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
  

  ▶ 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 characteristics $z \in Z \equiv \{z_1, z_2, \ldots, z_M\}$.
Methodology

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

- We exploit the panel structure of the HRS to estimate:
  - Age-specific survival rates conditional 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.
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 conditional on $z$
Methodology

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

- We exploit the panel structure of the HRS to estimate:
  - Age-specific survival rates conditional 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 characteristics $z \in Z \equiv \{z_1, z_2, ..., z_M\}$

- We exploit the panel structure of the HRS to estimate:
  - Age-specific **survival rates** conditional 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 characteristics \( z \in Z \equiv \{ z_1, z_2, ..., z_M \} \)

- We exploit the panel structure of the HRS to estimate:
  - Age-specific survival rates conditional 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 characteristics $z \in Z \equiv \{z_1, z_2, \ldots, z_M\}$

- We exploit the panel structure of the HRS to estimate:
  - Age-specific **survival rates** conditional 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.

(a) Males, 2004

Life expectancy (NVSS): 78.8
Life expectancy (HRS): 78.6

(b) Females, 2004

Life expectancy (NVSS): 82.4
Life expectancy (HRS): 82.2
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:

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

$$x_{i+1}(z') = \sum_{z \in Z} p_i(z'|z) \gamma_i(z) x_i(z) \quad \forall z' \in Z, \forall i \geq 50$$

$$x_{50}(z_j) = 1; \quad x_{50}(z) = 0 \quad \forall z \neq z_j$$
## Expected longevities at age 50

<table>
<thead>
<tr>
<th>Expected Longevities</th>
<th>LE</th>
<th>edu</th>
<th>wea</th>
<th>lms</th>
<th>mar</th>
<th>smok</th>
<th>m-s</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>78.1</td>
<td>5.8</td>
<td>3.1</td>
<td>1.4</td>
<td>2.2</td>
<td>2.2</td>
<td>4.9</td>
<td>5.6</td>
</tr>
<tr>
<td>Female</td>
<td>81.8</td>
<td>5.8</td>
<td>2.6</td>
<td>0.7</td>
<td>1.2</td>
<td>1.8</td>
<td>2.9</td>
<td>4.7</td>
</tr>
</tbody>
</table>
**Expected longevities at age 50**

<table>
<thead>
<tr>
<th>Expected Longevities</th>
<th>LE</th>
<th>Longevity differentials</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>edu</td>
</tr>
<tr>
<td>Male</td>
<td>78.1</td>
<td>5.8</td>
</tr>
<tr>
<td>Female</td>
<td>81.8</td>
<td>5.8</td>
</tr>
</tbody>
</table>

- We uncover a very large amount of heterogeneity
## Expected longevities at age 50

<table>
<thead>
<tr>
<th>Expected Longevities</th>
<th>LE</th>
<th>edu</th>
<th>wea</th>
<th>lms</th>
<th>mar</th>
<th>smok</th>
<th>m-s</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>78.1</td>
<td>5.8</td>
<td>3.1</td>
<td>1.4</td>
<td>2.2</td>
<td>2.2</td>
<td>4.9</td>
<td>5.6</td>
</tr>
<tr>
<td>Female</td>
<td>81.8</td>
<td>5.8</td>
<td>2.6</td>
<td>0.7</td>
<td>1.2</td>
<td>1.8</td>
<td>2.9</td>
<td>4.7</td>
</tr>
</tbody>
</table>

- We uncover a very large amount of heterogeneity
- Less so for females (except for education)
Expected longevities at age 50

<table>
<thead>
<tr>
<th></th>
<th>LE</th>
<th>edu</th>
<th>wea</th>
<th>lms</th>
<th>mar</th>
<th>smok</th>
<th>m-s</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Longevities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>78.1</td>
<td>5.8</td>
<td>3.1</td>
<td>1.4</td>
<td>2.2</td>
<td>2.2</td>
<td>4.9</td>
<td>5.6</td>
</tr>
<tr>
<td>Female</td>
<td>81.8</td>
<td>5.8</td>
<td>2.6</td>
<td>0.7</td>
<td>1.2</td>
<td>1.8</td>
<td>2.9</td>
<td>4.7</td>
</tr>
<tr>
<td>Life Expectancies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>78.1</td>
<td>5.8</td>
<td>10.6</td>
<td>9.2</td>
<td>4.9</td>
<td>6.7</td>
<td>10.9</td>
<td>22.3</td>
</tr>
<tr>
<td>Female</td>
<td>81.8</td>
<td>5.8</td>
<td>9.3</td>
<td>6.7</td>
<td>3.1</td>
<td>5.2</td>
<td>7.2</td>
<td>20.3</td>
</tr>
</tbody>
</table>

