What if international comparisons of health expectancies and socio-economic disparities do depend on the choice of the survey used? An approach for consolidation

Erasmus MC, Department of Public Health

Jose Ricardo Rubio
Wilma Nusselder
Johan Mackenbach
Health Expectancy

- Population based indicator that incorporates **mortality** and **health status** measures in a single statistic

- Health expectancy adds information about health status to life expectancy measure

- **Average Health Status** – requires a multi-attribute description of health states for a **representative sample** of the population
  - Global Activity Limitation Index (GALI)
  - Self-reported health (SRH)

- Health status generally derived from survey data
Measuring health status - Data

- No gold standard survey
- Several potential candidate surveys
  - EU-SILC (2008/ 2012)
  - European Social Survey (2008 / 2010 / 2012)
  - EHIS wave 1 (2008)
  - SHARE Wave 4 (2011)

- If nationally representative and adequately used (i.e sampling design, sampling error, non-response error) – unbiased estimate

- Issues of small numbers when stratifying by sex, age and SES (education, occupation, income) - Particularly at older ages
Measuring GALI Prevalence – An example

Prevalence of Disability by Survey (2008-2012) - GALI
Spain - Males

Prevalence (0-1)
Age Group

with 95% confidence interval
Measuring GALI Prevalence – An example

Prevalence of Disability by Survey (2008-2012) - GALI
Spain - Females

with 95% confidence interval
Objectives of the Analysis

1) Quantify the **difference in prevalence** of GALI disability **across different surveys** incorporating information from 7 countries and 11 age groups (30-84 years old)

2) Quantify the **educational disparities** in GALI disability **across low and high educated**

3) Quantify the **differences in educational disparities** in GALI disability when using one survey over another.

4) **Predict GALI disability** across age groups (and educational attainment) – An Example (Spain)
Summary Statistics

Number of Respondents by Survey and Country

<table>
<thead>
<tr>
<th>Survey</th>
<th>Belgium</th>
<th>Czech Republic</th>
<th>Estonia</th>
<th>Spain</th>
<th>France</th>
<th>Hungary</th>
<th>Poland</th>
<th>Slovenia</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>EHISw1</td>
<td>8,477</td>
<td>1,918</td>
<td>6,398</td>
<td>21,437</td>
<td>23,537</td>
<td>4,963</td>
<td>34,562</td>
<td>2,102</td>
<td>103,394</td>
</tr>
<tr>
<td>ESS(08/10/12)</td>
<td>5,231</td>
<td>6,334</td>
<td>5,721</td>
<td>6,211</td>
<td>5,606</td>
<td>5,026</td>
<td>5,199</td>
<td>3,858</td>
<td>43,186</td>
</tr>
<tr>
<td>EU-SILC (08/12)</td>
<td>23,346</td>
<td>40,064</td>
<td>22,536</td>
<td>58,292</td>
<td>42,756</td>
<td>42,524</td>
<td>60,473</td>
<td>49,008</td>
<td>338,999</td>
</tr>
<tr>
<td>SHAREw4</td>
<td>4,993</td>
<td>5,269</td>
<td>6,550</td>
<td>3,277</td>
<td>5,459</td>
<td>2,975</td>
<td>1,654</td>
<td>2,650</td>
<td>32,827</td>
</tr>
<tr>
<td>Total</td>
<td>42,047</td>
<td>53,585</td>
<td>41,205</td>
<td>89,217</td>
<td>77,358</td>
<td>55,488</td>
<td>101,888</td>
<td>57,618</td>
<td>518,406</td>
</tr>
</tbody>
</table>

Summary Statistics for Relevant Variables

<table>
<thead>
<tr>
<th>Education</th>
<th>Freq</th>
<th>%</th>
<th>Cum %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (ISCED I-II)</td>
<td>178,273</td>
<td>34.85</td>
<td>34.85</td>
</tr>
<tr>
<td>Medium (ISCED IIIab-IV)</td>
<td>235,020</td>
<td>45.94</td>
<td>80.79</td>
</tr>
<tr>
<td>High (ISCED V1-V2)</td>
<td>98,259</td>
<td>19.21</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GALI (moderate + severe)</th>
<th>Freq</th>
<th>%</th>
<th>Cum %</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>339,181</td>
<td>70.97</td>
<td>70.97</td>
</tr>
<tr>
<td>Yes</td>
<td>138,736</td>
<td>29.03</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sex</th>
<th>Freq</th>
<th>%</th>
<th>Cum %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>242,173</td>
<td>46.72</td>
<td>46.72</td>
</tr>
<tr>
<td>Female</td>
<td>276,218</td>
<td>53.28</td>
<td>100</td>
</tr>
</tbody>
</table>
Countries Included in the Analysis

Graphs by Country
Methodological Approach

\[ \text{GALI}_{ic} = \alpha + \beta_k \text{AgeGroup}_k + \beta_s \text{Survey}_s + \beta_c \text{Country}_c + \beta_e \text{Education} + \beta_i \text{Education} \times \text{Survey} + \epsilon \]

\( \text{GALI}_{ic} \) is an indicator variable for whether individual \( i \) in country \( c \) reports moderate or severe disability.

