151.
- Hausman, Jerry A., and David A. Wise, (1985) "Social Security Health Status and Retirement," in D. Wise (ed.), *Pensions, Labor and Individual Choice*. Chicago: University of Chicago Press.
- Haveman, Robert, Barbara Wolfe and Fung Mey Huang, (1989) "Disability Status as an Unobservable: Estimates From a Structural Model," NBER working paper # 2831.
- Hogarth, Jeanne M., (1988) "Accepting an Early Retirement Bonus: An Empirical Study," in *Journal of Human Resources* XXIII(1), pp.21-33.
- Hurd, Michael D., and Michael J. Boskin, (1981) "The Effect of Social Security Retirement in the

- Early 1970's," in *Quarterly Journal of Economics*, November, pp. 767-790.
- Johnson, William G., (1977) "The Effect of Disability on Labor Supply: Comment," *Industrial Labor Relations Review*, April, 30, 380-381.
- Joreskog, N., and A. Goldberger, (1975) "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable Model," *Journal of the American Statistical Association*, 56, 235-245.
- Kapteyn, Arie and Tom Wansbeek, (1983) "Identification in the Linear Errors in Variables Model," *Econometrica* 51, 1847-1849.
- Klepper, Steven, and Edward E. Leamer, (1984) "Consistent Sets for Regression with Errors in All Variables," *Econometrica*, 52, 162-183.
- Klepper, Steven, and David C. Stapleton, (1985) "Consistent Sets of Estimates for Restricted Regression with Errors in All Variables," Carnegie-Mellon University, Pittsburgh, Pennsylvania.
- Kotlikoff, L. J., and D. E. Smith (1983) *Pensions in the American Economy*, Chicago, The University of Chicago Press.
- Kotlikoff, L. J., and D. A. Wise, (1985) "Labor Compensation and the Structure of Private Pension Plans: Evidence of Contractual versus Spot Labor Markets," in D. Wise (ed.), *Pensions, Labor and Individual Choice*. Chicago: University of Chicago Press.
- Krasker, William S. and John W. Pratt, (1986) "Bounding the Effects of Proxy Variables on Regression Coefficients," *Econometrica*, 54(3), May, 641-655.
- Lambrinos, James, (1981) "Health: A Source of Bias in Labor Supply Models," *Review of Economics and Statistics*, May, 20 6-212.
- LaRue, Asenath, Lew Bank, Lissy Jarvic, and Monte Hewtland, (1979) "Health in Old Age: How Physicians' Ratings and Self-ratings Compare?," *Journal of Gerontology*, September, 34(5).
- Lazear, E., (1985) "Pensions as Severance Pay", in Z. Bodie and J.B. Shoven (ed.), *Financial Aspects of the United States Pension System*. Chicago: University of Chicago Press.
- Lee, L. F., (1979) "Identification and Estimation in Binary Choice Models with Limited (Censored) Dependent Variables," *Econometrica*, 47, 977-996.
- Lee, L. F., (1982a) "Health and Wage: A Simultaneous Equation Model with Multiple Discrete Indicators," *International Economic Review*, 3(1), 199-221.
- Lee, L. F., (1982b) "Simultaneous Equations Models with Discrete and Censored Variables," in C. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data: With Econometric Applications*. Cambridge, MA: MIT Press.
- Luft, Harold, (1975) "The Impact of Poor Health on Earnings," *Review of Economics and Statistics*, February, 57(1), 43-57.
- MaCurdy, Thomas, (1985) "Interpreting Empirical Models of Labor Supply in an Intertemporal Framework with Uncertainty", in J. Heckman and B. Singer, *Longitudinal Analysis of Labor Market Data*, Cambridge U. Press.

