

**Effects of Health Status on the Decision to Retire
with Corrections for Measurement Error**

By

Kyung-Rae Hyun

**A dissertation submitted to the Graduate Faculty in Economics in partial
fulfillment of the requirements for the degree of Doctor of Philosophy,**

The City University of New York

2008

©2008

Kyung-Rae Hyun

All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Economics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Michael Grossman

Date

Chair of Examining Committee

Thom Thurston

Date

Executive Officer

Professor Michael Grossman

Professor Henry Saffer

Professor Thom Thurston

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Effect of Health Status on the Decision to Retire

with Corrections for Measurement Error

By

Kyung-Rae Hyun

Adviser: Professor Michael Grossman

Many studies on the relationship between labor supply and health status generally do not illustrate a unanimous agreement on the usage of subjective and objective health due to the difficulty in employing the right measures to obtain the precise and up-to-date information on the health status of subjects in question. Thereby, to resolve this quandary, we propose two models; the first model addresses how a researcher could obtain more appropriate estimate in health measure with measurement error; the other model examines whether probabilities questions within and across waves could potentially result in another measurement error.

In the first model allowing to justification bias, we found the health and financial compensation' lower and upper bounds for labor force exit and in views of bounds concept, health status plays a larger role in labor force exit than financial compensation.

In the second model, we found that measurement error from survey effect has an impact on the survival probability along with the measurement errors from an inherent errors-in-variable bias by using two methods: the former employing the averaged survival probabilities and the latter incorporating the instrumental variables.

Acknowledgements

I would like to thank my dissertation advisor, Professor Michael Grossman. He was always available to me to figure out my problems concerning this dissertation. I have been helped and encouraged by many members of the Graduate Center's Economics Department. Professor Thom Thurston and Professor Henry Saffer were very helpful to me.

On a personal note, I gratefully acknowledge my father Jong-Yun Hyun and my mother Gwang-Ok Kim's indispensable support. My two brothers, Kyung-Hwan Hyun and Kyung-Suk Hyun and my two sisters, Mi-Ja Hyun and Mi-Suk Hyun provided invaluable inspiration. I dedicated this work to my family.

Table of Contents

1. Introduction	1
2. Econometric Model	7
3. Data	20
4. Results	24
5. Conclusion	31
6. Appendix 1	43
7. Appendix 2	47
8. Appendix 3	49
9. Appendix 4	51
10. Bibliography	52

List of Tables

Table 1. Survey Sequence Ordering of Key Questions in HRS	33
Table 2. Sample Statistics	34
Table 3. Regression using Reduced Form	35
Table 4. Correlations of Reduced Form Regression	36
Table 5. Labor Force Exit using Proxy Variables of Health	37
Table 6. Labor Force Exit in Wave (1) based on Reduced Form Model	38
Table 7. Labor Force Exit in Wave (3) based on Reduced Form Model	39
Table 8. Correlation Matrices for Probability Variables	40
Table 9. OLS, Fixed-Effect, Avg and IV of Labor Force Participation in Wave (1)	41
Table 10. OLS, Fixed-Effect, Avg and IV of Labor Force Participation in Wave (3)	42
Appendix 4. Description of Variables	51

1. Introduction

Many studies for the relationship between labor supply and health status do not have a unanimous agreement on the usage of subjective health because of the difficulty of correct measures of health status. Considering the tripod relationship among health, working hours, and leisure time taking up household production function, the effect of labor supply is caused by either improvement or exacerbation in health status¹. Financial circumstances covering post-retirement may also wield significance influence on labor supply as well. It is doubtful whether proxy variable of subjective health gives us a reasonable result in consideration of the correlation between health status and financial compensation. The decision of labor force participation for older persons could be more affected by the survival probability recognized during the survey period than the subjective or objective health with holding other factors constant. When the older age group is likely to leave the labor market due to their health condition it is associated not only with financial factors but also with observed health problems and unobserved subjective excuse for non-work. Here it is also a point in question whether survival probability could be affected by measurement error from survey sequence ordering's change.

The first issue is measurement error from each health status such as subjective or objective health and survival probability. There are questions whether subjective health measure has a more accurate power in the relationship between labor force exit and health status. It has been hypothesized that those who answer to leave the labor market due to bad health condition do not provide an exact reason to

¹ An improvement by lowering stress or investing health care could make leisure and household time increase while an exacerbation by feeling a tediousness after retirement could make leisure and household time decrease.

leave the labor market. It's justified as a simple acceptable excuse for non-work. McGarry(2002) claimed that this phenomenon has been termed as "justification bias"². The justification bias occurs to endogenous problem about health measure. Generally respondents are also asked about subjective health status but there has been controversy as to whether this measurement can be comparable across persons. There may be unobserved differences across individuals that are correlated with health status and labor market behavior³. Non-comparable judgment across individuals refers to measurement error that leads to underestimate the effect of health on labor force exit, while the endogenous problem of subjective health measure leads to overestimate the effect. Those biases of health measures could produce negative or positive effect on the financial factors since they are correlated with subjective health measure.

In the past, many economists have traditionally supported for the use of objective health status, such as health conditions, activity limitation, or mortality⁴. Although they have employed objective health indicators in lieu of their more subjective equivalents, they have yet to provide concrete evidence as to why employing the objective health measure would have more accurate assessment on the labor force exit than the subjective health measure. Because they ignored errors-in-variable bias from objective health as insignificant matter this phenomenon occurs.

² McGarry(2002) said that although the justification bias is the most frequently cited problem, there are another problems. Suppose that health measure is taken after retirement. This measure may be poor proxies for health after retirement if health is affected by employment, or if bad health is the result of unexpected illness. Nor is it clear that health measure prior to retirement is the appropriate measure.

³ Bound (1991) said that there is no reason to expect that subjective health judgment will be comparable across individuals and respondents may not be independent of the labor market behavior.

⁴ Anderson and Burkhauser (1984, 1985) used mortality information and Bartlet and Taubman (1979) used the specific health condition and Chirikos and Nestel (1981, 1983) used activity limitation.

Because economists failed at demonstrating an unambiguous correlation between objective health and all involved factors the current methodology suffers from its inherent inaccuracy, and consequently, a certain amount of congenital error in measurement of its variables. Comparatively speaking, there could be only one inherent bias underestimating objective health in the regression and as long as the correlation between objective health proxy and real health is not perfect, this bias will affect any correlated coefficients such as the financial factors. However for subjective health measures, since biases are able to work in the opposite direction among each other, there is a possibility that these biases can be potentially canceled out in some way.

Alternatively it seems possible that men rationalize retirement in view of the expected survival probability⁵. Unlike the subjective health's measurement error, the expected survival probability might have only an inherent errors-in-variable bias which underestimates the effect of health on labor force exit⁶. Hurd and McGarry (1995) claimed that survival probability varies in a reasonable way with diseases problems relating to longevity, socio-economic status and financial factors therefore, it is a good variable of choice that relates to retirement decision, which potentially allows us to control the element of heterogeneity in a given observation.

Here one may question how a researcher can procure more accurate estimates in health measure associated with measurement error. Some economic studies have

⁵ Hurd and McGarry(1995) suggested that the measures of expected probabilities in the Health and Retirement Study(HRS) are good approximations to population probabilities. However, even if these probabilities are internally consistent and covary with other variables, it is not clear in view of the rapid change in mortality rates how much accurate they should be because researcher does not know how respondents make their views about mortality risk.

⁶ Survival probability is internally consistent. However to the extent that it is not, the researcher should construct model that causes the inconsistency. This process surely includes measurement error from errors-in-variable bias in the covariation with other variables.

found that when objective health measure is used⁷health has a smaller effect on labor force participation than financial factors. However, they have ignored measurement error in objective health by assuming that an inherent errors-in-variables bias of objective health proxy is less than measurement error resulting from the endogenous problem by the use of subjective health proxy. Other economic studies have argued that subjective health was the reliable health indicator by empathizing the defect of objective health measure. However, they have also ignored the fact that the correlation between retirement decision and health proxies could produce another unobserved measurement error⁸. Therefore, our goal is to realize how an analyst could reliably compromise the immanent oversight of the instruments of all health measures.

Another issue is whether survey sequence ordering's change for a similar question within or across waves has an effect on labor force decision. The fact that all respondents might not have a logical and reasonable intuitiveness during survey period within or across waves is likely to induce unobserved measurement error with regards to survey sequence ordering. If there were a perfect correlation between health and labor force decision an illogical survey response would overestimate (or underestimate) our coefficients by a systematical unobserved bias. In Health and Retirement Study (HRS) survey, there are two key questions related to survey sequence ordering within or across waves "Thinking about work generally and not just your present job, what do you think are the chances that you will be working full-time after you reach age 62 or 65?" "What do you think are the chances that you

⁷ See Anderson and Burkhauser (1984, 1985)

⁸ Those out of work or non-worker are more likely to report bad health as the excuse for non-work than current workers. This information gives rise to measurement error for misleading information.

will live to be 75 or 85?”⁹. I denote the first query as *pw62* or *pw65* (probability work) and the second as *pl75* or *pl85* (survival probability).

The Table 1 summarizes survey’s key question ordering. In wave (1), *pw*’s questions were asked in employment section and then *pl*’s questions are followed in the cognition section prior to questions about financial factors. In wave (2), *pw* and *pl*’s questions were asked in the cognition section together but these questions are reported between financial factors questions. The survey sequence ordering in wave (3) has the same ordering as that in the wave (4). In wave (3) and (4), *pw* and *pl*’s questions were asked in the expectation section together such as wave (2) but were asked prior to questions about financial factors such as wave (1).

As long as respondents have interviewed continuously from the first survey period they did not recognized a discrepancy in the survey sequence ordering within or across waves. This fact could make a controversy about whether survey sequence ordering could induce to overestimate (or underestimate) the survival probability of labor force participation. If there were positive survey effect, our estimate of survival probability on probability work would be overestimated; otherwise, it would be underestimated. We could make an adjusted structural model to eliminate the survey effect. If this adjusted model proves that there exists survey effect, it will be another issue that researchers should not ignore along with measurement error.

We construct two statistical models in section 2. One is the labor force exit model with justification bias and the other is the labor force participation model with survey effect and then in section 3 and 4 we describe our data set from Health and

⁹ On average, the expected probabilities will correctly predict some of the variation in outcomes in the HRS. From the cross-section, the measures of expected probabilities in the HRS show a substantial contribution to decision making under uncertainty.

Retirement Study (HRS) and results of our estimation. Finally, a brief conclusion follows.

2. Econometric Model

1) Labor Force Exit Model with Justification Bias

The first model is based on Bound(1991)'s methodology. Suppose labor force exit model of older persons in which retirement decision relies on the financial compensation, w , and unobserved bad health status, h^* :

$$r = \alpha_1 h^* + \beta_1 w + \varepsilon_1 \quad (1)^{10}$$

We can expect β_1 to be negative but α_1 to be positive¹¹. With regard to the correlation between health and financial compensation we added the relationship between explanatory variables:

$$w = \alpha_2 h^* + \varepsilon_2 \quad (2)$$

where α_2 is ambiguous¹². We assume that h^* is very weakly correlated with w because labor market behaviors have impact on wage rate decisively.

Consider that the unobserved health, h^* , consists of two factors. One is health problems, l , which people thought, could threaten their future life and the other is diseases problems, d , which are originally fatal components for older persons:

$$h^* = l + d$$

where l and d are assumed to be uncorrelated with each other.

Now suppose subjective health status, h^s , depends on the financial compensation and unobserved bad health status:

¹⁰ In our results, we include demographic variables and retirement plans as explanatory variables.

¹¹ Implicitly, as long as larger value of h^* is associated with bad health condition we could expect α_1 to be positive.

¹² Generally we can say that α_2 is negative. If we assumed that α_1 is positive, since bad health causes labor supply to decrease, this causality relationship also induces to decrease wage rate. Therefore we can expect α_2 to be negative.

$$h^s = \alpha_3 h^* + \beta_3 w + \varepsilon_3 \quad (3)$$

As long as h^s approaches to worst health status a respondent is likely to report himself reducing labor supply. The simultaneous relation between equations (1) and (3) allows “justification bias” that those who report retirement because of failing health are providing a socially acceptable excuse for non-worker rather than an accurate description of retirement¹³. We can expect α_3 to be positive but β_3 to be ambiguous¹⁴. With allowance to justification bias for retirement, ε_1 and ε_3 will be positively correlated. We assume that h^* is orthogonal to ε_1 and ε_3 .

