Health and Heterogeneity

Josep Pijoan-Mas   José-Víctor Ríos-Rull

CEMFI, CEPR

Minnesota, Mpls Fed, CAERP

Chicago Fed, May 2013

Part I

Data

(The socio-economic gradient of longevity)

Mortality and life expectancy differences

- **Mortality rates** are strongly associated to **socio-economic status**

  Kitagawa and Hauser (1973); Elo and Preston (1996)

Mortality and life expectancy differences

- **Mortality rates** are strongly associated to **socio-economic status**

  Kitagawa and Hauser (1973); Elo and Preston (1996)

- Differences are large when aggregated into **life expectancies**

  Brown (2002); Lin et al (2003); Meara et al (2008)

Mortality and life expectancy differences

- **Mortality rates** are strongly associated to **socio-economic status**

  Kitagawa and Hauser (1973); Elo and Preston (1996)

- Differences are large when aggregated into **life expectancies**

  Brown (2002); Lin et al (2003); Meara et al (2008)

▷ However, these results present a **static picture** of the relationship between longevity and SES:

SES measures may and do change over time

Mortality and life expectancy differences

- **Mortality rates** are strongly associated to **socio-economic status**

  Kitagawa and Hauser (1973); Elo and Preston (1996)

- Differences are large when aggregated into **life expectancies**

  Brown (2002); Lin et al (2003); Meara et al (2008)

▷ However, these results present a **static picture** of the relationship between longevity and SES:

  SES measures may and do change over time

- Need to **develop a methodology** to compute expected longevity at age 50 conditional on a given socio-economic characteristic at age 50

Objective of the project

1. *Compute expected longevities* conditional on individual characteristics at age 50
   - Measure the importance of life-cycle changes of these characteristics for the longevity differentials at age 50

Objective of the project

1. *Compute expected longevities* conditional on individual characteristics at age 50
   - Measure the importance of life-cycle changes of these characteristics for the longevity differentials at age 50
2. *Decompose the longevity differentials* at age 50 into
   - health differences already present at 50
   - health evolution after 50
   - mortality differences not related to measured health

Objective of the project

1. *Compute expected longevities* conditional on individual characteristics at age 50
   - Measure the importance of life-cycle changes of these characteristics for the longevity differentials at age 50
2. *Decompose the longevity differentials* at age 50 into
   - health differences already present at 50
   - health evolution after 50
   - mortality differences not related to measured health

▷ Eventually, try to **understand the determinants** of this level of individual heterogeneity

The Health and Retirement Study

- Bi-annual panel, 10 waves, from 1992 to 2010
- Initial HRS cohort aged 50-61 in 1992 and 68-79 in 2010
- Two additional younger cohorts and two additional older cohorts
- This gives around 140,000 individual-year observations
  (white, aged 50-92, non-missing)
- Rich socio-economic data
  (marital status, education, income, wealth, labor market)
- Rich health-related data:
  - health stock: self-assessed and diagnostics
  - health investment: expenditures and behavior
  - mortality: keeps track of mortality

Methodology

- We use the HRS to compute expected longevities at age 50 conditional on different socio-economic charactersitcs $z \in Z \equiv \{z_1, z_2, ..., z_M\}$

Methodology

- We use the HRS to compute expected longevities at age 50 conditional on different socio-economic charactersitcs $z \in Z \equiv \{z_1, z_2, ..., z_M\}$
- We exploit the panel structure of the HRS to estimate:

Methodology

- We use the HRS to compute expected longevities at age 50 conditional on different socio-economic charactersitcs $z \in Z \equiv \{z_1, z_2, ..., z_M\}$
- We exploit the panel structure of the HRS to estimate:
  - Age-specific *survival rates* condtional on z

Methodology

- We use the HRS to compute expected longevities at age 50 conditional on different socio-economic charactersitcs $z \in Z \equiv \{z_1, z_2, ..., z_M\}$
- We exploit the panel structure of the HRS to estimate:
  - Age-specific *survival rates* condtional on z
  - Age-specific *transition probabilities* for z

