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

28th REVES Meeting, Vienna 8-10 June 2016

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

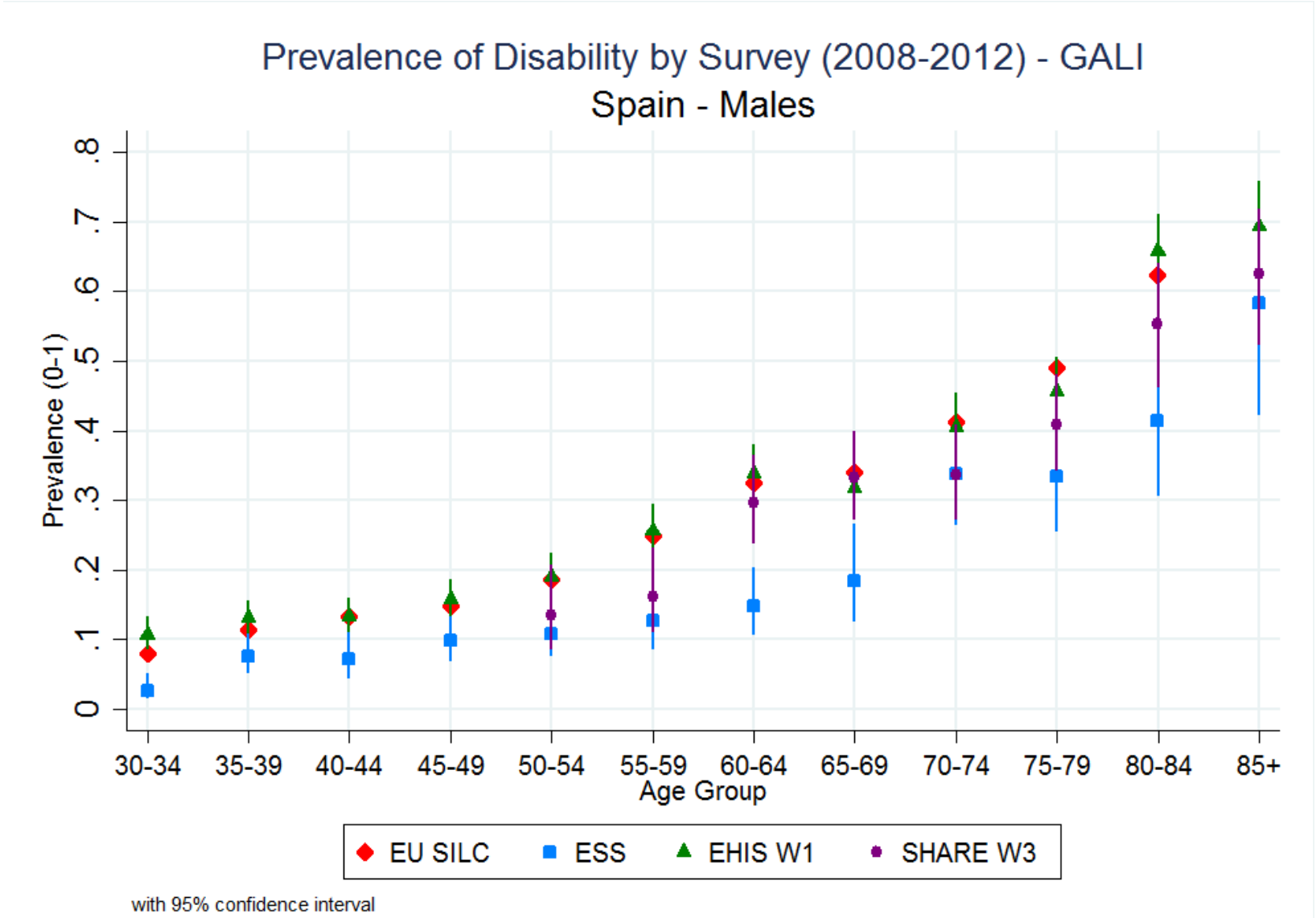
- **Average Health Status –** requires a multi-attribute description of health states for a **representative sample** of the population
  - Global Activity Limitation Index (GALI)
  - Self-reported health (SRH)

