"The Effect of Social Security on the Income, Work and Mortality of Retirees"

Mark Votruba
Department of Economics
Case Western Reserve University
October 2002

Objectives

- Establish conditions under which Social Security MBA is an *exogenous* determinant of health or income.
- Estimate causal effect of MBA on mortality of retirees.
- Extrapolate implied marginal effects of work and income on mortality using IV techniques.

Motivation

- Large literature spanning several disciplines documents the correlation between (lower) income and (poorer) health
- Great interest in isolating the causal relationship of income on health
- The difficulties of estimating this relationship are well known:
  - Reverse causality ~ health shocks increase out-of-pocket health care and decrease capacity to earn, both reducing future income.
  - Omitted variables ~ both poor health and low income can result from high discount rate, that reduces investment in human capital (Fuchs, 1982; Farrell and Fuchs, 1982).

Outline of Talk

- Discuss identification strategy for study:
  - under what conditions is MBA an exogenous determinant of mortality/work/income?
  - what can/can't we learn about the effect of income on retirees' mortality using this strategy?
- Describe dataset used in study (NBS)
- Main analytic results

I. Identification Strategy

Two approaches to solving the inference problem (i.e. identify causal relationship among endogenously determined variables)

1 Structural approach:

- Assume particular relationship between the unobservable determinants of income and health
- Problem: estimates are dependent on assumptions one makes

2 Instrumental Variables (e.g. 2SLS) approach:

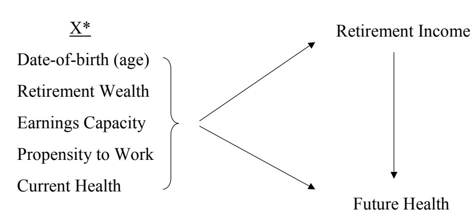
- Identify *exogenous* determinant of income, i.e. variable that affects income but is assumed to have no independent effect on health
- Problem: exogenous determinants of income are difficult to come by
- Promising source: administrative formulas used by public redistribution/benefit programs
- Example: Social Security benefit formula

What makes SS Benefit Formula Promising?

- Social Security comprises about 50% of retirees' total household income
- Problem: Social Security monthly benefit amount (MBA) is function of retirees' earnings history, retirement age and year of birth, all potentially correlated with unobserved determinants of health.
- Under what conditions is MBA a reasonable instrument for (or exogenous determinant of) income?

Conceptual Model 1:

Endogenous Income and Future Health



- We expect retirement income to be correlated with future health (mortality) because both are determined by the same underlying characteristics of the individual.
- If these underlying determinants could be measured precisely (X^*), OLS would provide unbiased estimation of the income effect:

$$\text{Mortality} = \beta X^* + \phi \text{ Income} + \varepsilon$$

- Measurement of X^* is generally imperfect (X_1):

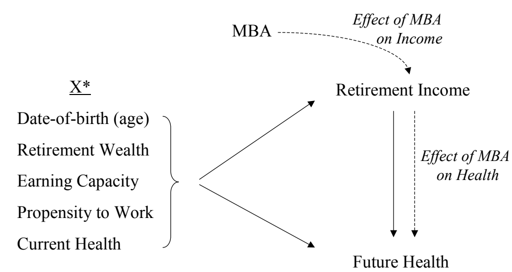
$$X^* = X_1 + \mu$$

- Results in biased estimation of ϕ if Income is correlated with μ:

$$\text{Mortality} = \beta X_1 + \phi \text{ Income} + (\beta \mu + \varepsilon)$$

Conceptual Model 2:

Exogenous MBA



- Conceptual Model 2 suggests that MBA is an exogenous determinant of Income, and MBA only affects Mortality through its effect on Income.
- Not true! MBA is a known function of Earnings History, Retirement Age and YOB, which we expect to be correlated with X*.
- However, *conditional* on a person's Earnings History, Retirement Age and date-of-birth, MBA *is* uncorrelated with X*.
- That is, if we appropriately control for a person's Earnings History, Retirement Age and date-of-birth, we can use OLS to generate unbiased estimates of the effect of MBA on Income and on Mortality.
- If MBA only affects Mortality *through* its effect on Income, we can back out the effect of Income on Mortality...