Since h^s is endogenous, we can regress reduced form with h^s as dependent variable and then derive the parameters from the estimated reduced form. We use h^s as proxy for h^* . The OLS regression of the following two equations is the same implication as the regression of reduced form for r and h^s :

$$r = \pi_1 w + u_1 \quad (1)'$$

$$h^s = \pi_3 w + u_3 \quad (3)'$$

where $\pi_1 = \alpha_1 \gamma + \beta_1$, $\pi_3 = \alpha_3 \gamma + \beta_3$ and $u_1 = \alpha_1 v + \varepsilon_1$, $u_3 = \alpha_3 v + \varepsilon_3$.

In the classical measurement error model using the reduced form equations, the OLS estimates of α_1 and β_1 are¹⁵:

$$\text{plim } a_1 = \text{Cov}(u_1, u_3) / \text{Var}(u_3) \text{ and } \text{plim } B_1 = \text{plim } \pi_1 - \text{plim } \pi_3 \text{plim } a_1$$

By normalization restriction that α_3 is equal to 1 and $\gamma \equiv \sigma_{h^*,w} / \sigma_w^2$

¹³ See Kathleen McGarry (2002)

¹⁴ Suppose that we expect β_3 to be negative, that is, those who are compensated more at the job are likely to report themselves having better health status. There is higher possibility that those who have a higher wage rate and report a better health are not likely to leave employment.

¹⁵ Our derivation is based on Bound(1991)'s methodology. See Bound(1991) and Newey (1987) and Appendix 1.

$$plim a_1 = (\alpha_1 \sigma_v^2 + \sigma_{\varepsilon_1} \sigma_{\varepsilon_3} \rho) / (\sigma_v^2 + \sigma_{\varepsilon_3}^2)$$

$$plim B_1 = \beta_1 + (\alpha_1 - plim a_1) (\sigma_{h^*,w} / \sigma_w^2) - plim a_1 \beta_3$$

where σ_v^2 is related to the correlation between h^* and w , and ρ the correlation between ε_1 and ε_3 . The correlation between ε_1 and ε_3 induces bias from endogenous problem in subjective health while variance in ε_3 relating to the reliance of h^s on h^* induces an inherent errors-in-variable bias. Bound said that the bias from endogenous problem is called to simultaneity bias.

As long as the simultaneity bias will give rise to upward bias and the errors-in-variable bias do to downward bias it could be possible to cancel out both biases in the scale respect. However if h^s approach to the true h^* , the bias of α_1 would be greater and substantial since $\sigma_{\varepsilon_3}^2$ disappears. The existence of simultaneity bias has a similar implication of justification bias that prevents providing an accurate information about retirement decision while the existence of errors-in-variable bias has an opposite implication of justification bias in terms that the more reliable is subjective health (the smaller is $\sigma_{\varepsilon_3}^2$), the greater will be the bias of α_1 . The bias of β_1 relies on α_1 's measurement error and β_3 's sign. Even though we estimate a perfect α_1 , when α_1 has positive and β_3 has negative sign β_1 would be overestimated¹⁶. If we assume that h^* is very weakly correlated with w β_3 's sign is determined by estimate of π_3 's sign in reduced form because covariance $\sigma_{h^*,w}$ approaches to zero and then γ also approaches to zero¹⁷.

¹⁶ Usually when we construct different model with different explanatory variables we should not insist β_3 to be negative sign. Therefore the sign of β_3 depends on how analyst construct his(her) structural model.

¹⁷ See Appendix 1.

Now suppose the objective health, h^o , depends on the health problems, l :

$$h^o = \alpha_4 l + \varepsilon_4 \quad (4)$$

where ε_4 is assumed to be uncorrelated with either ε_1 , ε_2 , or ε_3 .

If r and h^s have dichotomous variables and the objective health problem, h^o , and survival probability, pl , which we will discuss in the later part are continuous variables it is not possible to use h^o and pl as a proxy for h^* or as an instrumental variable for h^* . The approach to this problem is that we choose the estimation of the reduced form parameters. When we use h^o or pl as a proxy for h^* we can make the following reduced form equations:

$$h^o = \pi_4 w + u_4 \quad (4)'$$

$$1-pl = \pi_5 w + u_5 \quad (5)'$$

where $\pi_4 = \alpha_1 \gamma_1$, $\pi_5 = \alpha_5 \gamma_d + \beta_5$ and $u_4 = \alpha_4 \zeta + \varepsilon_4$, $u_5 = \alpha_5 \tau + \varepsilon_5$.

After regressing equation (1)' and (4)', we can derive the OLS estimates of α_1 and β_1 from the reduced form regression with π_1 , π_4 and u_1 , u_4 ¹⁸. Using h^o as proxy for h^* in the reduced form equations, the OLS estimates of α_1 and β_1 are:

$$plim a_1 = Cov(u_1, u_4) / Var(u_4) \text{ and } plim B_1 = plim \pi_1 - plim \pi_4 plim a_1$$

By normalization restriction that α_4 is equal to 1 and $\gamma_1 \equiv \sigma_{l,w} / \sigma_w^2$

$$plim a_1 = \alpha_1 \sigma_\zeta^2 / (\sigma_\zeta^2 + \sigma_{\varepsilon_4}^2)$$

$$plim B_1 = \beta_1 + \alpha_1 (\sigma_{h^*,w} / \sigma_w^2) - plim a_1 (\sigma_{l,w} / \sigma_w^2)$$

where σ_ζ^2 is related to the correlation between l and w .

¹⁸ See Appendix 1.

For α_l , only the variance of ε_4 , $\sigma_{\varepsilon_4}^2$ induces an inherent errors-in-variable bias that made α_l to be underestimated. For β_l , even though errors-in-variable bias makes α_l be underestimated, controlling h^o with only l could reserve an omitted variable bias on β_l . Because of errors-in-variable bias and omitted variable bias we could not make decision whether β_l is overestimated or underestimated.

Suppose survival probability is pl , hence $1-pl$ denotes mortal probability. To explain our model easily, we use $1-pl$ as another health measure because subjective and objective health denote bad health measures. Now suppose mortal probability, $1-pl$, depends on diseases problems, d , and the financial component, w :

$$1-pl = \alpha_5 d + \beta_5 w + \varepsilon_5 \quad (5)$$

where ε_5 is uncorrelated with ε_1 and ε_3 . We can expect α_5 to be positive and β_5 to be negative. It is the different point from equation (4) that $1-pl$ relies on diseases problems and financial compensation while h^o only relies on health problems. Likewise h^o 's case, after regressing equation (1)' and (5)', we can derive the OLS estimates of α_l and β_l from the reduced form regression with π_1 , π_5 and u_1 , u_5 . The OLS estimates of α_l and β_l are:

$$plim a_1 = Cov(u_1, u_5) / Var(u_5) \text{ and } plim B_1 = plim \pi_1 - plim \pi_5 plim a_1$$

By normalization restriction that α_5 is equal to 1 and $\gamma_d \equiv \sigma_{d,w} / \sigma_w^2$

$$plim a_1 = \alpha_1 \sigma_{\tau}^2 / (\sigma_{\tau}^2 + \sigma_{\varepsilon_5}^2)$$

$$plim B_1 = \beta_1 + \alpha_1 (\sigma_{h^*,w} / \sigma_w^2) - plim a_1 \{ (\sigma_{d,w} / \sigma_w^2) + \beta_5 \}$$

where σ_{τ}^2 is related to the correlation between d and w .

The existence of errors-in-variable bias induces α_l to be underestimated. However even though errors-in-variable bias makes α_l be underestimated, controlling $l-pl$ with only d and w could leave an omitted variable bias on β_l . Because of errors-in-variable bias and omitted variable bias we could not make decision whether β_l is overestimated or underestimated.

To estimate α_l and β_l appropriately and decrease various biases from measurement error we use the various proxies of h^* as instrumental variables. Since h^s is correlated with ε_l it's impossible to use h^s as instrument variable for h^o or $l-pl$. However, it's possible to use h^o or $l-pl$ to instrument for h^s because ε_4 and ε_5 are uncorrelated with ε_l . As we can see in the Appendix 1, we could obtain an identified coefficient of subjective health by 2SLS in which employing objective health problem proxy and mortal probability proxy as instrumental variables. The followings are the estimates of subjective health proxy (h^s) and financial compensation (w) after subjective health proxy is instrumented for objective health problem proxy or mortal probability proxy (h^o or $l-pl$).

After normalizing α_3 to equal to 1,

$$\mathbf{plim} a_{IV} = \alpha_l$$

$$\mathbf{plim} B_{IV} = \beta_l - \alpha_l \beta_3$$

By processing an instrument variables method, α_l can be identified but β_l is still overestimated (or underestimated) depending on the signs of α_l and β_3 . When we summarize the above results we can produce the following judgment conditions to find more appropriate estimates:

Condition (i) After using various proxies for the true health and comparing subjective health proxy with that in IV method by accepting objective health problem proxy and mortal probability proxy as valid instrumental variables, as long as we confirm that α_l of subjective health proxy is overestimated in OLS regression with positive sign of β_3 , that is $0 < \alpha_l < \text{plim } a_l$ with $\beta_3 > 0$, β_l might be underestimated in OLS regression. If β_3 has a negative sign, we will not confirm that β_l is either overestimated or underestimated. Therefore, when we face this first condition we can get more appropriate estimates by finding a smaller value of α_l than an upper bound of OLS regression and a larger value of β_l than a lower bound of OLS regression.

Condition (ii) After using various proxies for the true health and comparing subjective health proxy with that in IV method by accepting objective health problem proxy and mortal probability proxy as valid instrumental variables, as long as we confirm that α_l of subjective health proxy is underestimated in OLS regression with negative sign of β_3 , that is, $0 < \text{plim } a_l < \alpha_l$ with $\beta_3 < 0$, β_l might be overestimated. If β_3 has a positive sign, we will not confirm that β_l is either overestimated or underestimated in OLS regression. Therefore, when we face this second condition we can get more appropriate estimates by finding a larger value of α_l than a lower bound of OLS regression and a smaller value of β_l than an upper bound of OLS regression.

For subjective health, sometimes without further information the parameters of the structural model are not identified since subjective health is correlated with financial compensation and error term. Therefore, to identify subjective health we

need another proxy variable for h^* as much as possible. In other words, if we have additional variables that are redundant in the structural model and they are utilized as instrumental variables for h^s these additional variables help produce the consistent estimates¹⁹. If r and h^s are binary variables this case is different from the instrumental variables method using the various proxies of h^* as instrumental variables. Since a binary endogenous explanatory variable (h^s) exists in the discrete response models we have the four possible outcomes of (r, h^s) . We derive log-likelihood function for maximum likelihood analysis to find estimates of α_l , β_1 , and β_3 in Appendix 2.

2) Labor Force Participation Model with Survey Effect

The second model is based on Ashenfelter and Krueger(1994)'s methodology. We constructed the first model in which survival probability has measurement error from an inherent errors-in-variables bias that underestimates the coefficient. Multiple measures for survival probability and probability work within and across waves can assess survey effect relating to survey's sequence ordering.

Suppose $v_i^j = pl_i^j - pl_i$ ($i= 75,85$ and $j = \text{wave } 1,2$ or $\text{wave } 3,4$) where pl_i^j refers to observed measure of survival probability to live to be 75 (or 85) years old as reported at j th wave and pl_i is the true and unobserved survival probability and v_i^j is measurement errors which are uncorrelated with pl_i and with each other. Therefore the correlation between pl_i^1 and pl_i^2 (or pl_i^3 and pl_i^4) is the ratio to how much pl_i^j depends on pl_i .

¹⁹ See Griliches and Mason (1972).

Now we assume that there is no distinction between survival probabilities within or across waves and there is no distinction between probability work and survival probability within or across waves. That is,

$$\text{Corr} [pl^1_{75}, pl^2_{75}] = \text{Corr} [pl^1_{85}, pl^2_{75}] \text{ and } \text{Corr} [pl^1_{75}, pl^2_{85}] = \text{Corr} [pl^1_{85}, pl^2_{85}],$$

$$\text{Corr} [pw62, pl^1_{75}] = \text{Corr} [pw65, pl^1_{85}] \text{ and } \text{Corr} [pw62, pl^2_{75}] = \text{Corr} [pw65, pl^2_{85}]$$

where $pw62$ and $pw65$ are defined as probability work after 62 years old (or 65).