- Health status generally derived from survey data

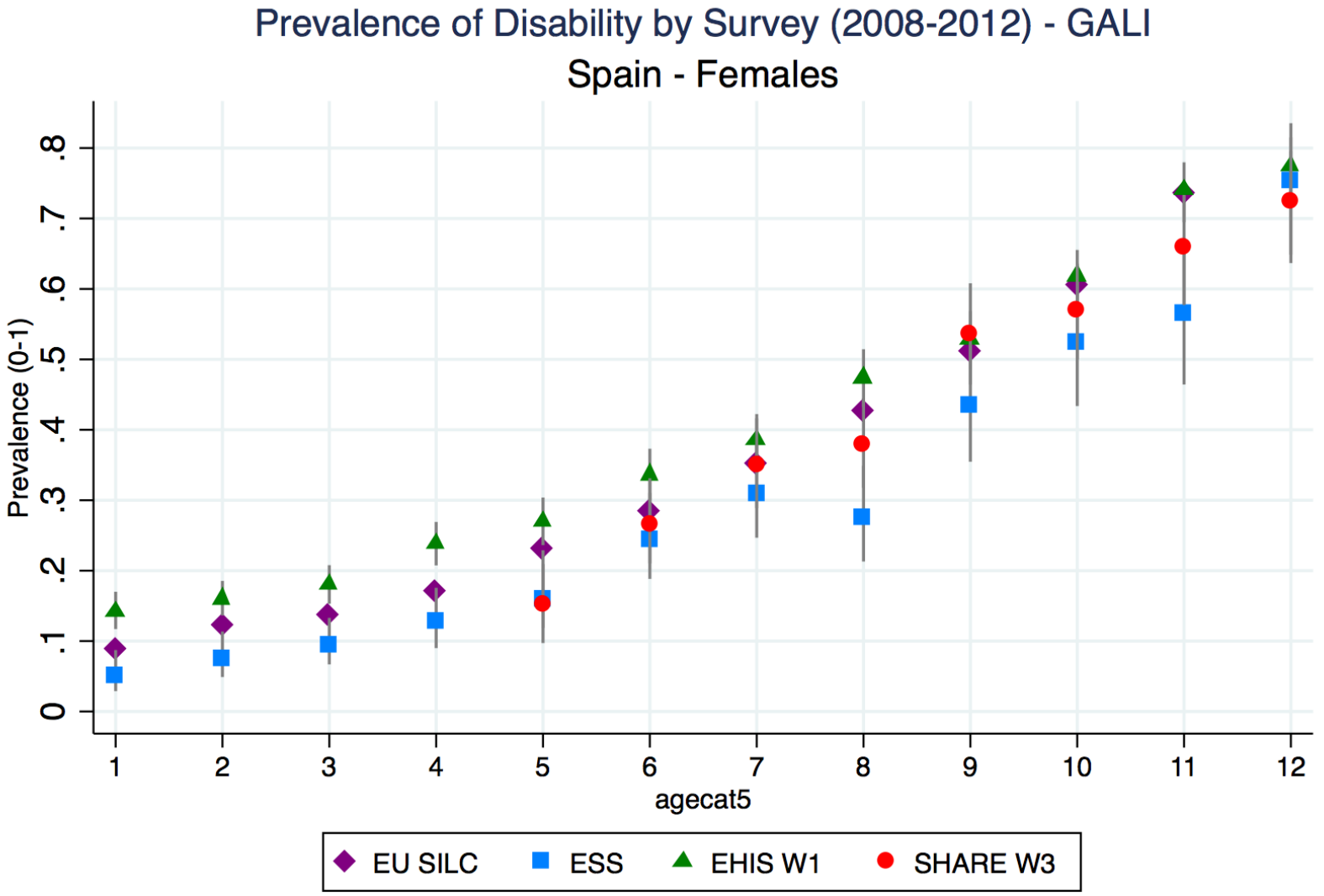
Measuring health status - Data

- 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



Measuring GALI Prevalence – An example



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

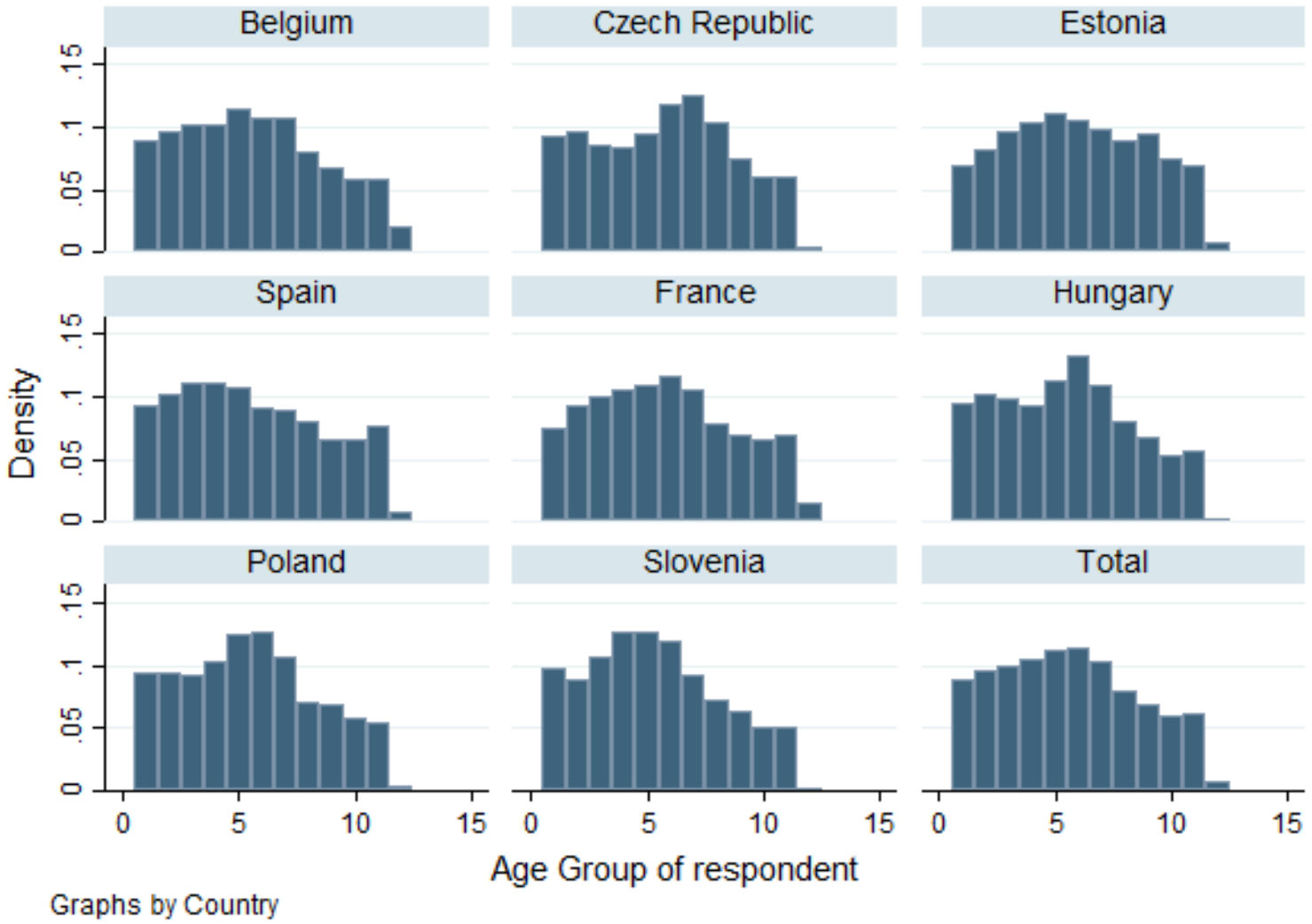
Number of Respondents by Survey and Country