- Maddox, G., and E. Douglas, (1973) "Self-assessment of Health: A Longitudinal Study of Elderly Subjects," *Journal of Health and Social Behavior*, 14(1), 87-93.
- Mossey, Jana M., and Evelyn Shapiro, (1982) "Self-rated Health: A Predictor of Mortality among the Elderly," *American Journal of Public Health*, August, 72(8), 800-808.
- Muthen, B., (1983) "Latent Variable Structural Equation Modeling with Categorical Data," *Journal of Econometrics*, 22:43-65.
- Muthen, B., (1979) "A Structural Probit Model with Latent Variables" *Journal of the American Statistical Association*, 74:807-811.
- Muthen, B., (1979) "Contributions to Factor Analysis of Dichotomous Variables", *Psychometrika*, 43:551-560.
- Myers, Robert J., (1982) "Why Do People Retire from Work Early?," *Aging and Work*, 5(2), 83-91.
- Myers, Robert J., (1983) "Further about Controversy on Early Retirement Study," *Aging and Work*, 6(2), 10 5-10 9.
- Nagi, Saad Z., (1969a) "Congruency in Medical and Self-assessment of Disability," *Industrial Medicine*, 27-36.
- Newey, Whitney K., (1987), "Efficient Estimation of Limited Dependent Models with Endogenous Explanatory Variables", *Journal of Econometrics*, 36:231-250, November.
- Olsson, Ulf, (1979), "Maximum Likelihood Estimation of the Polychoric Correlation Coefficient", *Psychometrika*, Vol. 44, no. 4, pp. 443-460, December.
- Parsons, Donald O., (1977) "Health, Family Structure, and Labor Supply," *American Economic Review*, September, 67, 70 3-712.
- Parsons, Donald O., (1980a) "The Decline of Male Labor Force Participation," *Journal of Political Economy*, February, 88, 117-134.
- Parsons, Donald O., (1980b) "Racial Trends in Male Labor Force Participation," *American Economic Review*, December, 70, 911-920 .
- Parsons, Donald O., (1982a) "The Male Labor Force Participation Decision: Health, Reported Health, and Economic Incentives," *Economica*, February, 49, 81-91.
- Parsons, Donald O., (1984a) "Disability Insurance and Male Labor Force Participation: A Response to Haveman and Wolfe," *Journal of Political Economy*, June, 92(3), 542-549.
- Pearson, K., and Pearson E. S., (1922) "On Polychoric Coefficients of Correlation", *Biometrika*, 14:17-156.
- Quinn, Joseph F., (1977) "Microeconomic Determinants of Early Retirement: A Cross-sectional View of White Married Men," *Journal of Human Resources*, Summer, 12(3), 329-346.
- Reno, Virginia, (1971) "Why Men Stop Working at or before Age Sixty- five," *Social Security Bulletin*, June, 34, 3-16.
- Sammartino, Frank J., (1982) "The Timing of Social Security Acceptance by Older Men: Examining

the Financial Incentives." U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, Technical Analysis Paper No. 24, January.

Schwab, Karen, (1974) "Early Labor Force Withdrawal of Men: Participants and Non-participants Aged 58-63," *Social Security Bulletin*, August, 37, 24-38.

Sherman, Sally R., (1985) "Reported Reasons Retired Workers Left Their Last Job: Findings from the New Beneficiary Survey," *Social Security Bulletin*, March, 48(3), 22-30 .

Sickles, Robin C. and Paul Taubman, (1986) "An Analysis of the Health and Retirement Status of the Elderly", *Econometrica*54:1339-1356.

Stern, Stephen, (1988) "Measuring the Effect of Disability on Labor Force Participation" Manuscript, the University of Virginia.

U.S. Department of Health and Human Services, Social Security Administration, (1982) *Social Security Bulletin, Annual Statistical Supplement*.

U.S. Department of Health and Human Services, Social Security Administration, (1984) *Vital and Health Statistics: Current Estimates from the National Health Interview Survey, United States*.

Table 1: Descriptive Statistics

variable	mean	min	max	std
Ln(wage)	8.80	4.63	11.59	.78
Ln(income)	12.03	8.79	13.76	.77
Age	60.40	58.00	63.00	1.71
Education	9.78	0.0	18.00	3.57
Nonwhite	.09			
Married	.88			
Pension	.42			
Good health	.35			
Poor health	.20			
Health limits work	.35			
Labor force participant	.82			

#### Variable Definitions

ln(wage)	log of the total compensation for working an extra year (see the text for details)
ln(inc)	log of discounted lifetime earnings (see the text for details)
age	age as of 1969
education	educational attainment in 1969
married	1 if married spouse present in 1969, 0 otherwise
nonwhite	1 if nonwhite, 0 otherwise
pension	1 if eligible for a pension in 1969; 0 otherwise
health limits work	1 if respondent, as of 1969, reports that health limits his ability to work, 0 otherwise.
good health	1 if respondent, as of 1969, reports that his health is better than that of other men his age, 0 otherwise.
poor health	1 if respondent, as of 1969, reports that his health is worse than that of other men his age, 0 otherwise.
labor force participant	1 if respondent was in the labor force as of 1969, 0 otherwise.