2SLS Estimation

X ~ vector of covariates, includes *flexible* controls for earnings history, retirement age and date-of-birth, as well as other variables potentially correlated with X*

D* ~ probability of person's dying by some time

I ~ income

Income determination model (first-stage):

(1) $I = b_0 X + b_3 \text{ MBA} + e_i$

Mortality determination model (second-stage):

(2) $D^* = a_0 X + a_1 I + e_d$

Substituting I from eq1 into eq2:

(2') $D^* = d_0 X + d_3 MBA + e'_d$ ~ reduced form

where $d_3 = a_1 * b_3$

(-) (+)

Implies that we can estimate effect of I on D* (a_1), by dividing the estimated effect of MBA on D* (d_3) by the estimated effect of MBA on I (b_3).

Also implies expected effect of MBA on D* is negative.

Critical assumption:

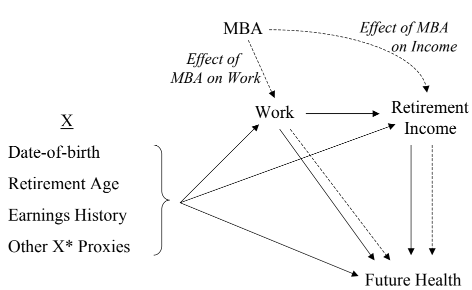
- Cov (MBA, e_d) =0
- i.e. conditional on Xs, MBA is uncorrelated with unobservable determinants of Mortality.

Associated implications:

- MBA must be uncorrelated with the unobservable determinants of Income.
- MBA only affects Mortality *through* Income.

Conceptual Model 3:

Exogenous MBA affects both Income and Work



Implies three equation model is needed...

Work determination model:

(3) $W = c_0 X + c_3 MBA + e_w$

Income determination model:

(4) $I = b_0 X + b_2 W + b_3 MBA + e_i$

Mortality determination model:

(5) $D^* = a_0 X + a_1 I + a_2 W + e_d$

(5') $D^* = d_0 X + d_3 MBA + e_d$ ~ reduced form

where $d_3 = a_1 * (b_2 * c_3 + b_3) + a_2 * c_3$

\+ + - + - -

- Model is "under-identified":
    - We *cannot* back out the effect of income and work from equation 5' -- we have one equation and two unknowns (a_1 and a_2).
    - We *can* estimate the (unbiased) effect of MBA on Mortality, but won't know how much of the effect is due to the Income effect and how much is due to the Work effect.
- Well-known rule about IV -- need at least as many instruments as endogenous variables.
- *Is there anything we can do?*

- Suppose MBA has different effect on the income and/or work of different types of recipients, T={1 , 2}.
- We can construct *two* instruments, MBA∗T, to satisfy the identification requirement.
- *Example:*

  If two types displayed similar effect on MBA on Mortality and Income, but different effect on Work, we would could conclude the Mortality effect was due to Income.
- Requires additional assumption:

  The effect of Income and Work on Mortality is the same across the two types.

II. MBA Formula

- Primary Insurance Amount (PIA) is size of MBA for those retiring (starting benefits) at exactly age 65.
- MBA of those retiring before age 65 decreases by 6.67% for every year of earlier retirement.
- MBA of those retiring after age 65 increases by 1% for every year of delayed retirement.
- Exception: Married recipients can receive MBA based own PIA or up to 50% of spouse's PIA (decreased for early retirement).

Historical Changes in PIA Calculation

- PIA originally calculated based on nominal covered earnings of the worker during work years.
- The use of nominal earnings led to gradual increases in PIA for later birth cohorts as result of wage inflation.
- In addition, Congress periodically adjusted the PIA formula upward to account for inflationary erosion of benefits over time.
- In the early 1970s, Congress implemented automatic COLA increases in PIA (based on the CPI) in response to high inflation at the time.

- Indexing PIA to inflation combined with the existing PIA increases due to nominal wage inflation had the effect of *double-indexing* benefits to inflation.
- During period of high inflation and low wage growth in the mid-1970s, Congress quickly saw their mistake: benefit projections suggested an impending financial crisis for the SS Trust Fund.
- Congress reacted (in 1977) by implementing a new PIA formula applying to those born after 1916.
- New formula was designed to reduce benefit amounts and eliminate double-indexing. PIAs would be function of *indexed* covered earnings rather than nominal.