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

Summary Statistics for Relevant Variables

Education		Freq	%	Cum %
	Low(ISCED I-II)	178,273	34.85	34.85
	Medium (ISCED IIIab-IV)	235,020	45.94	80.79
	High (ISCED V1-V2)	98,259	19.21	100
GALI (moderate + severe)				
	No	339,181	70.97	70.97
	Yes	138,736	29.03	100
Sex				
	Male	242,173	46.72	46.72
	Female	276,218	53.28	100

Countries Included in the Analysis



Methodological Approach

$$GALI_{ic} = \alpha + \beta_k AgeGroup_k + \beta_s Survey_s + \beta_c Country_c + \beta_e Education + \beta_i Education * Survey + \varepsilon$$

$GALI_{ic}$ is an indicator variable for whether individual *i* in country *c* reports moderate or severe disability

$AgeGroup_k$ is an ordered categorical (11 levels) variable for age groups (30-34; 35-39; 40-44*;…; 80-84). 85+ are excluded*

$Survey_s$ *is a categorical variable that denotes the survey of individual i (4 levels)*

$Country_c$ *is a categorical variable that denotes the country of individual i (7 lvl.)*

$Education * Survey$ *is an interaction term between education and survey*

ε *is the standard error term*

Methodological Approach

$$GALI_{ic} = \alpha + \beta_k AgeGroup_k + \beta_s Survey_s + \beta_c Country_c + \beta_e Education + \beta_i Education * Survey + \varepsilon$$

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

$$GALI_{ic} = \alpha + \beta_k AgeGroup_k + \beta_s Survey_s + \beta_c Country_c + \beta_e Education + \beta_i Education * Survey + \varepsilon$$

Model 1 - OLS	
EU-SILC (10/12)	Ref
EHISw1 (08)	0.09*** [0.01]
ESS (08/10/12)	-0.04*** [0.01]
SHAREw4 (2011)	0.06*** [0.01]

Model 2 – Logistic Regression	
EU-SILC (10/12)	Ref
EHISw1 (08)	0.10*** [0.00]
ESS (08/10/12)	-0.02*** [0.00]
SHAREw4 (2011)	0.07*** [0.01]

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

Males

N= 175,303

Coefficient of Interest

1) Difference in prevalence of GALI disability across different surveys

$$GALI_{ic} = \alpha + \beta_k AgeGroup_k + \beta_s Survey_s + \beta_c Country_c + \beta_e Education + \beta_i Education * Survey + \varepsilon$$

Model 1 - OLS	
EU-SILC (10/12)	Ref
EHISw1 (08)	0.12*** [0.01]
ESS (08/10/12)	-0.02*** [0.01]
SHAREw4 (2011)	0.08*** [0.01]

Model 2 – Logistic Regression	
EU-SILC (10/12)	Ref
EHISw1 (08)	0.12*** [0.00]
ESS (08/10/12)	-0.01*** [0.00]
SHAREw4 (2011)	0.10*** [0.01]

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

Females

N= 211,320

Coefficient of Interest

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

$$GALI_{ic} = \alpha + \beta_k AgeGroup_k + \beta_s Survey_s + \beta_c Country_c + \beta_e Education + \beta_i Education * Survey + \varepsilon$$

Model 1 - OLS	
Low Educated (ISCED I-II)	Ref
Medium Educated (ISCED III-IV)	-0.09*** [0.00]
Highly Educated (ISCED V-VI)	-0.15*** [0.00]

Model 2 – Logistic Regression	
Low Educated (ISCED I-II)	Ref
Medium Educated (ISCED III-IV)	-0.09*** [0.00]
Highly Educated (ISCED V-VI)	-0.16*** [0.00]

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

Males

N= 175,303

Coefficient of Interest

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

$$GALI_{ic} = \alpha + \beta_k AgeGroup_k + \beta_s Survey_s + \beta_c Country_c + \beta_e Education + \beta_i Education * Survey + \varepsilon$$

Model 1 - OLS	
Low Educated (ISCED I-II)	Ref
Medium Educated (ISCED III-IV)	-0.10*** [0.00]
Highly Educated (ISCED V-VI)	-0.16*** [0.00]

Model 2 – Logistic Regression	
Low Educated (ISCED I-II)	Ref
Medium Educated (ISCED III-IV)	-0.09*** [0.00]
Highly Educated (ISCED V-VI)	-0.17*** [0.00]