Methodology

- We use the HRS to compute expected longevities at age 50 conditional on different socio-economic charactersitcs $z \in Z \equiv \{z_1, z_2, ..., z_M\}$
- We exploit the panel structure of the HRS to estimate:
  - Age-specific *survival rates* condtional on z
  - Age-specific *transition probabilities* for z
- Both mortality rates and transition matrices are estimated with parametric models

  Logit and multinomial logits with z-specific age terms

Methodology

- We use the HRS to compute expected longevities at age 50 conditional on different socio-economic charactersitcs $z \in Z \equiv \{z_1, z_2, ..., z_M\}$
- We exploit the panel structure of the HRS to estimate:
  - Age-specific *survival rates* condtional on z
  - Age-specific *transition probabilities* for z
- Both mortality rates and transition matrices are estimated with parametric models

  Logit and multinomial logits with z-specific age terms
- We link estimates of different cohorts to estimate expected longevities at age 50

Methodology

- We use the HRS to compute expected longevities at age 50 conditional on different socio-economic charactersitcs $z \in Z \equiv \{z_1, z_2, ..., z_M\}$
- We exploit the panel structure of the HRS to estimate:
  - Age-specific *survival rates* condtional on z
  - Age-specific *transition probabilities* for z
- Both mortality rates and transition matrices are estimated with parametric models

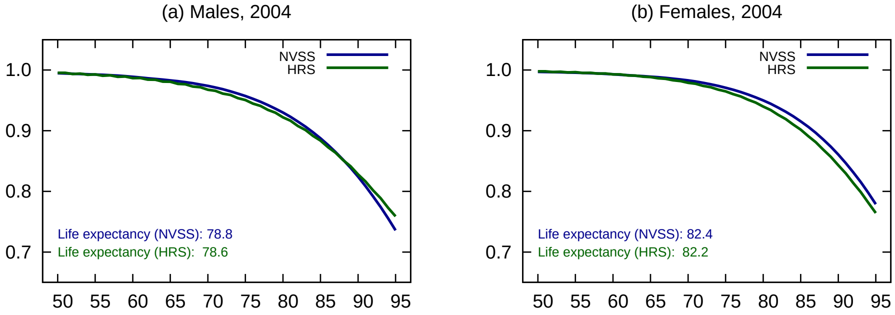
  Logit and multinomial logits with z-specific age terms
- We link estimates of different cohorts to estimate expected longevities at age 50
- We use data of all HRS years to increase sample size

Mortality rates and the National Vital Statistics System

We can compute life tables for 2004 and compare them to the NVSS

Mortality rates and the National Vital Statistics System

We can compute life tables for 2004 and compare them to the NVSS



Expected longevity at age 50

- Compute EL conditional on a characteristic $z \in Z$ at age 50:
  - Education: college *vs* high school dropout
  - Wealth: top *vs* bottom quintile
  - Labor market status: strongly attached *vs* inactive
  - Marital status: married *vs* non-married
  - Smoking: non-smoker *vs* smoker

Expected longevity at age 50

- Compute EL conditional on a characteristic $z \in Z$ at age 50:
  - Education: college *vs* high school dropout
  - Wealth: top *vs* bottom quintile
  - Labor market status: strongly attached *vs* inactive
  - Marital status: married *vs* non-married
  - Smoking: non-smoker *vs* smoker

- Estimate the following elements:
  - Survival Rates: $\gamma_i(z)$
  - Transition probabilities: $p_i(z'|z)$

Expected longevity at age 50

- Compute EL conditional on a characteristic $z \in Z$ at age 50:
  - Education: college *vs* high school dropout
  - Wealth: top *vs* bottom quintile
  - Labor market status: strongly attached *vs* inactive
  - Marital status: married *vs* non-married
  - Smoking: non-smoker *vs* smoker