Considering the cross equation restriction on the parameters:

$$pw62 = \alpha X + \beta pl_{75} + \mu + \varepsilon_1$$

where X is observable components which varies across respondents and pl_{75} is unobservable variable and μ is unobserved variable affected by survey effect from survey sequence ordering within or across waves, and ε_1 is unobservable error. When we plug $v^j_i = pl^j_i - pl_i$ into the above equation:

$$pw62 = \alpha X + \beta pl^j_{75} - \beta v^j_{75} + \mu + \varepsilon_1 \quad (1)$$

We assume that ε_1 is uncorrelated with v^j_{75} but v^j_{75} is correlated with pl^j_{75} . Since we assumed that there is no distinction between probability work and survival probability within or across waves, it is possible to assume that unobserved variable(μ) affected by survey effect from survey sequence ordering in the relationship between $pw62$ and pl^j_{75} in the equation (1) is identical with that in the relationship between $pw65$ and pl^j_{85} . This assumption enables us to make the following equation.

$$pw65 = \alpha X + \beta pl^j_{85} - \beta v^j_{85} + \mu + \varepsilon_2 \quad (2)$$

The above process implies that those two equations are identical for each respondent and that he (or she) has the same survey sequence ordering effect within or across waves. In the classical errors-in-variable model, OLS estimate of β in equation (1) will be attenuated because of an inherent errors-in-variable bias and omitted variable bias ignoring survey effect from survey sequence ordering²⁰:

$$\text{plim } B_{OLS} = \beta \{1 - [\sigma_{vj75}^2 / (\sigma_{vj75}^2 + \sigma_{pl75}^2)]\}$$

where σ_{pl75}^2 is the variance of the true survival probability and σ_{vj75}^2 is the variance of measurement error in survival probabilities. The inherent errors-in-variable bias of survival probability makes the estimate to be underestimated. If those who have a high probability work are more likely to report a higher survival probability it could be that there exists a positive survey effect. Also as long as there exists a positive survey effect, the coefficient of survival probability will be overestimated. Because errors-in-variable bias tending to underestimate could be canceled out by a positive survey effect tending to overestimate it has possibility that the estimate of β approached to the true value by an action of cancellation.

Now we assume that pl_i^j and μ are uncorrelated with X . Therefore we can get the following equation that represents the general correlations between survival probabilities and unobserved variable affected by survey sequence ordering:

$$\mu = \gamma pl_{75}^j + \gamma pl_{85}^j + \varepsilon_3 \quad (3)$$

where γ means survey effect occurring in the correlation between survival probabilities and probability work within or across waves. By substituting equation (3) into (1), we can obtain the following equation.

²⁰ See Appendix 3 for derivation.

$$pw62 = \alpha X + (\beta + \gamma)pl_{75}^j + \gamma pl_{85}^j - \beta v_{75}^j + \varepsilon_1' \quad (4)$$

where $\varepsilon_1' = \varepsilon_1 + \varepsilon_3$.

As long as equation (1) is identical for each respondent and survival probability is uncorrelated with other observable variables, we enable survival probability at different ages to enter into this equation. That is, pl_{85}^j is added in equation (4) newly. As we include pl_{85}^j we control survey effect by removing omitted variable bias in equation (1). If there exists a positive survey effect the estimate of β in equation (1) has larger value than the estimate of β in equation (4) because β in equation (1) is not controlled from survey effect a downward bias from errors-in-variable could be canceled out by an upward bias from a positive survey effect.

Now consider the difference between equations (1) and (2) to eliminate the survey effect.

$$pw62 - pw65 = \beta(pl_{75}^j - pl_{85}^j) - \beta(v_{75}^j - v_{85}^j) + (\varepsilon_1 - \varepsilon_2) \quad (5)$$

In the classical errors-in-variable model, the general fixed-effect estimate of β will be more attenuated than the OLS estimate because of the correlation between true survival probabilities within waves, $(1 - \rho_{pl})^{21}$:

$$plim B_{FE} = \beta (1 - \{ \sigma_{vj75}^2 / [\sigma_{vj75}^2 + \sigma_{pl75}^2(1 - \rho_{pl})] \})$$

where ρ_{pl} is the correlation between true survival probabilities within and across waves.

As method to reduce measurement error from errors-in-variable bias and survey effect from omitted variable bias we will use two methods. One is to use

²¹ See Appendix 3 for derivation.

averaged survival probabilities across waves and the other is to use instrumental variables²².

First, we use averaged survival probabilities across waves, that is, we will use $(pl^1_{75} + pl^2_{75})/2$, $(pl^1_{85} + pl^2_{85})/2$, $\{(pl^1_{75} + pl^2_{75}) - (pl^1_{85} + pl^2_{85})\}/2$ for each survival probability in equation (1), (4), and (5). Krueger claimed that this procedure reduces inherent errors-in-variable bias and omitted variable bias by making a modified OLS and fixed-effect model. Krueger said that all estimates of averaged survival probability is less underestimated by a fraction of total variance, $2\sigma^2_{pl}$, and the correlation between survival probabilities within and across wave, $(1 - \rho_{pl})$ than the estimates of OLS and fixed-effect respectively. When we compare OLS and fixed-effect estimates of β with the averaged OLS and fixed-effect estimates of β , the averaged OLS and fixed-effect estimates of β will always have larger values because of total variance and correlation between survival probabilities within and across wave, that is $plim B_{AVG} > plim B$.

Secondly, we will use the instrumental variables to reduce measurement error and survey effect. Previously we assumed that measurement errors of survival probabilities across waves are uncorrelated with each other. Since we have multiple survival probabilities across waves the corresponding survival probability($pl75$) in wave (2) is used as instrumental variables for $pl75$ in wave (1) and survival probability($pl75$) in wave (4) is used as instrumental variables for $pl75$ in wave (3).

If our result is that survey effect, γ , has negative sign, it implies that β in equation (1) produces serious underestimation because of downward bias from

²² See Alan Krueger (1994)

errors-in-variable as well as a negative survey effect. Therefore, OLS estimate of β in equation (1) neglecting omitted variable bias has smaller value than OLS estimate of β in equation (4) adding survey effect. To confirm this result we will use averaged survival probabilities and instrumental variable method. If the estimates of β from averaged method are larger than the corresponding estimates in OLS and fixed-effect from equation (1), (4), and (5) this result addresses further proof that β in equation (1) is seriously underestimated by errors-in-variable bias and survey effect because there does not exist any action to cancellation from biases. However, if we agree the corresponding survival probabilities in wave (2) and (4) as valid instruments and there exists negative survey effect and then IV estimate of β is much larger than OLS estimate, we could not confirm whether estimate of β in equation (1) is more seriously underestimated by negative survey effect or is more seriously underestimated by errors-in-variable bias.

Contrary, if our result is that survey effect, γ , has positive sign, it implies that because downward bias from errors-in-variable could be canceled out by upward bias from survey effect OLS estimate of β in equation (1) has larger value than OLS estimate in equation (4). Therefore, it has possibility that IV estimate has a similar value to OLS estimate by those two biases working in opposite direction. The fact that IV estimate has a similar value to OLS estimate provides whether estimate of β in equation (1) is more seriously overestimated by positive survey effect or is more seriously underestimated by errors-in-variable bias. This implication could be induced by checking a method to use averaged survival probability to reduce measurement error and survey effect.

3. Data

We used data from the Health and Retirement Study (HRS). The HRS is nationally representative longitudinal study that focuses on persons born in the years 1931 to 1941. It is the biennial survey in which contains information about wage, job characteristics, retirement plan, labor market behavior, demographic information and detailed health measures²³. The initial wave of interviews took place in 1992 and the second wave followed in 1994 and thereafter. Since our focus is on the effect of measurement error from health status and survey effect from survey sequence ordering within or across waves we used wave (1) and wave (3) data. The wave (1) consists of 12,652 observations and the wave (3) 10,964. We limited our samples to age eligible persons between 50 and 61 years old²⁴. With age restriction 2544 observations were dropped in wave (1) and 4410 in wave (3).

In wave (1) respondents were asked to answer probability on a scale of 0 to 10. The scale is expanded to run from 0 to 100 thereafter. Thus, there is likely to be some error across waves and we adjust to the rescaling between 0 and 1. As long as the rescaling is uncorrelated with other independent variables in the model, it will have negligible effect on our results.

In the first model, we use retirement decision as labor force exit variable. Since respondents answered “*question not relevant*” if they do not presently work for pay or are a home-worker, or have not worked for ten or more years, “*completely retired*” is referred to be equivalent to “*question not relevant*”. We made a

²³ See Juster and Suzman (1995) and <http://hrsonline.isr.umich.edu>

²⁴ Age restriction rules out Social Security early retirement benefits effect. Social Security early retirement benefits are one important source to retiree’s income that could enable some 62 – 64 years old to reduce labor supply prior to age 65 at which they can have Social Security normal retirement benefits.

dichotomous labor force exit variable in the way that categorizes “*completely retired*” and “*question not relevant*” as retirement status (Retire=1). HRS data include various health problem variables like asthma, back or leg pain, kidney, cholesterol, headache, etc and also various diseases problems like hypertension, diabetes, cancer, lung, stroke told by doctors. We made a dichotomous subjective health variable in the way that categorizes “*fair*” and “*poor*” among five level answers as bad health status (Bad=1). For retirement decision, various health condition, and demographic variables, 2871 missing data were dropped in wave (1) and 2052 in wave (3). Because only current workers respond to wage question, the last job or the previous wave’s wage rate was replaced for retirees or non-respondents. 1088 missing data in wage variables were dropped in wave (1) and 635 in wave (3). HRS data include various information relating to retirement plans like employer-provided health insurance (EPHI), life insurance, private insurance, and so on. If those variables prognosticate the prospective preferences of risk among respondents they have an important meaning in the labor force exit²⁵. For retirement plans, 78 missing data were dropped in wave (1) and 331 in wave (3). Finally 8247 sampling data were collected in wave (1) and 4806 in wave (3). Here, 485 respondents were added newly into wave (3) because of the time difference(temporal gap) between waves²⁶.

²⁵ A risk-averse individual is more likely to place a higher value on health insurance. If such an individual recognized that he lost his EPHI upon leaving the labor force, then he would be more likely to remain longer with his employer than he would. In contrast, an individual who lost their health insurance upon retiring could try to buy a private health insurance. However, since such insurances require much higher premiums for a given level of coverage than EPHI a private health insurance is not good retirement plan for some older persons who want to build a successful retirement plan.

²⁶ For example, 61 years old in wave (1) who is eligible in our sample become 65 years old when he(he) interviews again in wave (3), but he(he) will be excluded in our sample of wave (3) by age

In the second model, to avoid occurrence of justification bias that prevents collecting the correct information for non-workers, we limited to current workers among age eligible persons. To detect survey effect from survey sequence ordering, survival probability variables in wave (2) or wave (4) were also used with those corresponding variables in wave (1) or wave (3). Thus 4902 observations in wave (1) and 4353 in wave (3) were dropped. HRS reports two pension plans; one is defined benefits (DB) and the other is defined contributions (DC)²⁷. Some pension plans allow retirement at an early age, usually with reduced benefits. Pensions have complicated effects on the incentives that encourage workers to remain in employment for some period (substitution effect) or to leave the current job at earlier age than normal retirement age (income effect)²⁸. HRS defines that most pension plans have a normal retirement age at which full benefits could be received and pension incentives are more powerful in defined benefits plan. Since pension has a complicated impact on labor force participation we included each pension plan and a combined plan in our model. Also we included various job requirement variables. Finally, 4973 sampling data were collected in wave (1) and 1850 in wave (3). For more description of data, refer to Appendix 4.

Table 2 shows the means and standard error of several variables. In the sample (1), the mean of retirement is high in wave (3). The reason of a high percent of retirement in wave (3) could be described intuitively in that the high wage rate

restriction. In contrast, 46 years old in wave (1) who is not eligible in our sample become 50 years old when he/she interviews again in wave (3), and he/she will be included in the sample of wave (3).