\( \text{AgeGroup}_k \) is an ordered categorical (11 levels) variable for age groups (30-34; 35-39; 40-44;\ldots; 80-84). 85+ are excluded.

\( \text{Survey}_s \) is a categorical variable that denotes the survey of individual \( i \) (4 levels).

\( \text{Country}_c \) is a categorical variable that denotes the country of individual \( i \) (7 lvl.).

\( \text{Education} \times \text{Survey} \) is an interaction term between education and survey.

\( \epsilon \) is the standard error term.
Methodological Approach

\[ \text{GALL}_{ic} = \alpha + \beta_k \text{AgeGroup}_k + \beta_s \text{Survey}_s + \beta_c \text{Country}_c + \beta_e \text{Education} + \beta_i \text{Education} \ast \text{Survey} + \varepsilon \]

**Model 1** – Linear Probability Model (OLS)

**Model 2** – Logistic Regression

Both models include **robust standard errors**

Stratified by gender

*Coefficients of Interest*
### 1) Difference in prevalence of GALI disability across different surveys

\[ \text{GALI}_{ic} = \alpha + \beta_k \text{AgeGroup}_k + \beta_s \text{Survey}_s + \beta_c \text{Country}_c + \beta_e \text{Education} + \beta_i \text{Education} \ast \text{Survey} + \varepsilon \]

**Model 1 - OLS**

<table>
<thead>
<tr>
<th>Survey</th>
<th>Coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU-SILC (10/12)</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td>EHISw1 (08)</td>
<td>0.09***</td>
<td>[0.01]</td>
</tr>
<tr>
<td>ESS (08/10/12)</td>
<td>-0.04***</td>
<td>[0.01]</td>
</tr>
<tr>
<td>SHAREw4 (2011)</td>
<td>0.06***</td>
<td>[0.01]</td>
</tr>
</tbody>
</table>

**Model 2 – Logistic Regression**

<table>
<thead>
<tr>
<th>Survey</th>
<th>Coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU-SILC (10/12)</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td>EHISw1 (08)</td>
<td>0.10***</td>
<td>[0.00]</td>
</tr>
<tr>
<td>ESS (08/10/12)</td>
<td>-0.02***</td>
<td>[0.00]</td>
</tr>
<tr>
<td>SHAREw4 (2011)</td>
<td>0.07***</td>
<td>[0.01]</td>
</tr>
</tbody>
</table>

*significant at 10% level  
** significant at 5% level  
*** significant at 1% level

**Males**  
\[ N = 175,303 \]
1) Difference in prevalence of GALI disability across different surveys

\[ \text{GALI}_{ic} = \alpha + \beta_k \text{AgeGroup}_k + \beta_s \text{Survey}_s + \beta_c \text{Country}_c + \beta_e \text{Education} + \beta_i \text{Education} \ast \text{Survey} + \varepsilon \]

<table>
<thead>
<tr>
<th>Model 1 - OLS</th>
<th>Model 2 – Logistic Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU-SILC (10/12)</td>
<td>Ref</td>
</tr>
<tr>
<td>EHISw1 (08)</td>
<td>0.12*** [0.01]</td>
</tr>
<tr>
<td>ESS (08/10/12)</td>
<td>-0.02*** [0.01]</td>
</tr>
<tr>
<td>SHAREw4 (2011)</td>
<td>0.08*** [0.01]</td>
</tr>
</tbody>
</table>

Females

\[ N = 211,320 \]

*significant at 10% level
** significant at 5% level
*** significant at 1% level
2) Quantify the educational disparities in GALI disability across low and high educated

\[ \text{GALI}_{ic} = \alpha + \beta_k \text{AgeGroup}_k + \beta_s \text{Survey}_s + \beta_c \text{Country}_c + \beta_e \text{Education} + \beta_i \text{Education} \times \text{Survey} + \epsilon \]

<table>
<thead>
<tr>
<th>Model 1 - OLS</th>
<th>Model 2 – Logistic Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Educated (ISCED I-II)</td>
<td>Ref</td>
</tr>
<tr>
<td>Medium Educated (ISCED III-IV)</td>
<td>-0.09*** [0.00]</td>
</tr>
<tr>
<td>High Educated (ISCED V-VI)</td>
<td>-0.15*** [0.00]</td>
</tr>
<tr>
<td>Low Educated (ISCED I-II)</td>
<td>Ref</td>
</tr>
<tr>
<td>Medium Educated (ISCED III-IV)</td>
<td>-0.09*** [0.00]</td>
</tr>
<tr>
<td>High Educated (ISCED V-VI)</td>
<td>-0.16*** [0.00]</td>
</tr>
</tbody>
</table>

*significant at 10% level  
** significant at 5% level  
*** significant at 1% level