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

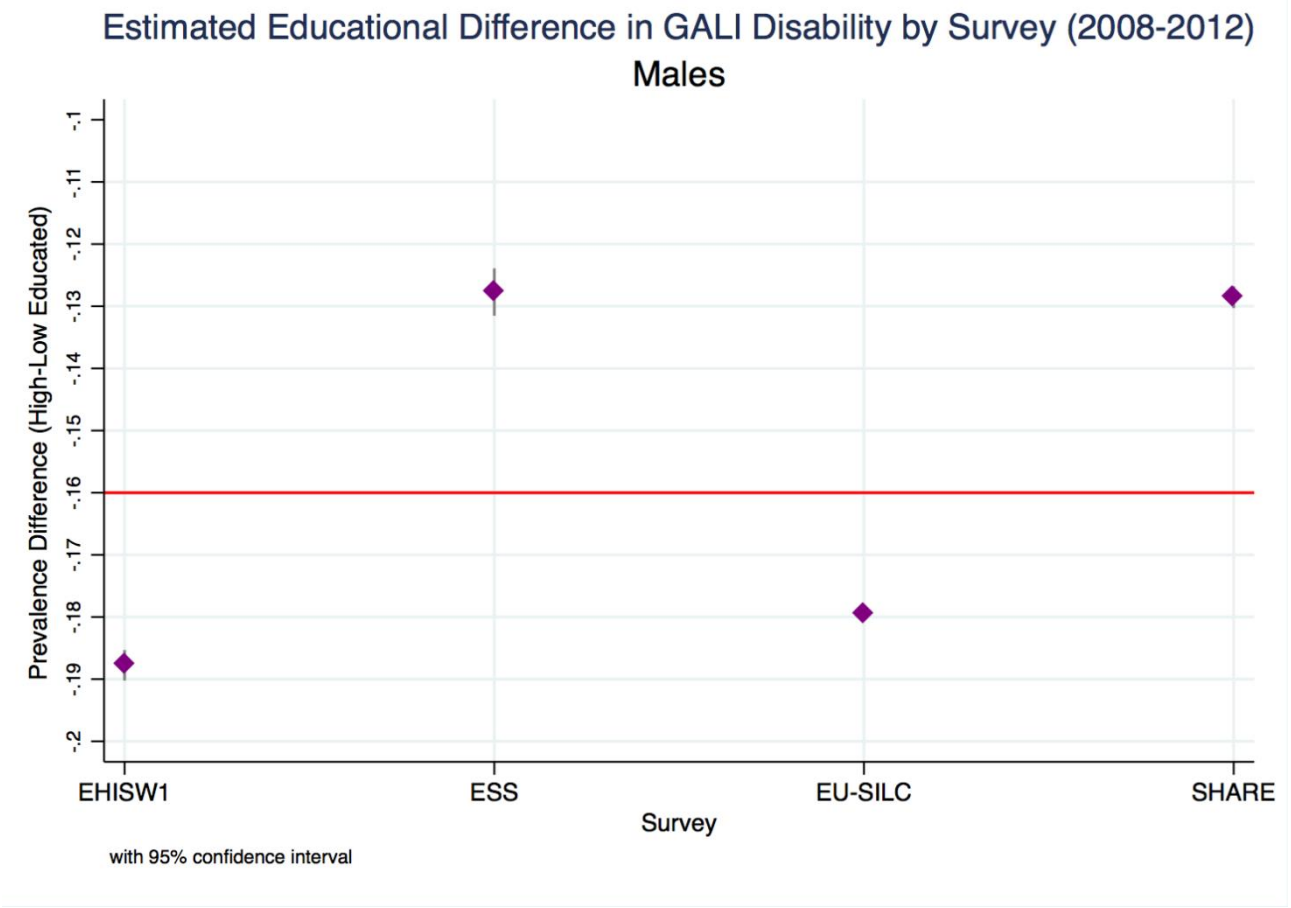
Females

N= 211,320

Coefficient of Interest

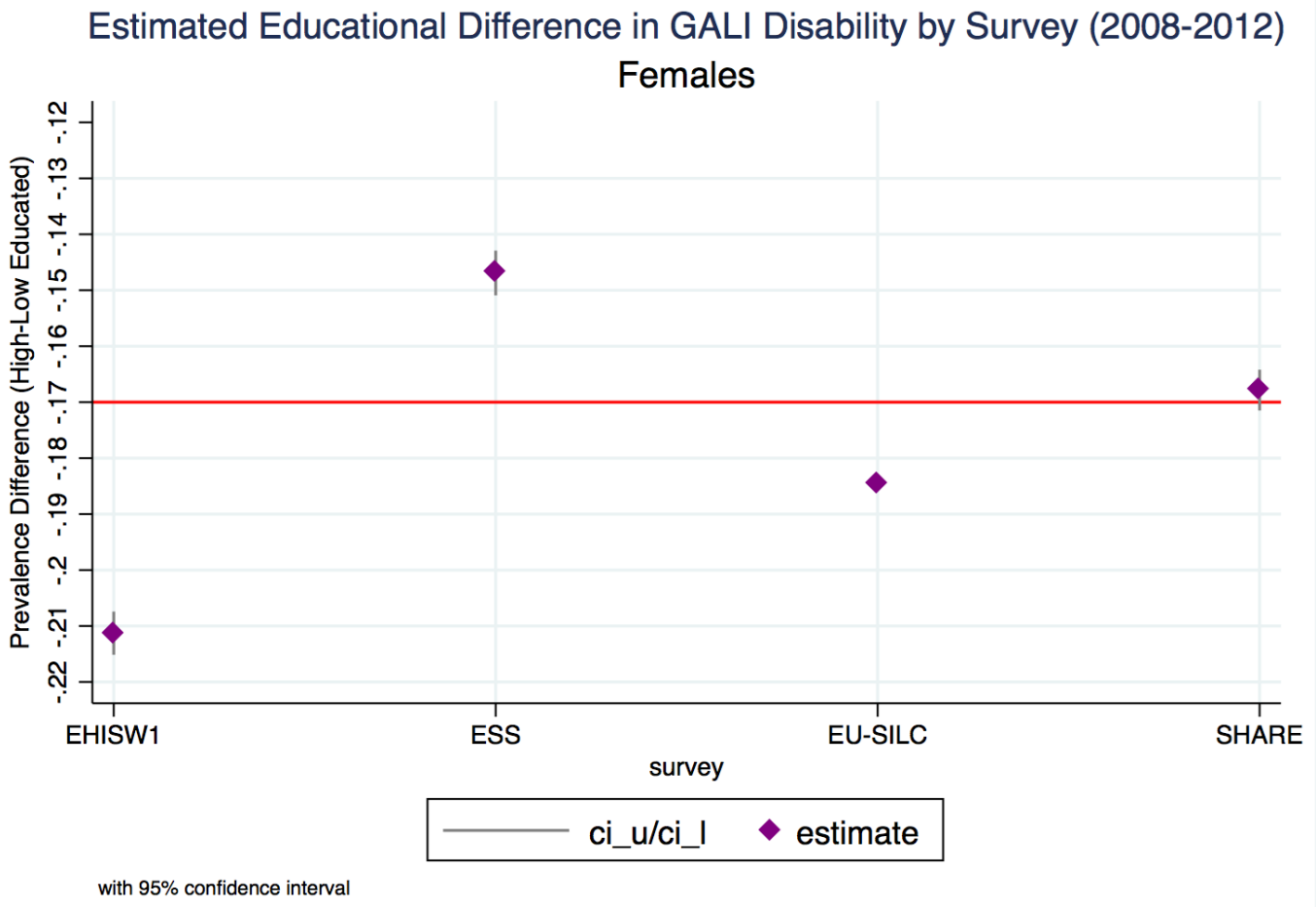
3) Quantify the differential in educational disparities in GALI disability across surveys