²⁷ Defined benefits are usually based on a formula involving age, years of service, and salary. Defined contributions are based on money accumulated in an account for workers.

²⁸ Christopher J. Ruhm (1994) said that most defined pension plans encourage early retirement, typically prior to age 65.

corresponds to the high labor supply, that is, the low wage rate corresponds to the high retirement and the high age corresponds to the high retirement. The mean of mortal probability ($1-p/75$) is same in both waves but the means of subjective health and health problems are a little high in wave (1). Usually the high wage rate might cause person to report a low mortal probability and subjective health (bad=1) and health problems. However, mortal probability is the same in both waves and subjective health and health problems are higher in wave (1). It is implied that if wage plays much a greater role in retirement decision than health status, this phenomenon of health measures could be neglected; otherwise there exists measurement error in all health variables.

In the sample (2), the high wage rate corresponds to the high probability work in wave (3) but there is almost same survival probability in both waves. Defined benefits and defined contributions have a higher percentage in wave (1). We should notice that both waves have almost same mean of survival probabilities but probability work in wave (1) is much lower than wave (3). If this disparity stems from survey effect we could describe this phenomenon by using the classical errors-in-variable model.

4. Results

The results of the first model are Table 3 to Table 7 and the results of the second model are Table 8 to Table 10. Table 3 is the result of regression using the reduced form and shows the effect on retirement decision, objective health(h^o), subjective health(h^s), and mortal probability to age 75($1-pl75$) respectively. In here, objective health is the total number of health problems. Therefore, objective health and mortal probability are continuous variables while retirement decision and subjective health are dichotomous variables with 1's representing labor force exit and bad health status respectively. The wage has much smaller effect on retirement decision than other explanatory variables in wave (1). The positive sign of wage means that those who are compensated in the work place are likely to retire and to report bad health in wave (1). However, the negative sign of wage in wave (3) means that they are not likely to retire and to report bad health. The fact that wage has different sign on retirement decision suggests that measurement error from health status causes to have influence on wage. Here we could predict β_3 's sign showing the relationship between wage and subjective health. Column (3) and (3)' show β_3 's sign. Since wage in wave (1) has a different sign from wave (3) we would apply our two judgment conditions. The estimates shown in Table 3 are the basis of our health proxy estimation in Table 6.

Table 4 reports correlations across equations after regression of the reduced form model in Table 3. All correlations are positive and strong relationship except the correlation between retirement decision and mortal probability in both waves. The fact that they have a positive small correlation comparatively can propose

implicitly that the diseases problems relating to longevity could be the less threatening components for retirement decision than health problems in HRS. The cross equation correlation pattern in wave (3) is a little higher.

Table 5 shows estimations of the various health measures. The wage in wave (1) has positive sign, but exerts *much* smaller effect than health measures while wage effect in wave (3) has negative sign, but larger than health measure. The fact that wage has small value in wave (1) implies that there exists a possibility that bias of endogenous problem or inherent errors-in-variable bias from health measure could have influence on wage variable. Social security encourages older persons to retire early while there is negative effect on retirement for EPHI, private health insurance, and life insurance. We should notice that the log-likelihood statistics have a small difference in wave (3). This means that the effect of each health status on labor force exit could not be ignored.

Table 6 and table 7 show the various estimations of subjective and objective health problem proxy and mortal probability proxy based on Table 3 and Table 5. The column (1) of both tables is the estimation without health measure. The column (2), (3) and (4) are results that use various health measures as the proxies for the true health. Here, objective health, h^o , is the number of health problem. As we said previously we would use the sign of wage rate in Table 3 as β_3 's sign. β_3 is positive in wave (1) but is negative in wave (3).

Table 6 shows the results of labor force exit in wave (1). In wave (1), the estimate of subjective health proxy in column (3) has smaller value than that of objective health problem proxy in column (2) and has larger value than that of mortal

probability proxy in the column (4). Remember that the estimate of subjective health proxy has simultaneity bias inducing to overestimate and errors-in-variable bias inducing to underestimate together, and that objective health problem proxy and mortal probability proxy have only errors-in-variables bias. The column (5) and (6) are reporting the use of objective health problems proxy and mortal probability proxy as instrumental variables respectively. The health coefficients (0.455, 0.278) of column (5) and (6) do not have larger values than that (0.686) of column (3). If we accept objective health problems proxy and mortal probability proxy as the reasonable instrumental variables the estimate of subjective health proxy, α_l , in column (3) is producing overestimate in OLS regression. As we said model section, it is implied that because α_l is overestimated subjective health could be more affected by simultaneity bias from endogenous problem than inherent errors-in-variable bias. By following condition (i) we could predict that β_l , estimate of wage in column (3), might be underestimated in OLS regression. We note that if we use IV method subjective health, α_l , is identified but wage, β_l , is underestimated when β_3 is positive. Compare wage estimate, β_l , in column (3) with column (5) and (6). Because wage estimate, β_l , in column (3) is underestimated and larger than column (5) and (6), it is implied that the estimate of β_l in column (3) gives us a more appropriate lower bound for wage.

Table 7 shows the results of labor force exit in wave (3). In wave (3), wage has negative sign and we proposed β_3 to be negative. The health coefficients (0.720, 0.649) of column (5) and column (6) have larger value than that (0.605) of column (3) unlike wave (1)'s case. If we accept proxy of objective health problem and proxy of

mortal probability as valid instrumental variables the estimate of subjective health proxy, α_l , in column (3) should be underestimated and β_l be overestimated when β_3 is negative by following condition (ii). Our result implied that because α_l in column (3) is underestimated subjective health could be more affected by inherent errors-in-variable bias than simultaneity bias and because β_l in column (3) is smaller than column (5) and (6) the estimate of β_l in column (3) gives us a more appropriate upper bound for wage. By deduction it is implied that our choice of instrumental variables is reasonable. Also we could address that the fact that wage of column (3) seems to have a big impact on retirement decision in OLS regression does not rule out a possibility that measurement error from subjective health has a more correlated effect on wage. In other words, we could not rule out that as long as the simultaneity bias will give rise to upward bias and errors-in-variable bias to downward bias both biases could be cancelled out in OLS regression.

In brief, because of problem occurring measurement error we could provide a lower or upper bound for health and wage variables. Suppose that our conditions are satisfied. When $\beta_3 > 0$ in wave (1), the estimate of subjective health in column (3) makes an upper bound for health measure and a lower bound for wage in conventional OLS regression while when $\beta_3 < 0$ in wave (3), the estimate of subjective health in column (3) makes a lower bound for health measure and an upper bound for wage.

Now we will mention another measurement error in the second model's result. The second model's results will be explained from Table 8. Table 8 shows the correlation between the probability work and survival probability. We make sure that

the independent measures of survival probability show a similar correlated relationship in both waves. We make sure again that the probability work and survival probability are weakly correlated, however, they are more highly correlated in wave (3). The measured variance between 48 and 56 percent in survival probability stems from measurement error in wave (1) while the measured variance between 45 and 54 percent stems from in wave (3). Because both waves have a similar percentage we can expect that they will have the same survey effect sign.

Table 9 shows the estimates of the effect of survival probability on probability work in wave (1). Column (1) shows OLS estimation for equation (1). Column (1) explains that 10% increase of survival probability causes to increase 1.5% of probability work. Wage effects are small while pension and union have a great impact on labor force participation. As we explained in the previous section the defined benefits have a powerful negative effect on labor force participation, that is, pension's income effect that encourages leaving the work place at the early age dominates substitution effect in our sample. The result of column (2) on the Table 9 corresponds to OLS estimation of equation (4) controlling survey effect. The coefficient of $pl85$ is a measure of the survey effect, γ , and has a positive sign. Because there exist positive survey effect(0.683) as long as downward bias from errors-in-variable could be canceled out by upward bias from survey effect, OLS estimate of $\beta(pl)$ in column (1) has larger value than column (2). Column (3) shows the general fixed-effect result that eliminates survey effect. Because $plim_{BFE}$ is more attenuated than $plim_{BOLS}$ the general fixed-effect's estimate of survival probability has a smaller value than OLS estimate in column (1) or (2). If we have a consistent

and acceptable result from the instrumental variables method to reduce measurement error and survey effect we could confirm whether the estimate of OLS estimates could be overestimated or underestimated. Column (4) and (5) are showing the results using survival probability, pl , in wave (2) and the first difference between wave (1) and (2) as the instrumental variables. Both coefficients of survival probability ($pl75$) in the column (4) and (5) have much larger values than that in column (1). If we accept the instrumental variables method as the valid estimation, the fact that IV estimates have much larger values than OLS estimates provides downward bias from errors-in-variable works more seriously than upward bias from survey effect. Column (6), (7), and (8) show the averaged survival probability method to reduce measurement error and survey effect. All estimates of averaged survival probability have larger values than the corresponding estimates in the column (1), (2), and (3). This result gives further evidence that errors-in-variable bias is making a more serious downward bias than upward bias from survey effect in OLS regression.

Table 10 shows the estimates of the effect of survival probability on probability work in wave (3). Wage has a very small effect while defined benefits have a strong income effect. Likewise wave (1), there exists positive survey effect. Because column (1) does not control survey effect β of column (1) has larger value than β of column (2), (3). The fact that all estimates of the averaged survival probability in column (6), (7), and (8) have larger values than the corresponding estimates in the column (1), (2), and (3), provides evidence that downward bias from errors-in-variable bias is canceled out by upward bias from survey effect in OLS

regression. Both coefficients of survival probability in column (4) and (5) using instrumental variables have a little larger values than the estimate of column (1). If we accept survival probability, pl , in wave (4) and the first difference from wave (4) as the acceptable instrumental variables there is a different result from wave (1). Remember that the positive survey effect overestimates but errors-in-variable bias underestimates our results. The coefficients of $pl75$ in column (4) and (5) have *a little* larger values than that of $pl75$ in column (1). However, in wave (1) of the table 8, the coefficients of $pl75$ in column (4) and (5) have *a great* larger values. This means that in wave (3), upward bias from survey effect and downward bias from errors-in-variable could be well-balanced and canceled out in the view of quantity, while in wave (1), those biases could be unbalanced and canceled out. Therefore in some way, bias from survey effect is the non-ignorable component for researchers as well as bias from measurement error.

5. Conclusion

Many studies for the relationship between labor supply and health status do not have a unanimous agreement on the usage of subjective health because of the difficulty of correct measures of health status. Alternatively, many researchers have employed the objective health measure and it was suggested by some of them that financial compensation plays a greater role in labor supply than does the subjective health. However, they have assumed previously that the errors-in-variable bias of objective health is not large. Thus, this bias can be ignored in the relationship between labor supply and health status. Although there is a reliable objective health measure we will face the identification problem of other coefficients.

In our research we added mortal probability as well as subjective and objective health and structured that mortal probability variable has an inherent errors-in-variable bias in the way that the objective health does. To approach true health status we used objective health problem proxy and mortal probability proxy as instrumental variables for subjective health proxy and then made two judgment conditions in which we could find more appropriate estimates. Considering health proxies and instrumental variables method, our results provide us more appropriate health measure. The disparities of results in wave (1) and wave (3) provide lower and upper bounds to health and wage for the labor force exit. In views of these measured bounds for health and wage, we can confirm that health plays a larger role in the labor force exit than wage.

In the second model, our research is whether probability questions(probability work and survival probability) within and across waves cause another measurement

error. The fact that in HRS, wave (1) and (2)'s survey sequence ordering has a little different ordering system each other while wave (3) and (4)'s ordering has the same ordering system each other enables us make a controversy about unobserved measurement error. We constructed the classical errors-in-variable model with cross equation restriction on the parameters and then applied a first difference model to describe measurement errors and survey effect. We found that survey effect for the survival probability exists in both waves by using two methods, one is to use averaged survival probabilities and the other is to use instrumental variables. However, the possibility that survey effect and errors-in-variable bias are well-balanced and canceled out, or they are unbalanced and canceled out suggests that survey effect is not a negligible component in the further researches.