Males
N= 175,303
2) Quantify the educational disparities in GALI disability across low and high educated

\[ GALI_{ic} = \alpha + \beta_k AgeGroup_k + \beta_s Survey_s + \beta_c Country_c + \beta_e Education + \beta_i Education \ast Survey + \epsilon \]

<table>
<thead>
<tr>
<th>Model 1 - OLS</th>
<th></th>
<th>Model 2 – Logistic Regression</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Educated</td>
<td>Ref</td>
<td>Low Educated (ISCED I-II)</td>
<td>Ref</td>
</tr>
<tr>
<td>(ISCED I-II)</td>
<td></td>
<td>Medium Educated (ISCED III-IV)</td>
<td>-0.09*** [0.00]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Highly Educated (ISCED V-VI)</td>
<td>-0.17*** [0.00]</td>
</tr>
<tr>
<td></td>
<td>-0.10*** [0.00]</td>
<td>Medium Educated (ISCED III-IV)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.16*** [0.00]</td>
<td>Highly Educated (ISCED V-VI)</td>
<td></td>
</tr>
</tbody>
</table>

*significant at 10% level  
** significant at 5% level  
*** significant at 1% level

Females  
N = 211,320
3) Quantify the differential in educational disparities in GALI disability across surveys

\[ \text{GALI}_{ic} = \alpha + \beta_k \text{AgeGroup}_k + \beta_s \text{Survey}_s + \beta_c \text{Country}_c \]
\[ + \beta_e \text{Education} + \beta_i \text{Education} \times \text{Survey} + \varepsilon \]
3) Quantify the differential in educational disparities in GALI disability across surveys

\[ GALI_{ic} = \alpha + \beta_k \text{AgeGroup}_k + \beta_s \text{Survey}_s + \beta_c \text{Country}_c + \beta_e \text{Education} + \beta_i \text{Education} \times \text{Survey} + \epsilon \]

Estimated Educational Difference in GALI Disability by Survey (2008-2012)

Females

Prevalence Difference (High-Low Educated)

with 95\% confidence interval

\( \text{ci}_u / \text{ci}_l \) • estimate
4) Predict GALI disability across age groups (and educational attainment) – An Example

Prevalence of Disability by Survey (2008-2012) - GALI
Spain - Males

with 95% confidence interval
4) Predict GALI disability across age groups (and educational attainment) – An Example (Spain)

**Prevalence of Disability by Survey (2008-2012) - GALI**

Spain - Females

- Prevalence (0-1)
- agecat5

Legend:
- **EU SILC**
- **ESS**
- **EHIS W1**
- **SHARE W3**
- **Estimated**

with 95% confidence interval
4) Predictions by Educational Status

Estimated Prevalence of Disability by Survey (2008-2012) - GALI
Spain - Males

Prevalence (0-1)

agecat

Low Educated  Medium Educated  Highly Educated

with 95% confidence interval
4) Predictions by Educational Status

Estimated Prevalence of Disability by Survey (2008-2012) - GALI
Spain - Females

Prevalence (0-1)

agecat

with 95% confidence interval

Low Educated | Medium Educated | Highly Educated
Limitations

- Further robustness checks are required
- **Goodness of fit** needs to be assessed and compared to other models
- Use of Wave 4 of SHARE means individuals have dropped out of the sample
- Weights have not been included in regression based analyses as they are not provided for all countries included
- Further interaction terms should be explored (survey*country ; age*country)
- Other GLMs might be more suitable for the task (i.e binomial log link function)
- Inference modelling might not be optimal for predicting
Conclusion

1) Very little data in last age group makes it hard to include it. Further complicated by stratification.

2) Holding other covariates constant, the surveys have statistically significant differences when measuring GALI disability.

3) There are non-trivial inequalities in GALI disability across high and low educated individuals.

4) These inequalities are influenced by the choice of survey. The degree to which they impact HLE inequalities is to be determined.

5) Further refinement of the modelling strategy is desirable for obtaining estimates particularly by educational level.
Data References

- **EUSILC UDB 2008 – version 7 of March 2015**
- **EUSILC UDB 2012 – version 1 of January 2016**
- ESS Round 5: European Social Survey Round 5 Data (2010). Data file edition 3.2. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- ESS Round 6: European Social Survey Round 6 Data (2012). Data file edition 2.2. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- EHIS Wave 1 2006/09

The responsibility for all conclusions drawn from the data lies entirely with the author(s)
Prevalence of Disability by Survey (2008-2012) - GALI

Spain - Males - Low Educated

Prevalence (0-1)

agecat

30-34  35-39  40-44  45-49  50-54  55-59  60-64  65-69  70-74  75-79  80-84  85+

with 95% confidence interval
Prevalence of Disability by Survey (2008-2012) - GALI
Spain - Males - Medium Educated

Prevalence (0-1) vs agecat

- EU SILC
- ESS
- EHIS W1
- SHARE W3
- Estimated

with 95% confidence interval