$$GALI_{ic} = \alpha + \beta_k\, AgeGroup_k + \beta_s Survey_s + \beta_c Country_c + \beta_e Education + \beta_i Education * Survey + \varepsilon$$

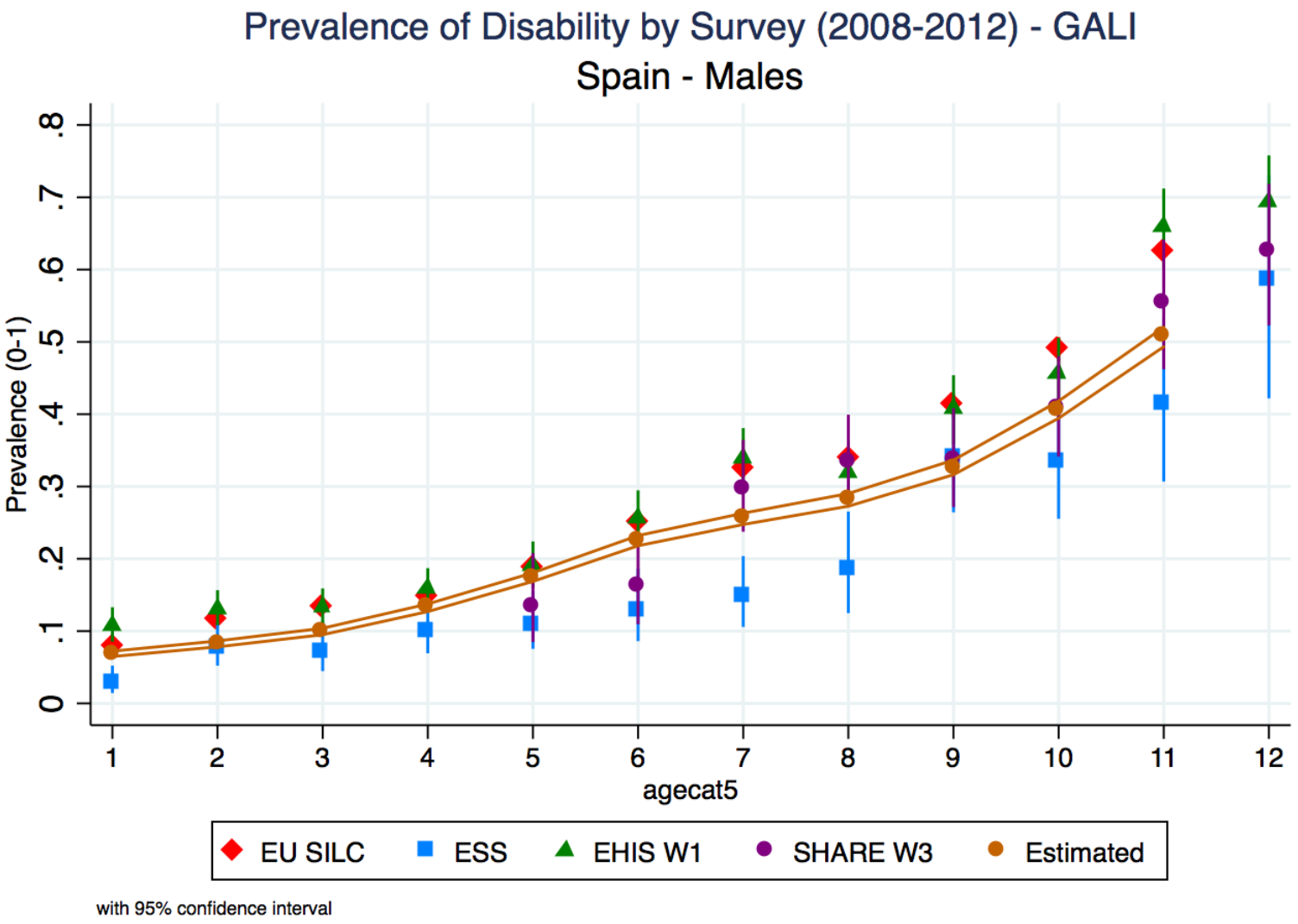


3) Quantify the differential in educational disparities in GALI disability across surveys