Table 1 : Survey Sequence Ordering of Key Questions in HRS

	Wave 1	Wave 2	Wave 3	Wave 4
Section Title	Demographics	Demographics	Demographics	Demographics
	Health	Health	Health	Health
	Employment (<i>pw</i>)	Employment	Employment	Employment
	Cognition (<i>pl</i>)	Net Worth	Expectation (<i>pw/pl</i>)	Expectation (<i>pw/pl</i>)
	Net Worth	Cognition (<i>pw/pl</i>)	Net Worth	Net Worth
	Income	Income	Income	Income
	Health Insurance	Health Insurance	Health Insurance	Health Insurance

* Expectation Section's questionnaires in wave (3) and wave (4) have the same as Cognition's questionnaires in wave (1) and wave (2).

Table 2 : Sample Statistics

	Sample 1 (Labor Force Exit)				Sample 2 (Labor Force Participation)			
	Wave 1		Wave 3		Wave 1		Wave 3	
<u>Dependent</u>	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
Retire	0.18	0.39	0.24	0.43				
<i>Pw62</i>					0.47	0.39	0.64	0.32
<i>Pw65</i>					0.25	0.33	0.34	0.34
<u>Health</u>								
Subjective Health	0.21	0.40	0.19	0.40				
1- <i>P175</i>	0.35	0.30	0.35	0.30				
<i>P175</i>					0.69	0.27	0.70	0.24
<i>P185</i>					0.46	0.31	0.47	0.30
Health Problem #	1.39	1.36	1.02	1.42				
Asthma	0.06	0.24	0.10	0.30				
Back	0.34	0.47	0.31	0.46				
Legs	0.35	0.48	0.14	0.35				
Kidney	0.10	0.30						
Stomach	0.09	0.29						
Cholesterol	0.24	0.43						
Breath			0.13	0.33				
Dizziness			0.08	0.28				
Headache			0.11	0.31				
Fatigue			0.15	0.36				
<u>Financial</u>								
Wage	10.83	42.84	8.62	73.06	4.05	43.79	5.08	111.48
<u>Demographic</u>								
Age	55.59	3.26	57.12	2.71	55.31	3.23	56.82	2.77
Sex	0.47	0.45	0.41	0.49	0.49	0.50	0.46	0.50
Race	0.73	0.44	0.81	0.39	0.78	0.42	0.85	0.36
Education	12.27	3.09	12.59	2.83	12.81	2.82	13.25	2.62
Married	0.74	0.44	0.74	0.44	0.75	0.43	0.73	0.44
<u>Retirement Plan</u>								
SS	0.08	0.27	0.13	0.33				
Federal HI	0.14	0.35	0.89	0.31				
Medicare	0.06	0.24	0.07	0.25				
EPHI	0.73	0.44	0.72	0.45				
Private HI	0.16	0.37	0.19	0.39				
Life Insurance	0.76	0.43	0.78	0.41				
<u>Job/Requirement</u>								
Union					0.23	0.42	0.04	0.20
DB					0.45	0.50	0.25	0.43
DC					0.24	0.42	0.18	0.38
DB+DC					0.03	0.17	0.04	0.19
Physical Effort					0.38	0.48	0.10	0.30
Heavy Lifting					0.15	0.36	0.04	0.18
Stooping					0.25	0.43	0.07	0.25
Good Eyesight					0.89	0.32	0.28	0.45
Attention					0.84	0.36	0.26	0.44
People Skill					0.87	0.33	0.27	0.44
Computer					0.30	0.50		
Analyzing					0.42	0.49		
Keep up with					0.54	0.50		
Repetition					0.63	0.48		
Learning					0.51	0.50		
Size	8247		4806		4973		1850	

Table 3 : Regression using Reduced Form

Dependent	Wave 1				Wave 3			
	Retire	h ^o	h ^s	1-P/75	Retire	h ^o	h ^s	1-P/75
	(1)	(2)	(3)	(4)	(1)'	(2)'	(3)'	(4)'
Wage	0.009*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.000 (0.000)	-0.626*** (0.032)	-0.000 (0.000)	-0.118*** (0.026)	0.000 (0.000)
Age	0.067*** (0.006)	0.008* (0.004)	0.011** (0.005)	-0.002* (0.001)	0.068*** (0.009)	-0.012 (0.007)	0.000 (0.009)	-0.001 (0.002)
Sex	-0.194*** (0.036)	-0.133*** (0.029)	0.066* (0.035)	0.050*** (0.007)	-0.096* (0.050)	-0.217*** (0.041)	0.042 (0.049)	0.043*** (0.009)
Race	0.042 (0.042)	0.087** (0.034)	-0.195*** (0.039)	0.028*** (0.008)	0.031 (0.061)	-0.012 (0.051)	-0.141** (0.057)	0.028** (0.011)
Education	-0.003 (0.006)	-0.058*** (0.005)	-0.079*** (0.006)	-0.014*** (0.001)	0.004 (0.009)	-0.075*** (0.007)	-0.089*** (0.009)	-0.015*** (0.002)
Married	0.231*** (0.045)	-0.223*** (0.035)	-0.131*** (0.041)	-0.014* (0.008)	0.271*** (0.057)	-0.159*** (0.047)	-0.080 (0.054)	-0.009 (0.010)
SS	0.560*** (0.073)	0.391*** (0.071)	0.431*** (0.074)	0.045*** (0.016)	0.369*** (0.070)	0.191*** (0.066)	0.228*** (0.072)	0.043*** (0.014)
Federal HI	0.582*** (0.060)	0.424*** (0.055)	0.434*** (0.060)	0.022* (0.012)	0.515*** (0.091)	0.205** (0.083)	0.128 (0.089)	0.019 (0.018)
Medicare	-0.001 (0.093)	0.247*** (0.089)	0.047 (0.093)	0.033 (0.020)	1.196*** (0.108)	1.028*** (0.092)	1.024*** (0.095)	0.118*** (0.020)
EPHI	-0.227*** (0.046)	-0.182*** (0.038)	-0.196*** (0.043)	-0.018** (0.009)	-0.331*** (0.069)	-0.347*** (0.062)	-0.266*** (0.070)	-0.015 (0.014)
Private HI	-0.240*** (0.052)	-0.094** (0.040)	-0.086* (0.049)	-0.010 (0.009)	-0.156** (0.066)	-0.136** (0.055)	-0.061 (0.066)	-0.025** (0.012)
Life Insurance	-0.156*** (0.043)	-0.061* (0.036)	-0.060 (0.041)	-0.014* (0.008)	-0.246*** (0.055)	-0.206*** (0.050)	-0.245*** (0.055)	-0.027** (0.011)
Size	8247				4806			

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

note that $\beta_3 > 0$ in wave (1) and $\beta_3 < 0$ in wave (3)

Table 4 : Correlations of Reduced Form Regression

	Wave 1				Wave 3			
	Retire	h^o	h^s	1- $P175$	Retire	h^o	h^s	1- $P175$
Retire	1.000				1.000			
h^o	0.790	1.000			0.716	1.000		
h^s	0.707	0.916	1.000		0.742	0.954	1.000	
1- $P175$	0.494	0.757	0.829	1.000	0.568	0.825	0.879	1.000

Table 5: Labor Force Exit using Proxy Variable of Health

	Wave 1			Wave 3		
	(1)	(2)	(3)	(1)'	(2)'	(3)'
Subjective Health	0.686*** (0.043)			0.605*** (0.058)		
1- <i>P175</i>			0.322*** (0.059)			0.310*** (0.076)
Asthma		0.174** (0.069)			-0.016 (0.081)	
Back		0.179*** (0.039)			0.136*** (0.052)	
Legs		0.223*** (0.040)			0.099 (0.068)	
Kidney		0.249*** (0.056)				
Stomach		0.154*** (0.059)				
Cholesterol		0.125*** (0.041)				
Breath					0.213*** (0.077)	
Dizziness					0.201** (0.086)	
Headache					0.007 (0.075)	
Fatigue					0.279*** (0.070)	
Wage	0.008*** (0.000)	0.009*** (0.000)	0.009*** (0.000)	-0.605*** (0.032)	-0.607*** (0.032)	-0.621*** (0.032)
Age	0.069*** (0.006)	0.068*** (0.006)	0.068*** (0.006)	0.072*** (0.009)	0.072*** (0.009)	0.068*** (0.009)
Sex	-0.221*** (0.037)	-0.167*** (0.037)	-0.212*** (0.037)	-0.120** (0.051)	-0.084* (0.051)	-0.110** (0.050)
Race	0.090** (0.044)	0.019 (0.043)	0.030 (0.043)	0.067 (0.063)	0.029 (0.062)	0.020 (0.062)
Education	0.015** (0.007)	0.006 (0.007)	0.002 (0.007)	0.022** (0.009)	0.016* (0.009)	0.010 (0.009)
Married	0.288*** (0.046)	0.279*** (0.046)	0.235*** (0.045)	0.303*** (0.058)	0.312*** (0.058)	0.279*** (0.057)
SS	0.486*** (0.074)	0.528*** (0.074)	0.550*** (0.073)	0.337*** (0.070)	0.349*** (0.071)	0.358*** (0.070)
Federal HI	0.506*** (0.062)	0.530*** (0.061)	0.576*** (0.060)	0.496*** (0.093)	0.485*** (0.092)	0.513*** (0.091)
Medicare	-0.010 (0.095)	-0.035 (0.094)	-0.010 (0.093)	1.033*** (0.112)	1.115*** (0.111)	1.173*** (0.109)
EPHI	-0.181*** (0.047)	-0.210*** (0.047)	-0.222*** (0.046)	-0.282*** (0.070)	-0.282*** (0.070)	-0.330*** (0.069)
Private HI	-0.220*** (0.053)	-0.227*** (0.052)	-0.232*** (0.052)	-0.130* (0.067)	-0.119* (0.067)	-0.147** (0.066)
Life Insurance	-0.145*** (0.044)	-0.155*** (0.043)	-0.153*** (0.043)	-0.210*** (0.056)	-0.229*** (0.056)	-0.239*** (0.055)
Log likelihood	-3069.414	-3116.532	-3180.010	-1890.028	-1896.788	-1934.824
Size		8247			4806	

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Table 6 : Labor Force Exit in Wave (1) based on Reduced Form Model

	Wave 1					
	None	h^0	h^s	1- <i>P175</i>	$h^s_{IV(a)}$	$h^s_{IV(b)}$
	(1)	(2)	(3)	(4)	(5)	(6)
Health		0.743*** (0.041)	0.686*** (0.043)	0.322*** (0.059)	0.455*** (0.025)	0.278*** (0.029)
Wage	0.001*** (0.000)	0.000 (0.000)	0.008*** (0.000)	0.009*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Age	0.015*** (0.000)	0.009*** (0.000)	0.069*** (0.006)	0.068*** (0.006)	0.013*** (0.000)	0.014*** (0.000)
Sex	-0.040*** (0.001)	0.041*** (0.005)	-0.221*** (0.037)	-0.212*** (0.037)	-0.049*** (0.001)	-0.045*** (0.001)
Race	0.007*** (0.002)	-0.046*** (0.003)	0.090** (0.044)	0.030 (0.043)	0.035*** (0.002)	0.024*** (0.002)
Education	-0.003*** (0.000)	0.035*** (0.002)	0.015** (0.007)	0.002 (0.007)	0.008*** (0.001)	0.004*** (0.001)
Married	0.042*** (0.002)	0.190*** (0.008)	0.288*** (0.046)	0.235*** (0.045)	0.068*** (0.002)	0.058*** (0.002)
SS	0.209*** (0.003)	-0.061*** (0.015)	0.486*** (0.074)	0.550*** (0.073)	0.141*** (0.005)	0.168*** (0.005)
Federal HI	0.165*** (0.003)	-0.115*** (0.016)	0.506*** (0.062)	0.576*** (0.060)	0.108*** (0.004)	0.130*** (0.004)
Medicare	0.031*** (0.004)	-0.152*** (0.011)	-0.010 (0.095)	-0.010 (0.093)	0.014*** (0.004)	0.020*** (0.004)
EPHI	-0.067*** (0.002)	0.054*** (0.007)	-0.181*** (0.047)	-0.222*** (0.046)	-0.032*** (0.003)	-0.046*** (0.003)
Private HI	-0.053*** (0.002)	0.010** (0.004)	-0.220*** (0.053)	-0.232*** (0.052)	-0.038*** (0.002)	-0.044*** (0.002)
Life Insurance	-0.041*** (0.002)	0.000 (0.003)	-0.145*** (0.044)	-0.153*** (0.043)	-0.029*** (0.002)	-0.034*** (0.002)
Size	8247					

a. subjective health is instrumented for proxy of health problems numbers

b. subjective health is instrumented for proxy of mortal probability

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Table 7 : Labor Force Exit in Wave (3) based on Reduced Form Model