- Estimate the following elements:
  - Survival Rates: $\gamma_i(z)$
  - Transition probabilities: $p_i(z'|z)$

- Then:

$$
\begin{aligned}
\ell_{50}(z_j) &= \sum_{i=50}^{92} i \sum_{z \in Z} [1 - \gamma_i(z)] \, x_i(z) + 1 \\
x_{i+1}(z') &= \sum_{z \in Z} p_i(z'|z) \;\; \gamma_i(z) \;\; x_i(z) \qquad \forall z' \in Z, \forall i \geq 50 \\
x_{50}(z_j) &= 1; \quad x_{50}(z) = 0 \;\; \forall z \neq z_j
\end{aligned}
$$

Expected longevities at age 50

	LE	Longevity differentials						
		edu	wea	lms	mar	smok	m-s	h
Expected Longevities								
Male	78.1	5.8	3.1	1.4	2.2	2.2	4.9	5.6
Female	81.8	5.8	2.6	0.7	1.2	1.8	2.9	4.7

Expected longevities at age 50

	LE	Longevity differentials						
		edu	wea	lms	mar	smok	m-s	h
Expected Longevities								
Male	78.1	5.8	3.1	1.4	2.2	2.2	4.9	5.6
Female	81.8	5.8	2.6	0.7	1.2	1.8	2.9	4.7

- We uncover a very large amount of heterogeneity

Expected longevities at age 50

	LE	Longevity differentials						
		edu	wea	lms	mar	smok	m-s	h
Expected Longevities								
Male	78.1	5.8	3.1	1.4	2.2	2.2	4.9	5.6
Female	81.8	5.8	2.6	0.7	1.2	1.8	2.9	4.7

- We uncover a very large amount of heterogeneity
- Less so for females (except for education)

Expected longevities at age 50

	LE	edu	wea	lms	mar	smok	m-s	h
		Longevity differentials						
Expected Longevities								
Male	78.1	5.8	3.1	1.4	2.2	2.2	4.9	5.6
Female	81.8	5.8	2.6	0.7	1.2	1.8	2.9	4.7
Life Expectancies								
Male	78.1	5.8	10.6	9.2	4.9	6.7	10.9	22.3
Female	81.8	5.8	9.3	6.7	3.1	5.2	7.2	20.3

- We uncover a very large amount of heterogeneity
- Less so for females (except for education)
- Life expectations computed only from cross-sections largely overstate the importance of the socio-economic conditions at age 50

Expected longevities at age 50

	LE	Longevity differentials						
		edu	wea	lms	mar	smok	m-s	h
Expected Longevities								
Male	78.1	5.8	3.1	1.4	2.2	2.2	4.9	5.6
Female	81.8	5.8	2.6	0.7	1.2	1.8	2.9	4.7
Life Expectancies								
Male	78.1	5.8	10.6	9.2	4.9	6.7	10.9	22.3
Female	81.8	5.8	9.3	6.7	3.1	5.2	7.2	20.3

- We uncover a very large amount of heterogeneity
- Less so for females (except for education)
- Life expectations computed only from cross-sections largely overstate the importance of the socio-economic conditions at age 50
- This tells us that there may be useful information contained in changes in characteristics z

Fine tunning

	Longevity differentials						
	edu	wea	lms	mar	smok	m-s	h
Male							
(1) All	5.8	3.1	1.4	2.2	2.2	4.9	5.6
(2) College graduates	—	2.0	0.9	1.5	1.3	3.5	5.2
(3) High school dropouts	—	2.6	1.5	2.6	2.5	5.7	4.8
Female							
(1) All	5.8	2.6	0.7	1.2	1.8	2.9	4.7
(2) College graduates	—	1.1	0.3	0.7	0.6	1.3	2.9
(3) High school dropouts	—	2.3	0.8	1.3	2.3	3.3	4.3