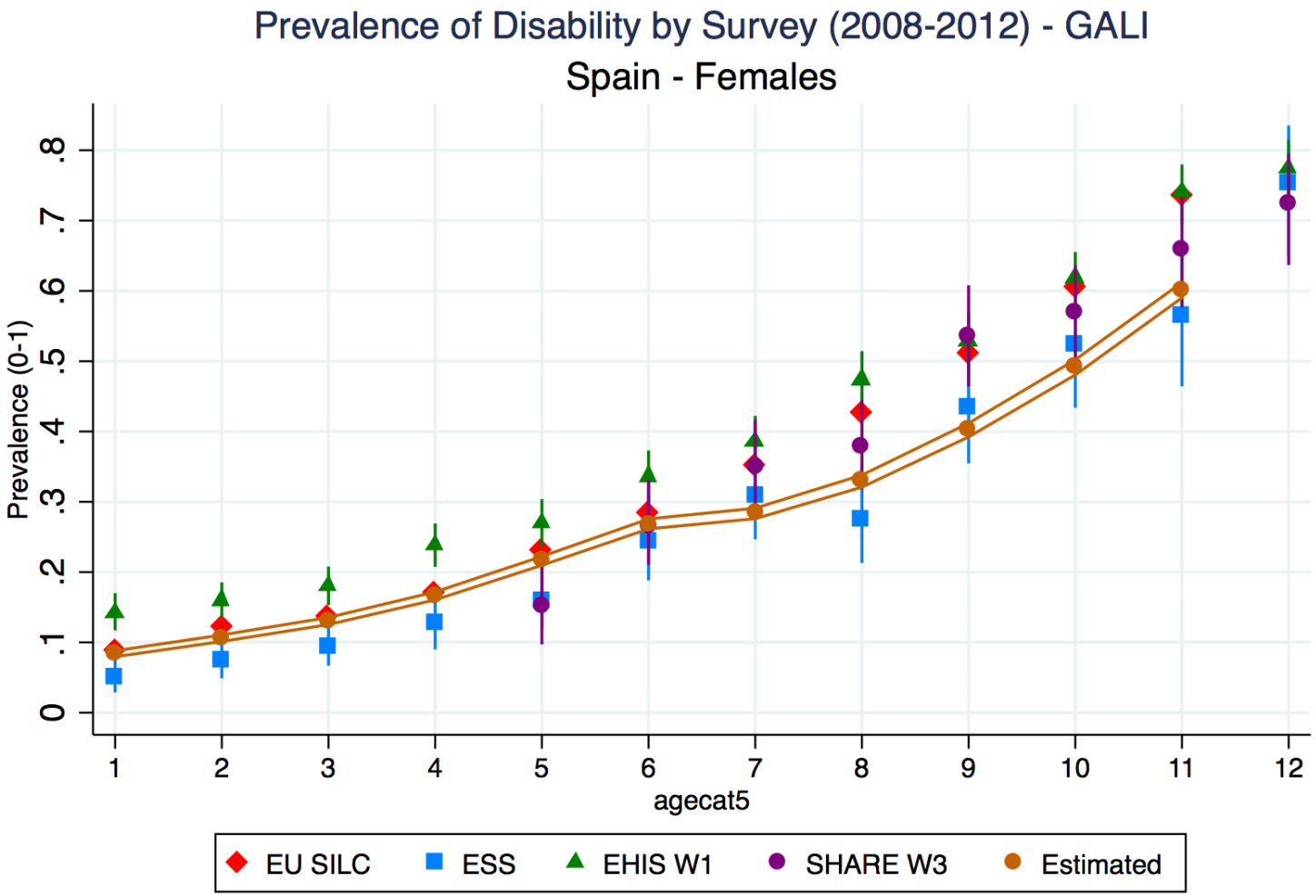
$$GALI_{ic} = \alpha + \beta_k AgeGroup_k + \beta_s Survey_s + \beta_c Country_c + \beta_e Education + \beta_i Education * Survey + \varepsilon$$