- To ease the transition to the new formula, beneficiaries born between 1917-1921 (the "Notch Babies") could receive benefits under the new system or under a (more generous) "transition formula."
- Transition formula was similar to the old formula but excluded earnings after age 61 and received no COLA adjustments until after age 61.
- Led to substantial drop in benefits for those born immediately after the "notch" and gradual declines thereafter.

- PIA if born < 1916:
  - Concave and "kinked" function of mean nominal CEs calculated over highest *N* years.
  - *N* = YOB + 6

- PIA if born ≥ 1917:
  - Transition PIA - same as old, excluding CEs after age 61 and COLA increases between 1977 and age 62.
  - New PIA - concave and "kinked" function of mean indexed CEs calculated over higher *N* year.
  - Notch babies received higher of the two calculations.

III. Data Description

- New Beneficiary Survey (NBS): random stratified sample of 18,599 "new" social security recipients (includes both SS and SSI recipients).
- Respondents received first benefit between 6/80-5/81.
- Respondents surveys between 10/82-3/83.
- Survey data merged to corresponding SSA records of respondent and spouse (contains CE and benefit history, mortality dates).

Sample Selection Criteria

- Married respondents receiving benefit based on own CE record or CE record of spouse.
- Husbands born 1910-1919, wives born 1911-1919.
- Eliminated any receiving SSI, DI, welfare, worker's comp or unemployment.
- Eliminated any receiving military of federal pension.
- Dropped handful of cases where PIA and CE history were grossly inconsistent.
- Remaining sample: 3138 total couples
  2881 both retired (92%)

Table 1: Respondent Composition

		Respondent Sex		
		Male	*Female*	*Total*
Sp Rec'ing Benefit	*Yes*	1498 (92.9)	1383 (90.63)	2881
	No	114 (7.07)	143 (9.37)	257

Table 2: Individual Characteristics (Means)

	Both Receiving		**Total Sample**	
	Husband	*Wife*	*Husband*	*Wife*
Age	67.2	65.6	67.1	65.6
Black	0.04	--	0.04	--
Hispanic	0.02	--	0.02	--
Education	11.4	11.5	11.4	11.5
Has Pension	0.65	0.29	0.65	0.30
Ret-age	63.9	63.1	63.9	63.1
CE yrs>0	24.6	11.1	24.6	11.4
CE yrs=max	15.9	0.83	15.9	0.92
Asset Wealth	$118k	--	$120k	--
HH MBA (rec'd)	$951	--		
HH MBA (cal'c)	$926	--		

(No significant differences between "both receiving" and "one receiving" couples. Variables for marital history and job history not shown.)

Table 3: Mean (SD) Monthly Income Variables

	Both Receiving		**Total Sample**	
Total	1840	(1267)	1916	(1316)
SS	860	(273)	825	(297)
Asset	389	(373)	395	(747)
Pension	270	(432)	260	(428)
H Earnings	175	(483)	252	(691)
W Earnings	71	(327)	111	(410)
Other	76	(549)	77	(530)

Table 4: Predicted versus Recorded MBA

	TOTMBA	TOTMBAf	Difference
Mean	953.2	929.5	23.7
(s.e.)	(4.17)	(4.42)	(1.15)

Correlation Coefficient = 0.966

Note: Predicted MBA calculated using CEs through year benefit established. Increases due to (1) accumulation of additional CEs after "retirement," and (2) periods of non-receipt contributing to increases in PIA multiplier.

Figure 1: Predicted versus Recorded MBA

IV. Data Analysis

Demographic/Household Covariates:

- H/W Age and difference
- H/W Ret-Age
- H/W Time Since Benefits Started
- Black/Hispanic (dummies)
- H/W Education and difference
- H/W/B Pension (dummies)
- H/W Longest Job Chars (self-employed, professional)
- Yrs Married
- H/W Previously Widowed
- H/W Previously Married, not widowed
- Number Children Had
- Adult Child/Other Adult present (dummies)
- Asset Wealth

Note: third-order controls included for all continuous variables.