- Larger heterogeneity for the less educated

Fine tunning

	Longevity differentials						
	edu	wea	lms	mar	smok	m-s	h
Male							
(1) All	5.8	3.1	1.4	2.2	2.2	4.9	5.6
(2) College graduates	–	2.0	0.9	1.5	1.3	3.5	5.2
(3) High school dropouts	–	2.6	1.5	2.6	2.5	5.7	4.8
(4) Education and z	–	7.4	6.9	8.2	7.3	10.5	9.2
Female							
(1) All	5.8	2.6	0.7	1.2	1.8	2.9	4.7
(2) College graduates	–	1.1	0.3	0.7	0.6	1.3	2.9
(3) High school dropouts	–	2.3	0.8	1.3	2.3	3.3	4.3
(4) Education and z	–	7.1	6.1	7.1	7.5	8.3	8.8

- Larger heterogeneity for the less educated
- Gaps between education-z categories are enormous

Time trends

	LE	Longevity differentials						
		edu	wea	lms	mar	smok	m-s	h
Male								
1992	77.3	5.1	2.5	1.2	1.7	1.7	3.8	5.5
2010	79.1	6.4	4.7	1.9	2.8	3.7	6.7	5.5
Δ	+1.8	+1.3	+2.2	+0.7	+1.1	+2.0	+2.9	0.0
Δ_{NVSS}	+2.6							
Female								
1992	82.0	5.1	2.3	0.4	0.4	1.4	1.6	4.1
2010	81.5	7.1	3.7	1.1	2.1	2.4	4.6	5.6
Δ	-0.5	+2.0	+1.4	+0.7	+1.7	+1.0	+3.0	+1.5
Δ_{NVSS}	+1.4							

- Longevity differentials are increasing over time

Measuring health

- Most of these factors *per se* do not kill people but affect health, which in turn determines survival rates

Measuring health

- Most of these factors *per se* do not kill people but affect health, which in turn determines survival rates
- We observe **self-assessed health**

Measuring health

- Most of these factors *per se* do not kill people but affect health, which in turn determines survival rates
- We observe **self-assessed health**
  - In our data it is the best predictor of survival

Measuring health

- Most of these factors *per se* do not kill people but affect health, which in turn determines survival rates
- We observe **self-assessed health**
  - In our data it is the best predictor of survival
  - It is also so in the large epidemiological literature (*See Idler and Benyamini, 1997 and 1999*)

Measuring health

- Most of these factors *per se* do not kill people but affect health, which in turn determines survival rates
- We observe **self-assessed health**
  - In our data it is the best predictor of survival
  - It is also so in the large epidemiological literature (*See Idler and Benyamini, 1997 and 1999*)
  - Indeed, it makes education level uninformative in 2-year survivals

Measuring health

- Most of these factors *per se* do not kill people but affect health, which in turn determines survival rates
- We observe **self-assessed health**
  - In our data it is the best predictor of survival
  - It is also so in the large epidemiological literature (*See Idler and Benyamini, 1997 and 1999*)
  - Indeed, it makes education level uninformative in 2-year survivals
  - It is present in many surveys: HRS, PSID, NLSY, ...

Measuring health

- Most of these factors *per se* do not kill people but affect health, which in turn determines survival rates
- We observe **self-assessed health**
  - In our data it is the best predictor of survival
  - It is also so in the large epidemiological literature (*See Idler and Benyamini, 1997 and 1999*)
  - Indeed, it makes education level uninformative in 2-year survivals
  - It is present in many surveys: HRS, PSID, NLSY, ...
- Observing individual health, we can determine whether:

Measuring health

- Most of these factors *per se* do not kill people but affect health, which in turn determines survival rates
- We observe **self-assessed health**
  - In our data it is the best predictor of survival
  - It is also so in the large epidemiological literature (*See Idler and Benyamini, 1997 and 1999*)
  - Indeed, it makes education level uninformative in 2-year survivals
  - It is present in many surveys: HRS, PSID, NLSY, ...
- Observing individual health, we can determine whether:
  - Life expectancy heterogeneity is due to factors present at age 50