- 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

<table>
<thead>
<tr>
<th>Expected Longevities</th>
<th>LE</th>
<th>edu</th>
<th>wea</th>
<th>lms</th>
<th>mar</th>
<th>smok</th>
<th>m-s</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>78.1</td>
<td>5.8</td>
<td>3.1</td>
<td>1.4</td>
<td>2.2</td>
<td>2.2</td>
<td>4.9</td>
<td>5.6</td>
</tr>
<tr>
<td>Female</td>
<td>81.8</td>
<td>5.8</td>
<td>2.6</td>
<td>0.7</td>
<td>1.2</td>
<td>1.8</td>
<td>2.9</td>
<td>4.7</td>
</tr>
<tr>
<td>Life Expectancies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>78.1</td>
<td>5.8</td>
<td>10.6</td>
<td>9.2</td>
<td>4.9</td>
<td>6.7</td>
<td>10.9</td>
<td>22.3</td>
</tr>
<tr>
<td>Female</td>
<td>81.8</td>
<td>5.8</td>
<td>9.3</td>
<td>6.7</td>
<td>3.1</td>
<td>5.2</td>
<td>7.2</td>
<td>20.3</td>
</tr>
</tbody>
</table>

- 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 tuning

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

- Larger heterogeneity for the less educated
## Fine tuning

<table>
<thead>
<tr>
<th></th>
<th>Longevity differentials</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>edu</td>
</tr>
<tr>
<td><strong>Male</strong></td>
<td></td>
</tr>
<tr>
<td>(1) All</td>
<td>5.8</td>
</tr>
<tr>
<td>(2) College graduates</td>
<td>—</td>
</tr>
<tr>
<td>(3) High school dropouts</td>
<td>—</td>
</tr>
<tr>
<td>(4) Education and $z$</td>
<td>—</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td></td>
</tr>
<tr>
<td>(1) All</td>
<td>5.8</td>
</tr>
<tr>
<td>(2) College graduates</td>
<td>—</td>
</tr>
<tr>
<td>(3) High school dropouts</td>
<td>—</td>
</tr>
<tr>
<td>(4) Education and $z$</td>
<td>—</td>
</tr>
</tbody>
</table>

- Larger heterogeneity for the less educated
- Gaps between education-$z$ categories are enormous
## Time trends

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

- 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 \textit{per se} do not kill people but affect health, which in turn determines survival rates

- We observe \textbf{self-assessed health}
  
  - In our data it is the best predictor of survival
  - It is also so in the large epidemiological literature \textit{(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:

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

$$x_{i+1}(z', h') = \sum_{h \in H, z \in Z} p_i(z', h'|z, h) \gamma_i(z, h) x_i(z, h) \quad \forall z' \in Z, \forall h' \in H, \forall i \geq 50$$

$$x_{50}(z_j, h) = \varphi_{50}(h|z_j) \text{ and } x_{50}(z, h) = 0 \quad \forall z \neq z_j$$
Expected Longevities: the role of self-rated health

How to decompose them

- Is it initial health differences by socio-economic type?
  \[ \{ \varphi_{50}(h|z), p_i(h'|h), \gamma_i(h) \} \]

- Is it type-specific health evolution?
  \[ \{ \varphi_{50}(h), p_i(z', h'|z, h), \gamma_i(h) \} \]

- Is it type-specific mortality?
  \[ \{ \varphi_{50}(h), p_i(h'|h), \gamma_i(z, h) \} \]
## Expected Longevities: the role of self-rated health

### Results

<table>
<thead>
<tr>
<th></th>
<th>EL</th>
<th>Longevity differentials</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>edu</td>
</tr>
<tr>
<td><strong>Male</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All type-specific</td>
<td>78.2</td>
<td>6.0</td>
</tr>
<tr>
<td>(a) type-specific initial health</td>
<td>1.6</td>
<td>1.2</td>
</tr>
<tr>
<td>(b) type-specific transition</td>
<td>4.7</td>
<td>1.7</td>
</tr>
<tr>
<td>(c) type-specific mortality</td>
<td>0.0</td>
<td>0.9</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All type-specific</td>
<td>81.8</td>
<td>5.9</td>
</tr>
<tr>
<td>(a) type-specific initial health</td>
<td>1.1</td>
<td>1.2</td>
</tr>
<tr>
<td>(b) type-specific transition</td>
<td>4.8</td>
<td>1.7</td>
</tr>
<tr>
<td>(c) type-specific mortality</td>
<td>0.3</td>
<td>0.8</td>
</tr>
</tbody>
</table>
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 status (and smoking)
  - small differences due to better health at age 50
  - larger health-protection of marital status over life
  - Mortality rates do depend on marital status once controlling for health:
    2/3 of gradient for men

- What is it that helps survival 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

<table>
<thead>
<tr>
<th>Logit regressions for survival (white males)</th>
<th>Odds ratios</th>
<th>LR test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cg vs hsd h2 vs h4</td>
<td>χ²</td>
</tr>
<tr>
<td>only education</td>
<td>65 75</td>
<td>65 75</td>
</tr>
<tr>
<td>only health</td>
<td>4.46 4.30</td>
<td>171.19 170.90</td>
</tr>
<tr>
<td>both together</td>
<td>1.08 1.16</td>
<td>202.29 202.02</td>
</tr>
</tbody>
</table>
Survival probabilities

Survival by health groups

Survival rates, white males

R² from 0.111 to 0.195

Source: HRS
Survival probabilities

Survival by education groups

Survival rates, white males

$R^2$ from 0.111 to 0.116

Source: HRS

Josep Pijoan-Mas, José-Víctor Ríos-Rull

Health and Heterogeneity
Survival probabilities

Survival by health and education groups

(a) all health categories

(b) top health

(c) average health

(d) worst health

Source: HRS