	Wave 3					
	None	h^0	h^s	$1-P/75$	$h^s_{IV(a)}$	$h^s_{IV(b)}$
	(1)	(2)	(3)	(4)	(5)	(6)
Health		0.498*** (0.056)	0.605*** (0.058)	0.310*** (0.076)	0.720*** (0.090)	0.649*** (0.075)
Wage	-0.000*** (0.000)	-0.000*** (0.000)	-0.605*** (0.032)	-0.621*** (0.032)	-0.000*** (0.000)	-0.000*** (0.000)
Age	0.019*** (0.001)	0.014*** (0.001)	0.072*** (0.009)	0.068*** (0.009)	0.019*** (0.001)	0.019*** (0.001)
Sex	-0.070*** (0.003)	-0.149*** (0.009)	-0.120** (0.051)	-0.110** (0.050)	-0.065*** (0.004)	-0.069*** (0.003)
Race	0.015*** (0.004)	0.009** (0.004)	0.067 (0.063)	0.020 (0.062)	-0.015** (0.006)	0.013** (0.005)
Education	-0.005*** (0.001)	-0.035*** (0.003)	0.022** (0.009)	0.010 (0.009)	-0.023*** (0.002)	-0.007*** (0.002)
Married	0.081*** (0.004)	0.017** (0.008)	0.303*** (0.058)	0.279*** (0.057)	0.065*** (0.005)	0.080*** (0.004)
SS	0.152*** (0.005)	0.230*** (0.010)	0.337*** (0.070)	0.358*** (0.070)	0.200*** (0.008)	0.155*** (0.007)
Federal HI	0.157*** (0.007)	0.244*** (0.012)	0.496*** (0.093)	0.513*** (0.091)	0.181*** (0.008)	0.159*** (0.007)
Medicare	0.471*** (0.007)	0.938*** (0.053)	1.033*** (0.112)	1.173*** (0.109)	0.753*** (0.036)	0.490*** (0.030)
EPHI	-0.158*** (0.005)	-0.300*** (0.017)	-0.282*** (0.070)	-0.330*** (0.069)	-0.223*** (0.010)	-0.162*** (0.008)
Private HI	-0.039*** (0.004)	-0.095*** (0.008)	-0.130* (0.067)	-0.147** (0.066)	-0.057*** (0.005)	-0.040*** (0.005)
Life Insurance	-0.095*** (0.004)	-0.179*** (0.010)	-0.210*** (0.056)	-0.239*** (0.055)	-0.146*** (0.008)	-0.098*** (0.007)
Size	4806					

a. subjective health is instrumented for proxy of health problems numbers

b. subjective health is instrumented for proxy of mortal probability

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Table 8 : Correlation Matrices for Probability Variables

1) wave 1

	<i>Pw62</i>	<i>Pw65</i>	<i>Pl75¹</i>	<i>Pl75²</i>	<i>Pl85¹</i>	<i>Pl85²</i>
<i>Pw62</i>	1.000					
<i>Pw65</i>	0.674	1.000				
<i>Pl75¹</i>	0.095	0.094	1.000			
<i>Pl75²</i>	0.074	0.062	0.473	1.000		
<i>Pl85¹</i>	0.084	0.120	0.725	0.439	1.000	
<i>Pl85²</i>	0.059	0.077	0.435	0.693	0.515	1.000

2) wave 3

	<i>Pw62</i>	<i>Pw65</i>	<i>Pl75³</i>	<i>Pl75⁴</i>	<i>Pl85³</i>	<i>Pl85⁴</i>
<i>Pw62</i>	1.000					
<i>Pw65</i>	0.560	1.000				
<i>Pl75³</i>	0.221	0.158	1.000			
<i>Pl75⁴</i>	0.147	0.115	0.496	1.000		
<i>Pl85³</i>	0.185	0.206	0.707	0.452	1.000	
<i>Pl85⁴</i>	0.133	0.148	0.439	0.696	0.551	1.000

$$\text{Corr}[pl^1_{75}, pl^2_{75}] = 0.473 < \text{Corr}[pl^3_{75}, pl^4_{75}] = 0.496$$

$$\text{Corr}[pl^1_{85}, pl^2_{75}] = 0.439 < \text{Corr}[pl^3_{85}, pl^4_{75}] = 0.452$$

$$\text{Corr}[pl^1_{75}, pl^2_{85}] = 0.435 < \text{Corr}[pl^3_{75}, pl^4_{85}] = 0.439$$

$$\text{Corr}[pl^1_{85}, pl^2_{85}] = 0.515 < \text{Corr}[pl^3_{85}, pl^4_{85}] = 0.551$$

$$\text{Corr}[pw62, pl^1_{75}] = 0.095 < \text{Corr}[pw62, pl^3_{75}] = 0.221$$

$$\text{Corr}[pw65, pl^1_{85}] = 0.120 < \text{Corr}[pw65, pl^3_{85}] = 0.206$$

$$\text{Corr}[pw62, pl^2_{75}] = 0.074 < \text{Corr}[pw62, pl^4_{75}] = 0.147$$

$$\text{Corr}[pw65, pl^2_{85}] = 0.077 < \text{Corr}[pw65, pl^4_{85}] = 0.148$$

Table 9 : OLS, Fixed-Effect, Avg and IV of Labor Force Participation in Wave (1)

	OLS	OLS	FE	IV(d)	FE _{IV(e)}	OLS _{Avg(a)}	OLS _{Avg(b)}	FE _{Avg(c)}
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>PI75</i>	1.466*** (0.204)	0.900*** (0.294)	0.793*** (0.195)	3.039** (1.510)	2.608*** (0.877)	1.852*** (0.248)	1.354*** (0.383)	1.117*** (0.247)
<i>PI85</i>		0.683*** (0.256)					0.546* (0.321)	
Age	0.848*** (0.168)	0.861*** (0.168)		0.807*** (0.173)		0.851*** (0.168)	0.856*** (0.168)	
Sex	11.435*** (1.158)	11.741*** (1.160)		11.618*** (1.237)		11.453*** (1.158)	11.625*** (1.159)	
Race	1.678 (1.159)	2.085 (1.364)		1.485 (1.529)		2.002 (1.361)	2.326* (1.370)	
Education	0.696*** (0.227)	0.674*** (0.227)		0.638*** (0.230)		0.688*** (0.227)	0.676*** (0.227)	
Married	-7.751*** (1.319)	-7.712*** (1.314)		-7.778*** (1.327)		-7.721*** (1.318)	-7.686*** (1.315)	
Wage	-0.002* (0.001)	-0.002* (0.001)		-0.002* (0.001)		-0.002* (0.001)	-0.002* (0.001)	
Union	-9.466*** (1.377)	-9.498*** (1.372)		-9.465*** (1.384)		-9.511*** (1.376)	-9.547*** (1.373)	
DB	-8.519*** (1.426)	-8.532*** (1.421)		-8.474*** (1.433)		-8.488*** (1.426)	-8.439*** (1.422)	
DC	-4.014*** (1.552)	-4.030*** (1.547)		-4.053*** (1.559)		-3.992*** (1.551)	-3.963*** (1.547)	
DB+DC	-6.919** (3.411)	-7.153** (3.401)		-7.028** (3.442)		-6.948** (3.409)	-7.064** (3.400)	
Physical Effort	0.490 (1.415)	0.437 (1.410)		0.421 (1.423)		0.409 (1.414)	0.363 (1.411)	
Heavy Lifting	4.862*** (1.868)	4.827*** (1.862)		4.940*** (1.878)		4.808*** (1.867)	4.777*** (1.862)	
Stooping	-0.642 (1.551)	-0.620 (1.546)		-0.550 (1.559)		-0.564 (1.551)	-0.538 (1.546)	
Good Eyesight	0.340 (1.779)	0.498 (1.774)		-0.069 (1.822)		0.478 (1.778)	0.442 (1.773)	
Attention	0.284 (1.601)	0.329 (1.596)		0.306 (1.610)		0.296 (1.600)	0.333 (1.596)	
People Skill	-0.559 (1.690)	-0.500 (1.685)		-0.781 (1.708)		-0.678 (1.690)	-0.661 (1.685)	
Computer	2.681* (1.373)	2.597* (1.369)		2.736** (1.387)		2.652* (1.372)	2.618* (1.368)	
Analyzing	4.141*** (1.334)	4.086*** (1.330)		4.212*** (1.345)		4.086*** (1.334)	4.067*** (1.330)	
Keep up with	1.243 (1.153)	1.341 (1.149)		1.093 (1.176)		1.345 (1.152)	1.398 (1.149)	
Repetition	-2.957** (1.198)	-2.984** (1.194)		-2.834** (1.208)		-2.927** (1.198)	-2.934** (1.194)	
Learning	3.342*** (1.183)	3.260*** (1.180)		3.304*** (1.194)		3.238*** (1.183)	3.182*** (1.180)	
R ²	0.074	0.075	0.003			0.074	0.075	0.004
Size	4973							

a. Averaged *PI75* is equal to $(PI75^1 + PI75^2)/2$.

b. Averaged *PI* is equal to $(PI75^1 + PI75^2)/2$ and $(PI85^1 + PI85^2)/2$.

c. Averaged fixed-effect is equal to $[(PI75^1 + PI75^2) - (PI85^1 + PI85^2)]/2$.

d. *PI75* is instrumented for *PI75* from wave (2).

e. The first difference is instrumented for each first difference from wave (2).

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Table 10 : OLS, Fixed-Effect, Avg and IV of Labor Force Participation in Wave (3)

	OLS	OLS	FE	IV(d)	FE _{IV(e)}	OLS _{Avg(a)}	OLS _{Avg(b)}	FE _{Avg(c)}
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>PI75</i>	2.959*** (0.298)	2.281*** (0.419)	1.605*** (0.336)	3.214* (1.726)	3.099*** (1.140)	3.290*** (0.345)	2.527*** (0.523)	1.930*** (0.415)
<i>PI85</i>		0.785** (0.344)					0.822* (0.425)	
Age	1.703*** (0.267)	1.716*** (0.265)		1.707*** (0.269)		1.711*** (0.268)	1.713*** (0.266)	
Sex	4.488*** (1.524)	4.828*** (1.521)		5.032*** (1.592)		4.654*** (1.528)	5.008*** (1.529)	
Race	2.276 (2.026)	2.902 (2.031)		2.957 (2.252)		2.337 (2.030)	2.976 (2.044)	
Education	-0.011 (0.291)	-0.002 (0.289)		-0.077 (0.300)		-0.047 (0.292)	-0.053 (0.291)	
Married	-6.978*** (1.697)	-6.913*** (1.685)		-6.667*** (1.706)		-6.687*** (1.702)	-6.727*** (1.691)	
Wage	0.001 (0.001)	0.001 (0.001)		0.001 (0.001)		0.001 (0.001)	0.001 (0.001)	
Union	-9.264** (3.813)	-8.849** (3.794)		-8.718** (3.903)		-9.231** (3.821)	-8.934** (3.800)	
DB	-3.513* (1.816)	-3.200* (1.809)		-3.127* (1.878)		-3.560* (1.820)	-3.235* (1.816)	
DC	1.795 (2.001)	2.054 (1.990)		1.928 (2.063)		1.580 (2.005)	1.840 (1.997)	
DB+DC	-5.906 (3.756)	-5.625 (3.733)		-5.671 (3.791)		-6.551* (3.765)	-6.180* (3.746)	
Physical Effort	-0.044 (3.236)	-0.118 (3.214)		0.199 (3.259)		0.274 (3.244)	0.086 (3.225)	
Heavy Lifting	5.659 (4.778)	6.205 (4.752)		5.711 (4.929)		5.469 (4.789)	5.962 (4.765)	
Stooping	-1.812 (3.672)	-2.220 (3.651)		-2.231 (3.733)		-1.897 (3.679)	-2.171 (3.658)	
Good Eyesight	2.801 (3.528)	3.084 (3.506)		3.017 (3.567)		2.152 (3.536)	2.356 (3.515)	
Attention	-2.751 (3.474)	-2.984 (3.452)		-2.579 (3.538)		-2.730 (3.481)	-2.852 (3.459)	
People Skill	-1.860 (3.201)	-1.853 (3.180)		-2.293 (3.226)		-1.067 (3.205)	-1.041 (3.185)	
R ²	0.110	0.113	0.012			0.107	0.108	0.011
Size	1850							

- Averaged *PI75* is equal to $(PI75^1 + PI75^2)/2$.
- Averaged *PI* is equal to $(PI75^1 + PI75^2)/2$ and $(PI85^1 + PI85^2)/2$.
- Averaged fixed-effect is equal to $[(PI75^1 + PI75^2) - (PI85^1 + PI85^2)]/2$.
- PI75* is instrumented for *PI75* from wave (4).
- The first difference is instrumented for each first difference from wave (4).