Measuring health

- Most of these factors *per se* do not kill people but affect health, which in turn determines survival rates
- We observe **self-assessed health**
  - In our data it is the best predictor of survival
  - It is also so in the large epidemiological literature (*See Idler and Benyamini, 1997 and 1999*)
  - Indeed, it makes education level uninformative in 2-year survivals
  - It is present in many surveys: HRS, PSID, NLSY, ...
- Observing individual health, we can determine whether:
  - Life expectancy heterogeneity is due to factors present at age 50
  - Or health conditions evolve differently for different people after age 50

Measuring health

- Most of these factors *per se* do not kill people but affect health, which in turn determines survival rates
- We observe **self-assessed health**
  - In our data it is the best predictor of survival
  - It is also so in the large epidemiological literature (*See Idler and Benyamini, 1997 and 1999*)
  - Indeed, it makes education level uninformative in 2-year survivals
  - It is present in many surveys: HRS, PSID, NLSY, ...
- Observing individual health, we can determine whether:
  - Life expectancy heterogeneity is due to factors present at age 50
  - Or health conditions evolve differently for different people after age 50
  - Or mortality rates are different even conditional on measured health

Expected Longevities: the role of self-rated health

How to build them

- Estimate the following elements:
  a) Initial distribution of health by characteristic z: $\varphi_{50}(h|z)$
  b) Joint health and z transitions: $p_i(z', h'|z, h)$
  c) z and health specific survival rates: $\gamma_i(z, h)$
- Then, compute:

$$\ell_{50}^{h}(z_j) = \sum_{i=50}^{92} i \sum_{h \in H, z \in Z} [1 - \gamma_i(z,h)]\ x_i(z,h) + 1$$

$$x_{i+1}(z', h') = \sum_{h \in H, z \in Z} p_i(z', h'|z, h)\ \gamma_i(z,h)\ x_i(z,h) \quad \forall z' \in Z,\ \forall h' \in H,\ \forall i$$

$$x_{50}(z_j, h) = \varphi_{50}(h|z_j) \text{ and } x_{50}(z,h) = 0 \ \ \forall z \neq z_j$$

Expected Longevities: the role of self-rated health

How to decompose them

- Is it initial health differences by socio-economic type?

$$\{\varphi_{50}(h|z), \quad p_i(h'|h), \quad \gamma_i(h)\}$$

- Is it type-specific health evolution?

$$\{\varphi_{50}(h), \quad p_i(z', h'|z, h), \quad \gamma_i(h)\}$$

- Is it type-specific mortality?

$$\{\varphi_{50}(h), \quad p_i(h'|h), \quad \gamma_i(z, h)\}$$

Expected Longevities: the role of self-rated health

Results

	EL	Longevity differentials					
		edu	wea	lms	mar	smok	m-s
Male							
All type-specific	78.2	6.0	3.7	3.3	2.4	2.9	6.0
(a) type-specific initial health		1.6	1.2	2.1	0.4	0.5	1.0
(b) type-specific transition		4.7	1.7	0.5	0.7	1.0	1.8
(c) type-specific mortality		0.0	0.9	0.4	1.4	1.4	3.3
Female							
All type-specific	81.8	5.9	3.6	1.3	1.4	2.3	3.4
(a) type-specific initial health		1.1	1.2	0.9	0.3	0.2	0.4
(b) type-specific transition		4.8	1.7	0.3	0.6	0.9	1.5
(c) type-specific mortality		0.3	0.8	0.2	0.4	1.2	1.6

Expected Longevities: the role of self-rated health

Education

- **Education** (and wealth)
  - 1/3 of gradient due to better health at age 50
  - 2/3 of gradient due to health-protection of education over life
  - Mortality rates independent of education once controlling for health

▷ More educated (and wealthy) experience a better evolution of health
  - More investment?
  - Better built?