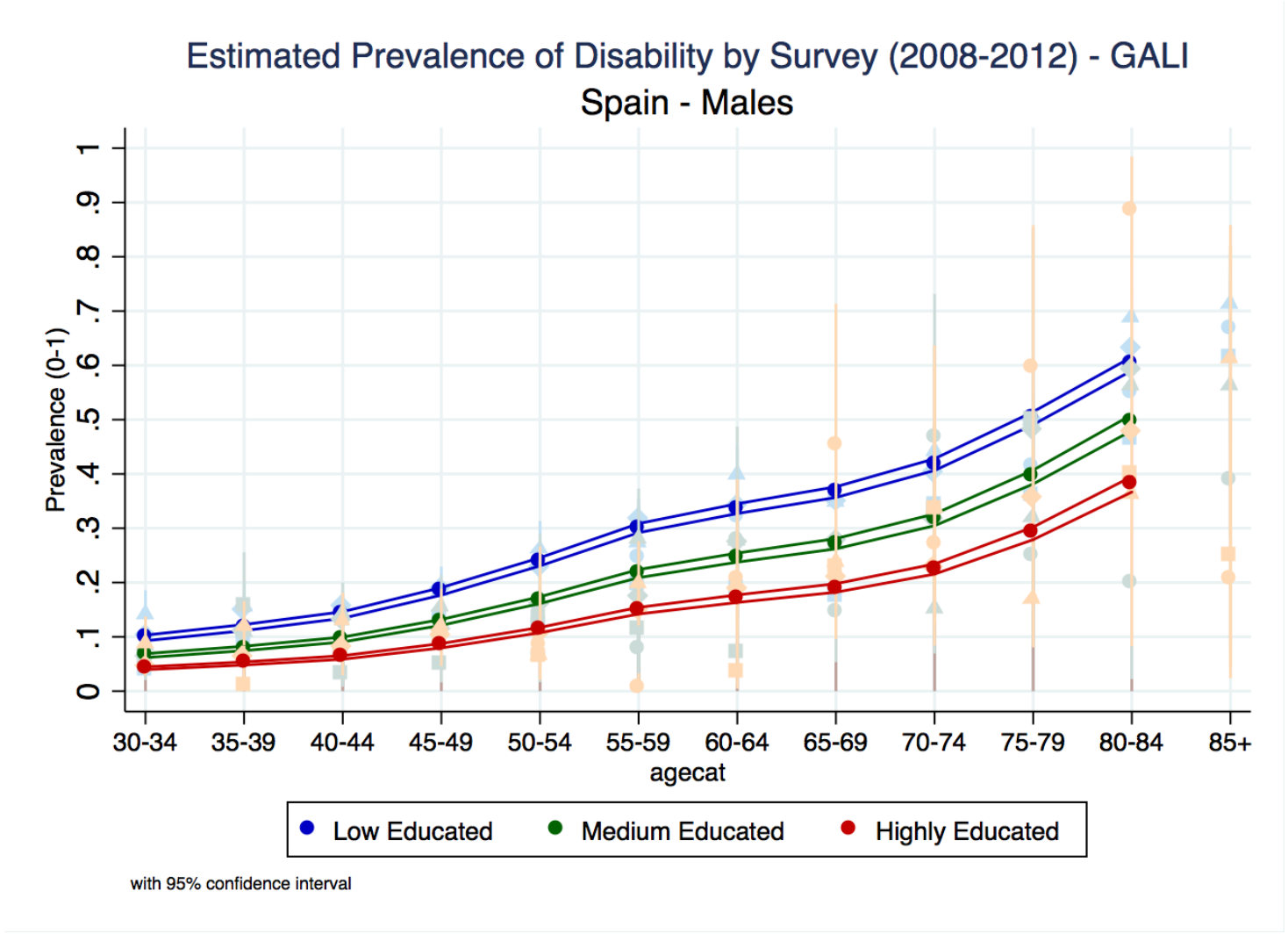
4) Predict GALI disability across age groups (and educational attainment) – An Example



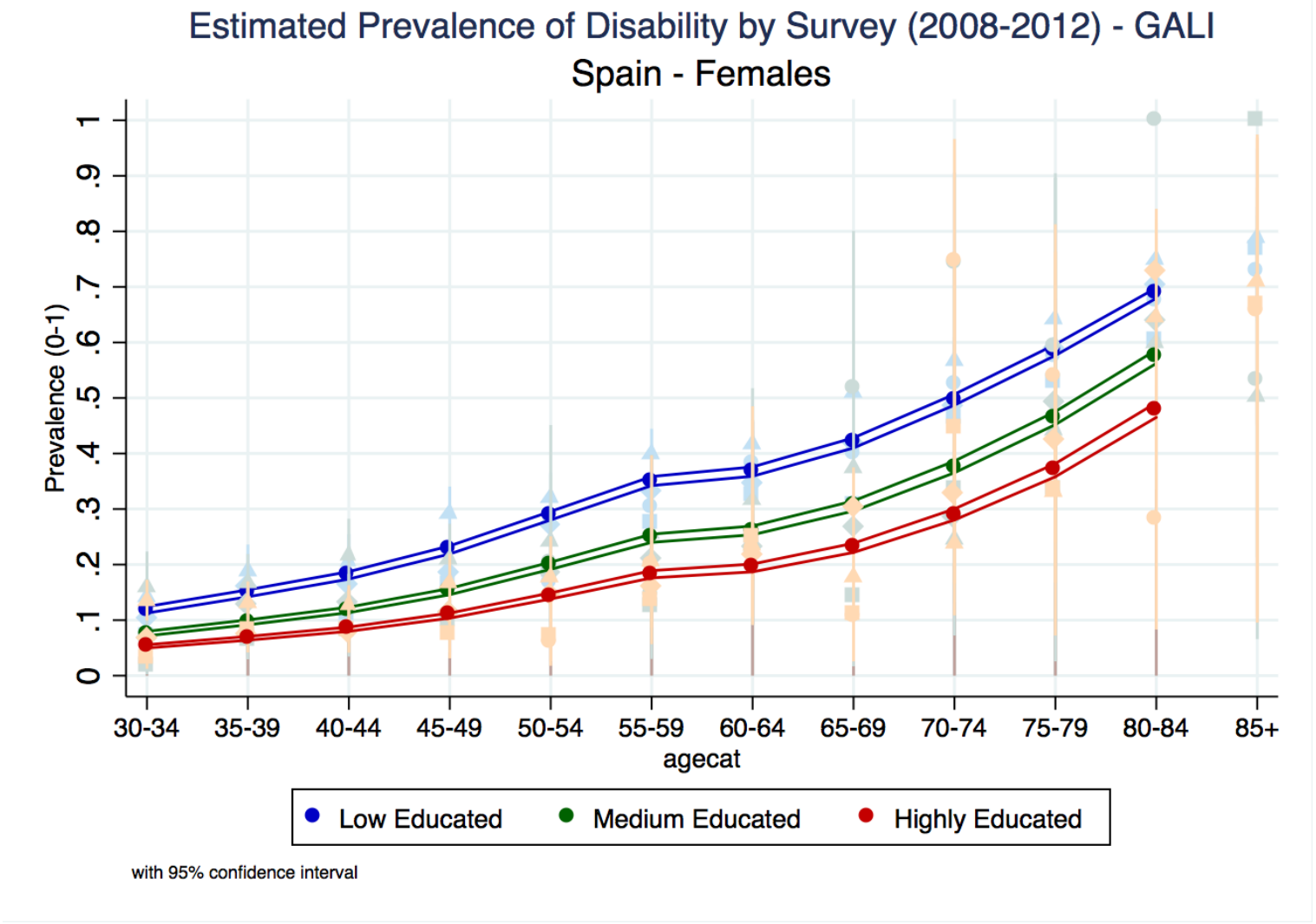
4) Predict GALI disability across age groups (and educational attainment) – An Example (Spain)