**Table 2: Percent with Health Limitations in 1969 and  
Percent out of the Labor Force in 1969 by Year of Death**

Date of Death	percent of the sample	% with health limitations '69	% in poor health '69	% out of the labor force '69
1969	1.2	81.9	58.3	52.8
1970	1.9	70.4	53.0	46.1
1971	1.7	68.3	46.5	45.5
1972	2.9	52.6	37.1	33.1
1973	2.9	51.7	27.3	27.3
1974	3.2	46.4	30.4	28.9
1975	2.9	51.7	33.7	26.2
1976	3.1	49.5	33.0	25.0
1977	3.2	48.4	32.3	26.0
1978	3.5	37.9	26.1	19.0
1979	1.1	40.0	26.2	20.0
alive in 1979	72.3	28.7	15.4	14.1
total	100.0	32.6	20.7	18.4

Table 3: Reduced Forms

Explanatory Variable	Equation			
	Labor Force	Health Limits	Poor Health	Date of Death
Log of Wage	.193 (.038)	.294 (.031)	.235 (.036)	.101 (.032)
Log of Income	-.061 (.038)	.073 (.033)	.043 (.037)	-.039 (.032)
Age	-.154 (.012)	-.063 (.010)	-.015 (.011)	-.062 (.010)
Education	.047 (.006)	.043 (.005)	.051 (.005)	.018 (.005)
Non white	-.034 (.069)	.169 (.064)	-.013 (.066)	.009 (.062)
Married	.447 (.055)	.038 (.054)	.049 (.058)	.172 (.053)

Cross Equation Correlations

Labor Force	1.00			
Health Limits	.707	1.00		
Poor Health	.717	.870	1.00	
Date of Death	.297	.337	.354	1.00
	Labor Force	Health Limits	Poor Health	Date of Death

**Table 4: Labor Force Non-Participation**

explanatory variable	labor force participation			
	1	2	3	4
Log of Wage	.193 (.037)	.075 (.041)	.112 (.041)	.184 (.037)
Log Income	-.061 (.038)	-.134 (.042)	-.115 (.042)	-.060 (.038)
Age	-.154 (.012)	-.157 (.013)	-.187 (.013)	-.149 (.012)
Education	.047 (.006)	.031 (.007)	.023 (.007)	.047 (.006)
NonWhite	-.034 (.068)	-.121 (.076)	-.023 (.076)	-.039 (.069)
Married	.447 (.056)	.541 (.061)	.541 (.060)	.441 (.057)
Health Limits		-1.366 (.046)		
Poor Health			-1.449 (.047)	
Mortality				
1974-1979				-.261 (.055)
1973				-.307 (.108)
1972				-.520 (.104)
1971				-.924 (.131)
1970				-.954 (.121)
1969				-1.017 (.155)
Log Lik:	-2654.0	-2153.4	-2147.9	-2567.9
N. obs	6022	6022	6022	6022

Table 5: Parameter Estimates Based on Reduced Forms

Control for Health	None	Date of Death	Health Limits	Bad Health	Limits Instrumented	Health Instrumented
Parameter Restrictions	$\lambda_1 \equiv 0$	$\lambda_3 \equiv 1$ $\sigma_{\epsilon_1}^2 \equiv 0$	$\beta_2 \equiv 0, \lambda_2 \equiv 0$ $\sigma_{\epsilon_2}^2 \equiv 0$		$\beta_1 \equiv 0$	
$\beta_1$	.192 (.037)	.163 (.036)	-.014 (.033)	.025 (.033)	-.065 (.040)	-.004 (.036)
$\lambda_1$	-	.296 (.022)	.711 (.016)	.719 (.016)	.876 (.147)	.762 (.128)

Table 6: Estimates Based on Various Assumptions about  $\lambda_2$

*Health Limits*

$\lambda_2$	$\beta_1$	$\beta_2$	$\lambda_1$	$\lambda_3\tau$	$\rho$
.334	.163	.257	.296	1.000	.675
.4	.150	.242	.354	.835	.659
.5	.125	.215	.443	.669	.624
.579	.103	.190	.513	.579	.584
.6	.096	.182	.532	.559	.571
.7	.062	.144	.620	.481	.486
.8	.023	.099	.709	.422	.330
.852	.000	.024	.755	.397	.187
.9	-.022	.049	.797	.377	-.033
.953	-.029	.128	.810	.371	-1.000
.989	-.065	.000	.876	.344	-1.786

*Bad Health*

$\lambda_2$	$\beta_1$	$\beta_2$	$\lambda_1$	$\lambda_3\tau$	$\rho$
.352	.163	.198	.296	1.000	.685
.4	.154	.188	.336	.881	.675
.5	.132	.162	.420	.706	.645
.595	.108	.132	.499	.595	.602
.6	.106	.131	.503	.590	.600
.7	.075	.094	.587	.507	.529
.8	.040	.052	.671	.445	.407
.9	.000	.004	.755	.397	.140
.908	-.004	.000	.762	.393	.103