▷ Education (and wealth) do little to help when bad conditions arise

Expected Longevities: the role of self-rated health

Marital status

- **Marital stauts** (and smoking)
  - small differences due to better health at age 50
  - larger health-protection of marital stauts over life
  - Mortality rates do depend on marital status once controlling for health: 2/3 of gradient for men

▷ What is it that helps survial of married over non-married when bad conditions arise?

(Recent widowhood kills you, but widows die less than other non-married)

Where is the advantage from education coming from?

- Our estimates say clear things
  - $\gamma^{i}(h)$ is independent of education
  - $\Gamma^{i,e}(h'|h)$ is NOT independent of education

Where is the advantage from education coming from?

- Our estimates say clear things
  - $\gamma^i(h)$ is independent of education
  - $\Gamma^{i,e}(h'|h)$ is NOT independent of education

- But still silent about the health-protection role of education:

Where is the advantage from education coming from?

- Our estimates say clear things
  - $\gamma^{i}(h)$ is independent of education
  - $\Gamma^{i,e}(h'|h)$ is NOT independent of education

- But still silent about the health-protection role of education:

a) Is it because educated *invest more* in their health?

  - If so, is it *expenditure* or *behavior*?
  - And is it because they are *richer* or because their *preferences* are different?

Where is the advantage from education coming from?

- Our estimates say clear things
  - $\gamma^{i}(h)$ is independent of education
  - $\Gamma^{i,e}(h'|h)$ is NOT independent of education

- But still silent about the health-protection role of education:

a) Is it because educated *invest more* in their health?
  - If so, is it *expenditure* or *behavior*?
  - And is it because they are *richer* or because their *preferences* are different?

b) Or rather, are educated intrinsically *better built*?

Where is the advantage from education coming from?

- Our estimates say clear things
  - $\gamma^{i}(h)$ is independent of education
  - $\Gamma^{i,e}(h'|h)$ is NOT independent of education

- But still silent about the health-protection role of education:
  a) Is it because educated *invest more* in their health?
  - If so, is it *expenditure* or *behavior*?
  - And is it because they are *richer* or because their *preferences* are different?

  b) Or rather, are educated intrinsically *better built*?

▷ Need **a model** to tease this mechanisms out

Where is the advantage from education coming from?

- Our estimates say clear things
  - $\gamma^{i}(h)$ is independent of education
  - $\Gamma^{i,e}(h'|h)$ is NOT independent of education

- But still silent about the health-protection role of education:
  a) Is it because educated *invest more* in their health?
  - If so, is it *expenditure* or *behavior*?
  - And is it because they are *richer* or because their *preferences* are different?

  b) Or rather, are educated intrinsically *better built*?

▷ Need **a model** to tease this mechanisms out

▷ More importantly, we need a way to identify the health production technology. We do not have it.

Conclusions

- There are large differences in education specific life expectancy.

Conclusions

- There are large differences in education specific life expectancy.
- These are associated to health as measured via self assessment.

Conclusions

- There are large differences in education specific life expectancy.
- These are associated to health as measured via self assessment.
- What matters is the health to health transition to which non-smoking and marriage contributes.