*** denotes significance at the 1 percent level
 ** denotes significance at the 5 percent level
 * denotes significance at the 10 percent level

Appendix 1

The following discussion is based on Bound (1991)'s formulation.

I) Derivation of subjective health proxy

Suppose the simple and concrete model about subjective health proxy.

$$r = \alpha_1 h^* + \beta_1 w + \varepsilon_1 \quad (1)$$

$$h^s = \alpha_3 h^* + \beta_3 w + \varepsilon_3 \quad (3)$$

where we give some assumptions applied to the above equations.

1) w and h^* are correlated with each other but uncorrelated with the ε 's.

2) ε_1 and ε_3 are correlated with each other (ρ).

Because w and h^* are correlated with each other Bound made $h^* = \gamma w + \nu$. By

assuming ν is uncorrelated with ε , $\gamma \equiv \sigma_{h^*,w} / \sigma_w^2$. See Bound(1991)'s paper. Now to

make the reduced form equations plug $h^* = \gamma w + \nu$ into equation (1), (3):

$$r = (\alpha_1 \gamma + \beta_1) w + (\alpha_1 \nu + \varepsilon_1) \quad (1)'$$

$$= \pi_1 w + u_1$$

$$h^s = (\alpha_3 \gamma + \beta_3) w + (\alpha_3 \nu + \varepsilon_3) \quad (3)'$$

$$= \pi_3 w + u_3$$

Bound said that by using the reduced form equation the estimates of α_1 and β_1 can be

written in terms of π 's and the covariance between u 's. If we assume that h^* is very

weakly correlated with w , β_3 's sign is determined by estimate of π_3 's sign in reduced

form because covariance $\sigma_{h^*,w}$ approaches to zero and then γ also approaches to zero.

By using h^s as a proxy for h^* , :

$$\text{plim } a_1 = \text{Cov}(u_1, u_3) / \text{Var}(u_3) = \text{Cov}(\alpha_1 \nu + \varepsilon_1, \alpha_3 \nu + \varepsilon_3) / \text{Var}(\alpha_3 \nu + \varepsilon_3)$$

$$= [\alpha_1 \alpha_3 \sigma_v^2 + \text{Cov}(\varepsilon_1, \varepsilon_3)] / (\alpha_3^2 \sigma_v^2 + \sigma_{\varepsilon_3}^2) = (\alpha_1 \alpha_3 \sigma_v^2 + \sigma_{\varepsilon_1} \sigma_{\varepsilon_3} \rho) / (\alpha_3^2 \sigma_v^2 + \sigma_{\varepsilon_3}^2)$$

by normalization restriction that α_3 is equal to 1

$$= (\alpha_1 \sigma_v^2 + \sigma_{\varepsilon_1} \sigma_{\varepsilon_3} \rho) / (\sigma_v^2 + \sigma_{\varepsilon_3}^2)$$

And then we use $\text{plim } a_1$ and normalization restriction and $\gamma \equiv \sigma_{h^*,w} / \sigma_w^2$.

$$\text{plim } B_1 = \text{plim } \pi_1 - \text{plim } \pi_3 \text{plim } a_1 = \beta_1 + \alpha_1 \gamma - \text{plim } a_1 \gamma - \text{plim } a_1 \beta_3$$

$$= \beta_1 + (\alpha_1 - \text{plim } a_1) (\sigma_{h^*,w} / \sigma_w^2) - \text{plim } a_1 \beta_3$$

II) Derivation of objective health and mortal probability proxy

$$r = \alpha_1 h^* + \beta_1 w + \varepsilon_1 \quad (1)$$

$$h^o = \alpha_4 l + \varepsilon_4 \quad (4)$$

$$l - pl = \alpha_5 d + \beta_5 w + \varepsilon_5 \quad (5)$$

$$h^* = l + d$$

1) l and d are uncorrelated with either ε 's or with each other.

2) ε_4 and ε_5 are uncorrelated with either ε_1 or with each other.

Because w and h^* are correlated with each other Bound suggest $l = \gamma_l w + \zeta$ and $d = \gamma_d w + \tau$ and $v = \zeta + \tau$. By assuming that ζ and τ are uncorrelated with either ε 's or with each other, $\gamma_l \equiv \sigma_{l,w} / \sigma_w^2$, and $\gamma_d \equiv \sigma_{d,w} / \sigma_w^2$. Now plug $l = \gamma_l w + \zeta$ and $d = \gamma_d w + \tau$ into equation (1), (4), (5) to make the reduced form.

$$r = \pi_1 w + u_1 \quad (1)'$$

$$h^o = \pi_4 w + u_4 \quad (4)'$$

$$l - pl = \pi_5 w + u_5 \quad (5)'$$

i) For objective health proxy, h^o :

$$\begin{aligned} \text{plim } a_1 &= \text{Cov}(u_1, u_4) / \text{Var}(u_4) = \text{Cov}(\alpha_1 v + \varepsilon_1, \alpha_4 \zeta + \varepsilon_4) / \text{Var}(\alpha_4 \zeta + \varepsilon_4) \\ &= \alpha_1 \alpha_4 \text{Cov}(v, \zeta) / (\alpha_4^2 \sigma_\zeta^2 + \sigma_{\varepsilon_4}^2) \text{ by normalization restriction that } \alpha_4 \text{ is equal to 1} \\ &= \alpha_1 \sigma_\zeta^2 / (\sigma_\zeta^2 + \sigma_{\varepsilon_4}^2) \end{aligned}$$

And then we use $\text{plim } a_1$ and normalization restriction and $\gamma_l \equiv \sigma_{l,w} / \sigma_w^2$.

$$\begin{aligned} \text{plim } B_1 &= \text{plim } \pi_1 - \text{plim } \pi_4 \text{plim } a_1 = \beta_1 + \alpha_1 \gamma - \text{plim } a_1 \gamma_l \\ &= \beta_1 + \alpha_1 (\sigma_{h^*,w} / \sigma_w^2) - \text{plim } a_1 (\sigma_{l,w} / \sigma_w^2) \end{aligned}$$

ii) For mortal probability proxy, l - pl :

$$\begin{aligned} \text{plim } a_1 &= \text{Cov}(u_1, u_5) / \text{Var}(u_5) = \text{Cov}(\alpha_1 v + \varepsilon_1, \alpha_5 \tau + \varepsilon_5) / \text{Var}(\alpha_5 \tau + \varepsilon_5) \\ &= \alpha_1 \alpha_5 \text{Cov}(v, \tau) / (\alpha_5^2 \sigma_\tau^2 + \sigma_{\varepsilon_5}^2) \text{ by normalization restriction that } \alpha_5 \text{ is equal to 1} \\ &= \alpha_1 \sigma_\tau^2 / (\sigma_\tau^2 + \sigma_{\varepsilon_5}^2) \end{aligned}$$

And then we use $\text{plim } a_1$ and normalization restriction and $\gamma_d \equiv \sigma_{d,w} / \sigma_w^2$

$$\begin{aligned} \text{plim } B_1 &= \text{plim } \pi_1 - \text{plim } \pi_5 \text{plim } a_1 = \beta_1 + \alpha_1 \gamma - \text{plim } a_1 \gamma_d - \text{plim } a_1 \beta_5 \\ &= \beta_1 + \alpha_1 (\sigma_{h^*,w} / \sigma_w^2) - \text{plim } a_1 \{ (\sigma_{d,w} / \sigma_w^2) + \beta_5 \} \end{aligned}$$

III) Derivation of using h^o and l - pl as instrumented variables for h^s :

When we apply 2SLS

$$\begin{aligned} \text{plim } a_{1IV} &= \text{Cov}(u_1, u_4) / \text{Cov}(u_3, u_4) \text{ or } \text{Cov}(u_1, u_5) / \text{Cov}(u_3, u_5) \\ &= \text{Cov}(\alpha_1 v + \varepsilon_1, \alpha_4 \zeta + \varepsilon_4) / \text{Cov}(\alpha_3 v + \varepsilon_3, \alpha_4 \zeta + \varepsilon_4) \\ &\quad \text{or } \text{Cov}(\alpha_1 v + \varepsilon_1, \alpha_5 \tau + \varepsilon_5) / \text{Cov}(\alpha_3 v + \varepsilon_3, \alpha_5 \tau + \varepsilon_5) \\ &= \alpha_1 \alpha_4 \text{Cov}(v, \zeta) / \alpha_3 \alpha_4 \text{Cov}(v, \zeta) \text{ or } \alpha_1 \alpha_5 \text{Cov}(v, \tau) / \alpha_3 \alpha_5 \text{Cov}(v, \tau) \\ &= \alpha_1 \quad \text{by normalization restriction that } \alpha_3 \text{ is equal to 1} \end{aligned}$$

$$plim B_{IV} = plim \pi_1 - plim \pi_3 plim a_{IV} = \beta_1 + \alpha_1 \gamma - plim a_{IV} \gamma - plim a_{IV} \beta_3$$

by normalization restriction that α_3 is equal to 1

$$= \beta_1 - \alpha_1 \beta_3$$

Appendix 2

Suppose subjective health, h^s , is binary endogenous variables.

$$r = 1[\alpha_1 h^s + \beta_1 z_1 + u_1 > 0]$$

$$h^s = 1[\beta_2 z + u_2 > 0] \quad z = (z_1, z_2)$$

where z_1 is all explanatory variables and z_2 is health problems or diseases problems as instrumental variables.

u_1, u_2 is independent of z and distributed as bivariate normal with zero mean and unit variance and $\rho_1 = \text{Corr}(u_1, u_2)$. If $\rho_1 \neq 0$, then u_1 and h^s are correlated and probit estimation is inconsistent for α_1 and β_1 . For conditional Maximum Likelihood, we need the joint distribution of (r, h^s) . Under joint normal distribution with unit variance, we can write $u_1 = \theta u_2 + e_1$ where e_1 is independent of z and u_2 , and also normally distributed with $E(e_1) = 0$ and $\text{Var}(e_1) = 1 - \rho_1^2$.

By the above process, we can write:

$$f(r, h^s | z) = f(r | h^s, z) f(h^s | z)$$

Since $u_2 = h^s - \beta_2 z$ and $r = 1[r^* > 0]$,

$$P(r = 1 | u_2, z) = \Phi[(\alpha_1 h^s + \beta_1 z_1 + \rho_1 u_2) / (1 - \rho_1^2)^{1/2}].$$

Since $h^s = 1$ only if $u_2 > -\beta_2 z$, we need a fact about the truncated normal distribution.

If u_2 has a standard normal distribution and is independent of z , then the density of u_2 given $u_2 > -\beta_2 z$ is

$$\phi(u_2) / P(u_2 > -\beta_2 z) = \phi(u_2) / \Phi(\beta_2 z)$$

Therefore,

$$\begin{aligned}
P(r=1 | h^s=1, z) &= E [P(r=1 | u_2, z) | h^s=1, z] \\
&= E \{ \Phi[\alpha_1 h^s + \beta_1 z_1 + \rho_1 u_2 / (1-\rho^2)^{1/2}] | h^s=1, z \} \\
&= \int_{-\beta_2 z}^{\infty} [1 / \Phi(\beta_2 z)] \Phi[(\alpha_1 h^s + \beta_1 z_1 + \rho_1 \psi^*_{2}) / (1-\rho^2)^{1/2}] \phi(\psi^*_{2}) d\psi^*_{2}
\end{aligned}$$

where ψ^*_{2} is the dummy argument of integration.