Covered Earnings Covariates:

- H/W Number Positive CE years (and difference)
- H/W Number Maximum CE years (and difference)
- H/W Mean Nominal CE (calculated over highest 22/18 years)
- Mean CE difference
- CE standard deviation (calculated over highest 26 years)
- H/W CE Dispersion (third-order interactions of Mean CE and SD)
- H/W Mean Indexed CEX (calculated over highest 22/19 years)
- Mean CEX difference
- CEX standard deviation (calculated over highest 26 years)
- H/W CEX Dispersion (third-order interactions of Mean CE and SD)

Note: Number of years used to calculate CE and CEX based on maximum fit of regressions on unearned, non-SS income. Various selection of years had little impact on later results.

Table 5A: Total Income Regressions

	OLS(1)	*OLS(2)*	*OLS(3)*	*OLS(4)*	*OLS(5)*
TotMBAf	1.72**	0.93**	0.52*	0.72	0.91*
(s.e.)	(.11)	(.15)	(.28)	(.47)	(.46)
Included Covars:					
Demographic/HH Chars		X	X	X	X
Pos Yrs/Max Yrs			X	X	X
Nominal CE				X	X
Indexed CEX					X
R-sqrd	.093	.352	.358	.369	.376

Notes: N=2881. Robust standard errors shown.

Table 5B: Log Income Regressions

	OLS(1)	*OLS(2)*	*OLS(3)*	*OLS(4)*	*OLS(5)*
ln(TotMBAf)	0.85**	0.54**	0.46**	0.48**	0.53**
(s.e.)	(.04)	(.05)	(.08)	(.19)	(.20)
Included Covars:					
Demographic/HH Chars		X	X	X	X
Pos Yrs/Max Yrs			X	X	X
Nominal CE				X	X
Indexed CEX					X
R-sqrd	.207	.481	.487	.496	.507

Notes: N=2881. Robust standard errors shown.

Table 6: Social Security Income Regressions

	Level Specification		Log Specification	
	OLS(1)	*OLS(2)*	*OLS(3)*	*OLS(4)*
TotMBAf	0.86**	0.84**		
(s.e.)	(.07)	(.04)		
ln(TotMBAf)			1.80**	0.92**
(s.e.)			(.35)	(.06)
Both Receiving		X		X
R-sqrd	.553	.801	.140	.737
N	2881	2700	2881	2700

Notes: Robust standard errors shown. Demographic/HH, Pos/Max Yrs, and CE covariates included in all regressions. Results robust to inclusion of CEX.

Table 7: Non-Social Security Income Regressions

	Total	Asset	Pension	H Earn	W Earn	Other
	OLS(1)	*OLS(2)*	*OLS(3)*	*OLS(4)*	*OLS(5)*	*OLS(6)*
Panel 1: Level Specification						
TotMBAf	-.15	-.31	-.11	-.19	.25*	.22
(s.e.)	(.47)	(.24)	(.17)	(.19)	(.15)	(.27)
R-sqrd	.309	.392	.341	.096	.076	.055
Panel 2: Log Specification						
ln(TotMBAf)	-.22	.01	.79	-2.70**	-.14	-.44
(s.e.)	(.72)	(.74)	(.94)	(1.17)	(.94)	(.73)
R-sqrd	.291	.519	.509	.144	.164	.056

Notes: N=2881. Robust standard errors shown. Demographic/HH, Pos/Max Yrs, and CE covariates included in all regressions. Results robust to inclusion of CEX.

Table 8: Probability Husband Working At Interview

	Level specification			Log specification		
	Logit(1)	*Logit(2)*	*Logit(3)*	*Logit(4)*	*Logit(5)*	*Logit(6)*
TotMBAf	-.27**	-.16	-.09			
(s.e.)	(.05)	(.12)	(.12)			
Marg Effect	*-.04*	*-.03*	*-.01*			
ln(TotMBAf)				-2.12**	-3.56**	-2.37*
(s.e.)				(.42)	(1.14)	(1.25)
Marg Effect				*-.04*	*-.06*	*-.04*
Included Covars:						
Nominal CE		X	X		X	X
Indexed CEX			X			X
Pseudo R-sqrd	.134	.148	.173	.134	.150	.174

Notes: N=2881. Robust standard errors shown. Demographic/HH and Pos/Max Yrs covariates included throughout. Mean dependent var = 0.28. Marginal effects represented implied mean effect of additional $100 of MBA.