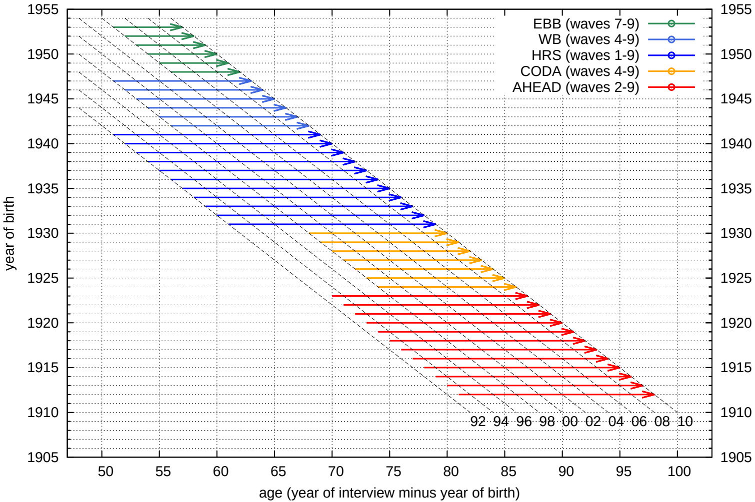
Conclusions

- There are large differences in education specific life expectancy.
- These are associated to health as measured via self assessment.
- What matters is the health to health transition to which non-smoking and marriage contributes.
- To identify the root of the advantages of education, we need to estimate rich models. Still a long way out.

Conclusions

- There are large differences in education specific life expectancy.
- These are associated to health as measured via self assessment.
- What matters is the health to health transition to which non-smoking and marriage contributes.
- To identify the root of the advantages of education, we need to estimate rich models. Still a long way out.
- However, we have some ideas of how to use the findings of this paper to learn important things about people.

HRS age-cohort structure



Survival probabilities

Health and education

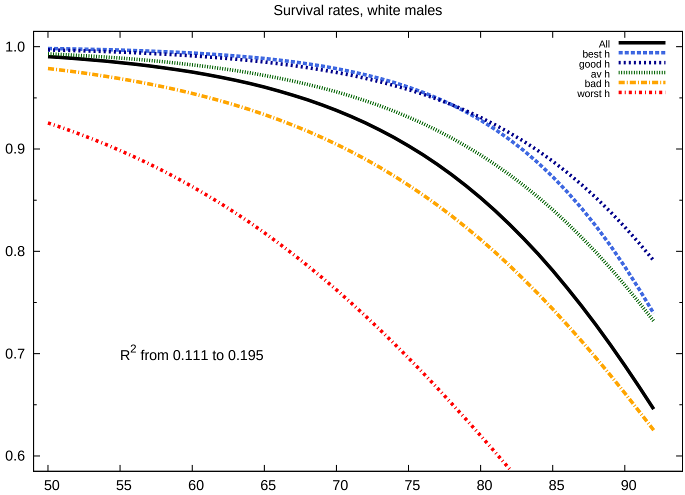
- **Education** is almost uninformative about two-year survival when we observe **health**:
  - Odds ratios for health much larger than for education
  - When health and education put together, education gives no advantage
  - LR test shows little value added by education

Logit regressions for survival (white males)

	Odds ratios				LR test	
	cg vs hsd		h2 vs h4		χ^2	p-value
	65	75	65	75		
only education	4.46	4.30			1897.89	0.000
only health			171.19	170.90	9.32	0.054
both together	1.08	1.16	202.29	202.02		

Survival probabilities

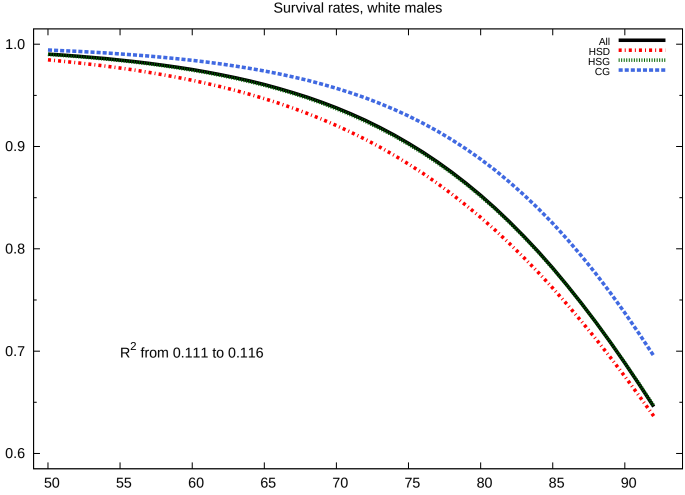
Survival by health groups



Source: HRS

Survival probabilities

Survival by education groups



Source: HRS

Survival probabilities

Survival by health and education groups