To simplify, we denote $\omega^* = (\alpha_1 h^s + \beta_1 z_1 + \rho_1 \psi^*_{2}) / (1-\rho^2)^{1/2}$.

By the same process,

$$P(r=0 | h^s=1, z) = \int_{-\infty}^{\infty} [1 - \Phi(\beta_2 z)] \Phi(\omega^*) \phi(\psi^*_{2}) d\psi^*_{2}$$

$$P(r=1 | h^s=0, z) = \int_{-\infty}^{-\beta_2 z} [1 - \Phi(\beta_2 z)] \Phi(\omega^*) \phi(\psi^*_{2}) d\psi^*_{2}$$

$$P(r=0 | h^s=0, z) = \int_{-\infty}^{-\beta_2 z} [1 - \Phi(\beta_2 z)] \Phi(\omega^*) \phi(\psi^*_{2}) d\psi^*_{2}$$

Therefore, $f(r, h^s | z) = f(r | h^s, z) f(h^s | z)$

$$= [P(r=1 | h^s=1, z) \times P(r=1 | h^s=0, z)]^r [P(r=0 | h^s=1, z) \times P(r=0 | h^s=0, z)]^{1-r}$$

$$= \int_{-\beta_2 z}^{\infty} [1 / \Phi(\beta_2 z)] \Phi(\omega^*) \phi(\psi^*_{2}) d\psi^*_{2} \times$$

$$\int_{-\infty}^{-\beta_2 z} [1 - \Phi(\beta_2 z)] \Phi(\omega^*) \phi(\psi^*_{2}) d\psi^*_{2} \}^r \times$$

$$\int_{-\beta_2 z}^{\infty} [1 - \Phi(\beta_2 z)] \Phi(\omega^*) \phi(\psi^*_{2}) d\psi^*_{2} \} \times$$

$$\int_{-\infty}^{-\beta_2 z} [1 - \Phi(\beta_2 z)] \Phi(\omega^*) \phi(\psi^*_{2}) d\psi^*_{2} \}^{1-r}$$

where ω^* depends on $\alpha_1, \beta_1, \rho_1, \beta_2$. By taking log gives the log likelihood function for Maximum Likelihood analysis.

Appendix 3

I) OLS estimate of β

Suppose $v^j_i = pl^j_i - pl_i$ ($i=75,85$ and $j = \text{wave } 1,2 \text{ or wave } 3,4$).

where v^j_i is measurement errors which are uncorrelated with pl_i and with each other.

Now suppose $pw62 = \alpha X + \beta pl^j_{75} - \beta v^j_{75} + \mu + \varepsilon_1$

where we assume that ε_1 is uncorrelated with v^j_{75} but v^j_{75} is correlated with pl^j_{75} .

$$\begin{aligned} \text{plim}_{BOLS} &= \beta + \{ \text{Cov}(pl^j_{75}, \varepsilon_1 + \mu - \beta v^j_{75}) / \text{Var}(pl^j_{75}) \} \\ &= \beta - \beta \{ \text{Cov}(pl^j_{75}, v^j_{75}) / \text{Var}(pl^j_{75}) \} \\ &= \beta \{ 1 - [\sigma^2_{vj75} / (\sigma^2_{vj75} + \sigma^2_{pl75})] \} \end{aligned}$$

In here, by assuming that ε_1 is uncorrelated with v^j_{75} but v^j_{75} is correlated with pl^j_{75}

and v^j_{75} is uncorrelated with pl_{75} $\text{Cov}(pl^j_{75}, \varepsilon_1 + \mu - \beta v^j_{75}) = -\beta \text{Cov}(pl^j_{75}, v^j_{75})$.

Also because pl^j_{75} and v^j_{75} must be correlated $\text{Cov}(pl^j_{75}, v^j_{75}) = \text{Cov}(pl_{75} + v^j_{75}, v^j_{75})$
 $= \sigma^2_{vj75}$ and $\text{Var}(pl^j_{75}) = \text{Var}(pl_{75} + v^j_{75}) = \sigma^2_{vj75} + \sigma^2_{pl75}$

II) Fixed-Effect estimate of β

If we do the first difference (Δ) to remove survey effect

$$pw62 - pw65 = \beta(pl^j_{75} - pl^j_{85}) - \beta(v^j_{75} - v^j_{85}) + (\varepsilon_1 - \varepsilon_2)$$

$$\text{or } \Delta pw = \beta \Delta pl^j - \beta \Delta v^j + \Delta \varepsilon$$

$$\begin{aligned} \text{plim}_{BFE} &= \beta + \{ \text{Cov}(\Delta pl^j, \Delta \varepsilon - \beta \Delta v^j) / \text{Var}(\Delta pl^j) \} \\ &= \beta - \beta \{ \text{Cov}(\Delta pl^j, \Delta v^j) / \text{Var}(\Delta pl^j) \} \\ &= \beta - 2\beta \{ [\sigma^2_{vj75} - \text{Cov}(v^j_{75}, v^j_{85})] / \text{Var}(\Delta pl^j) \} \\ &= \beta (1 - \{ \sigma^2_{vj75} (1 - \rho_v) / [\sigma^2_{vj75} (1 - \rho_v) + \sigma^2_{pl75} (1 - \rho_{pl})] \}) \end{aligned}$$

where $\rho_{pl} = \text{Corr}(pl_{75}, pl_{85})$ and $\rho_v = \text{Corr}(v_{75}^j, v_{85}^j)$.

To derive this formula we refer to Solon(1985) and we use $\text{Corr}(v_{75}^j, v_{85}^j) = \sigma_{vj75}^2 \rho_v$ and $\text{Var}(\Delta pl^j) = 2 [\sigma_{vj75}^2 (1 - \rho_v) + \sigma_{pl75}^2 (1 - \rho_{pl})]$. Because measurement error, v_i^j , is uncorrelated with each other, $\rho_v = 0$. Therefore, it is sure that measurement error of β is determined by correlation between true survival probabilities, $(1 - \rho_{pl})$.

Appendix 4 : Description of Variables

Variables	Description
<u>Dependent</u>	
Retire	1 if R consider themselves completely retired, or question not relevant to R (respondents)
<i>Pw62</i>	Probability of full-time work in age 62
<i>Pw65</i>	Probability of full-time work in age 65
<u>Health</u>	
Subjective Health (a)	1 if R currently reports bad health
<i>Pl75</i>	Survival probability live to age 75
<i>Pl85</i>	Survival probability live to age 85
Health Problems	Total number of health problems that R has
Asthma	1 if R currently has asthma
Back	1 if R currently has back problem
Legs	1 if R currently has problem with legs
Kidney	1 if R currently has kidney problem
Stomach	1 if R currently has stomach problem
Cholesterol	1 if R currently has high cholesterol
Breath	1 if R currently has breath problem
Dizziness	1 if R currently feels dizziness
Headache	1 if R currently has headache
Fatigue	1 if R currently feels tight fatigue
<u>Financial</u>	
Wage	Hourly wage
<u>Demographic</u>	
Age	Age of R during interview
Sex	1 if R is male
Race	1 if R is white
Education	Highest grade of R
Married	1 if R is married
<u>Retirement Plan</u>	
SS	1 if R receive Social Security
Federal HI	1 if R covered by federal government health insurance program, such as Medicare, Medicaid, CAMPUS, VA
Medicare	1 if R covered by Medicare
EPHI	1 if R has health insurance coverage from employer or union
Private HI	1 if R has private health insurance such as Medigap or from insurance company
Life Insurance	1 if R has life insurance
<u>Job/Requirement (b)</u>	
Union	1 if R joined into union
DB	1 if R has defined benefits
DC	1 if R has defined contributions
DB+DC	1 if R has both defined benefits and defined contributions
Physical Effort	1 if job require physical effort most or almost of the time
Heavy Lifting	1 if job require lifting heavy loads most or almost of the time
Stooping	1 if job require stooping or crouching most or almost of the time
Good Eyesight	1 if job require good eyesight most or almost of the time
Attention	1 if job require intense concentration most or almost of the time
People Skill	1 if job require skill in dealing with other people most or almost of the time
Computer	1 if job require working with computer most or almost of the time
Analyzing	1 if job require analyzing data or information most or almost of the time
Keep up with	1 if job require keeping up with the pace most or almost of the time
Repetition	1 if job require doing the same thing over and over again most or almost of the time
Learning	1 if job require learning new things most or almost of the time

(a) Respondents were asked "Would you say your health is excellent, very good, good, fair, or poor?" I categorized fair and poor as bad health to make dichotomous variable.

(b) Respondents were asked by 4 categories ; (1)all or almost of the time (2) most of the time (3) some of the time (4) none or almost none of the time. I categorized (1) and (2) as most or almost of the time to make dichotomous variable.

Bibliography

Aigner, D.J., (1973) "Regression With a Binary Independent Variable Subject to Errors of Observation," *Journal of Econometrics*, 1, 49-59.

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 Letter*, 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.

Ashenfelter, O., and A. Krueger, (1994) "Estimates of the Economic Return to Schooling from a New Sample of Twins," *American Economic Review*, Vol. 84, No. 5, 1157-1173.

Baker, M., and Mark Stabile, and Catherine Deri, (2001) "What Do Self-Reported, Objective, Measures of Health Measure?," NBER working paper, w8419.

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.

Bound, John, (1991) "Self-reported versus Objective Measures of Health in Retirement Models," *The Journal of Human Resources*, Vol.26: 106-138 Winter 91.

Bound, J., and A. Krueger, (1991), "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics*, Vol. 9, No.1, 1-24.

Chirikos, Thomas N., and Gilbert Nestel, (1981) "Impairment and labor Market Outcomes: A Cross-sectional and Longitudinal Analysis," in Hervert S. Parnes (ed.), *Work and Retirement: A Longitudinal Study of Men* (pp. 93-131). Cambridge, MA: MIT Press.

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.

Currie, J. and B.C. Madrian, (1999) "Health, Health Insurance and the Labor Market" in O. Ashenfelter and D. Cards (eds.), *Handbook of Labor Economics* v. 3c, New York: Elsevier, 3309-3416.

Dwyer, D. and O.S. Mitchell, (1999) "Health Problems as Determinants of Retirement: Are Self-Rated Measures Endogenous?," *Journal of Health Economics*, 18, 173-193.

Evans, W.N., and R.M. Schwab (1995), "Finishing High School and Starting College : Do Catholic Schools Make a Difference?" *Quarterly Journal of Economics* 110, 941-974.

Griliches, Z., and W. Mason, (1972) "Education, Income and Ability," *Journal of Political Economy*, Part II, 80, S74 – S103.

Gustman, Alan L. and Thomas L. Steinmeier, (2002) "Retirement Outcomes in The Health and Retirement Study," NBER w7588.

Harlow, S.D. and M.S. Linet, (1989) "Agreement Between Questionnaire Data and Medical Records," *American Journal of Epidemiology*, 129, 233-248.

Hurd, Michael and Kathleen McGarry, (1995) "Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study," *The Journal of Human Resources*, Vol. 30, S268 – S292.

Hyslop, Dean R. and Guido W. Imbens, (2000) "Bias From Classical and Other Forms of Measurement Error," NBER working paper, t0257.

Juster, Thomas and Richard Suzman, (1995) "An Overview of the Health and Retirement Study," *The Journal of Human Resources*, 30(supplement), S7 – S56.

Leamer, E., (1987), "Errors in Variables in Linear Systems," *Econometrica*, Vol.55, No. 5, 893-909.

Mathiowetz N.A. and Lair, T.J., (1994) "Getting Better? Change or Error in the Measurement of Functional Limitations," *Journal of Economic and Social Measurement*, 20, 237 – 262.

McGarry, Kathleen, (2002) "Health and Retirement: Do Changes in Health Affect Retirement Expectations?" NBER working paper, w9317.

Newey, Whitney K., (1987) "Efficient Estimation of Limited Dependent Models with Endogenous Explanatory Variables," *Journal of Econometrics*, 36:231-250, November.

Ruhm, Christopher J, (1994) "Do pensions increase the Labor Supply of older men?" NBER working paper, 4925.

Solon, G. (1985) "Comment on 'Benefits and Limitations of Panel Data' by C. Hsiao" *Econometric Reviews* 4, 183-186.