Table 9: Probability Husband Working (CE>0) by Year

	1983	*1984*	*1985*	*1986 …*	*1989*	*1990*	*1991*
Panel 1: Logit Regression with CE Covariates							
ln(TotMBAf)	-2.67**	-3.76**	-2.97**	-3.69**	-3.19**	-1.79	-3.87**
(s.e.)	(1.11)	(1.15)	(1.20)	(1.24)	(1.39)	(1.50)	(1.74)
Marg Effect	*-.05*	*-.07*	*-.05*	*-.06*	*-.04*	*-.02*	*-.04*
Pseudo R-sqrd	.098	.100	.095	.091	.094	.111	.121
Panel 2: Logit Regression with CE & CEX Covariates							
ln(TotMBAf)	-1.60	-3.42**	-2.37*	-3.34**	-2.64*	-2.02	-3.71**
(s.e.)	(1.20)	(1.22)	(1.31)	(1.35)	(1.52)	(1.70)	(1.83)
Marg Effect	*-.03*	*-.06*	*-.04*	*-.05*	*-.04*	*-.02*	*-.04*
Pseudo R-sqrd	.121	.114	.113	.105	.121	.139	.141
Mean	.32	.29	.27	.24	.18	.16	.15
N	2816	2755	2655	2567	2293	2200	2093

Notes: Robust standard errors shown. Demographic/HH and Pos/Max Yrs covariates included throughout. Marginal effects represented implied mean effect of additional $100.

Table 10: Probability Husband Dies by Year

	1985	*1986*	*1987*	*1988*	*1989*	*1990*	*1991*
Panel 1: Logit Regression with CE Covariates							
ln(TotMBAf)	-0.06	2.92	2.82*	2.07	2.29*	1.03	0.51
(s.e.)	(2.44)	(1.96)	(1.60)	(1.44)	(1.34)	(1.24)	(1.16)
Marg Effect	*-.00*	*.02*	*.03*	*.03*	*.03*	*.02*	*.01*
Pseudo R-sqrd	.126	.088	.082	.068	.070	.074	.066
Panel 2: Logit Regression with CE & CEX Covariates							
ln(TotMBAf)	-1.43	2.20	2.90*	2.10	2.33	1.01	0.45
(s.e.)	(2.97)	(2.05)	(1.70)	(1.35)	(1.42)	(1.31)	(1.22)
Marg Effect	*-.01*	*.02*	*.03*	*.03*	*.03*	*-.02*	*.01*
Pseudo R-sqrd	.158	.107	.101	.088	.088	.091	.080
Mean	.04	.08	.11	.14	.17	.20	.24

Notes: N=2881. Robust standard errors shown. Demographic/HH and Pos/Max Yrs covariates included throughout. Marginal effects represented implied mean effect of additional $100.

Preliminary Conclusions

Despite imprecision of estimates, strong indication that increased MBA leads to increase in male mortality (at least in the short term).

Specifically, a 0.10 log point increase in MBA (about 10.5% or $100 at the mean) raises cumulative mortality by 3% points for the period from 5-7 years following the first interview.

Assuming no effect of income on mortality and extrapolating from the largest of the work effect estimates (7%), this suggests that work reduces mortality by 40 *percentage* points for those who are just at the margin of working and not working.

Seems highly unlikely!

Other possible explanations:

- Estimation imprecision
- Income has a negative effect on health of the elderly:
  - E.g. extra income funds unhealthy consumption
- Estimated effect of MBA on work severely underestimated:
  - Currently working ≠ ever work
  - Using CE>0 as indicator for work in year fails to count persons working in uncovered jobs (or underground jobs)

Model misspecified:

  - Linear controls of retirement age and covered earnings without appropriate interactions causing a spurious correlation.
- Selection bias in sample:
  - Lower MBA leads some spouses (healthy ones) to delay retirement.