4) Predictions by Educational Status



4) Predictions by Educational Status



Limitations

- Further robustness checks are required
- **Goodness of fit** needs to be assesed 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

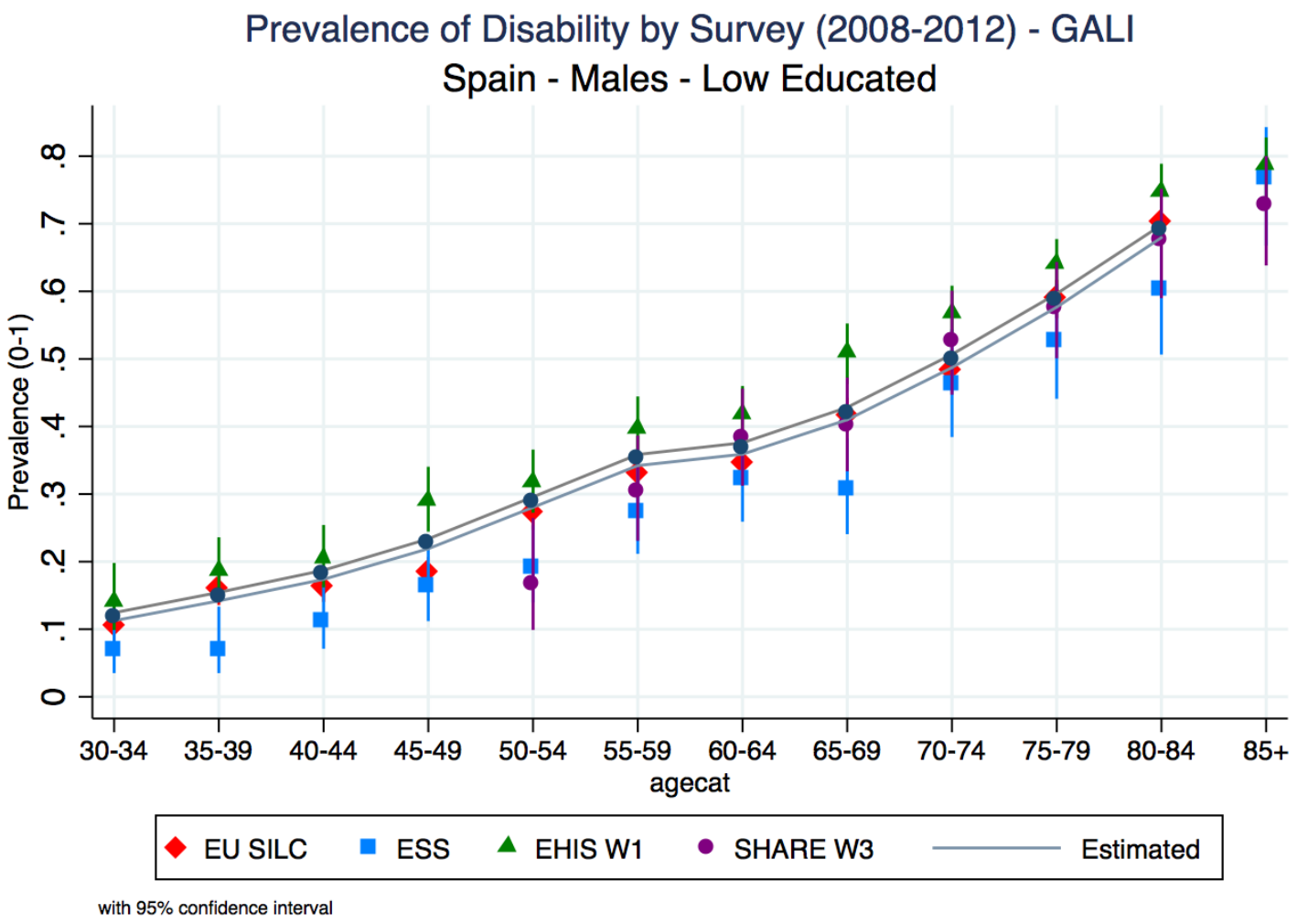
Conclusion

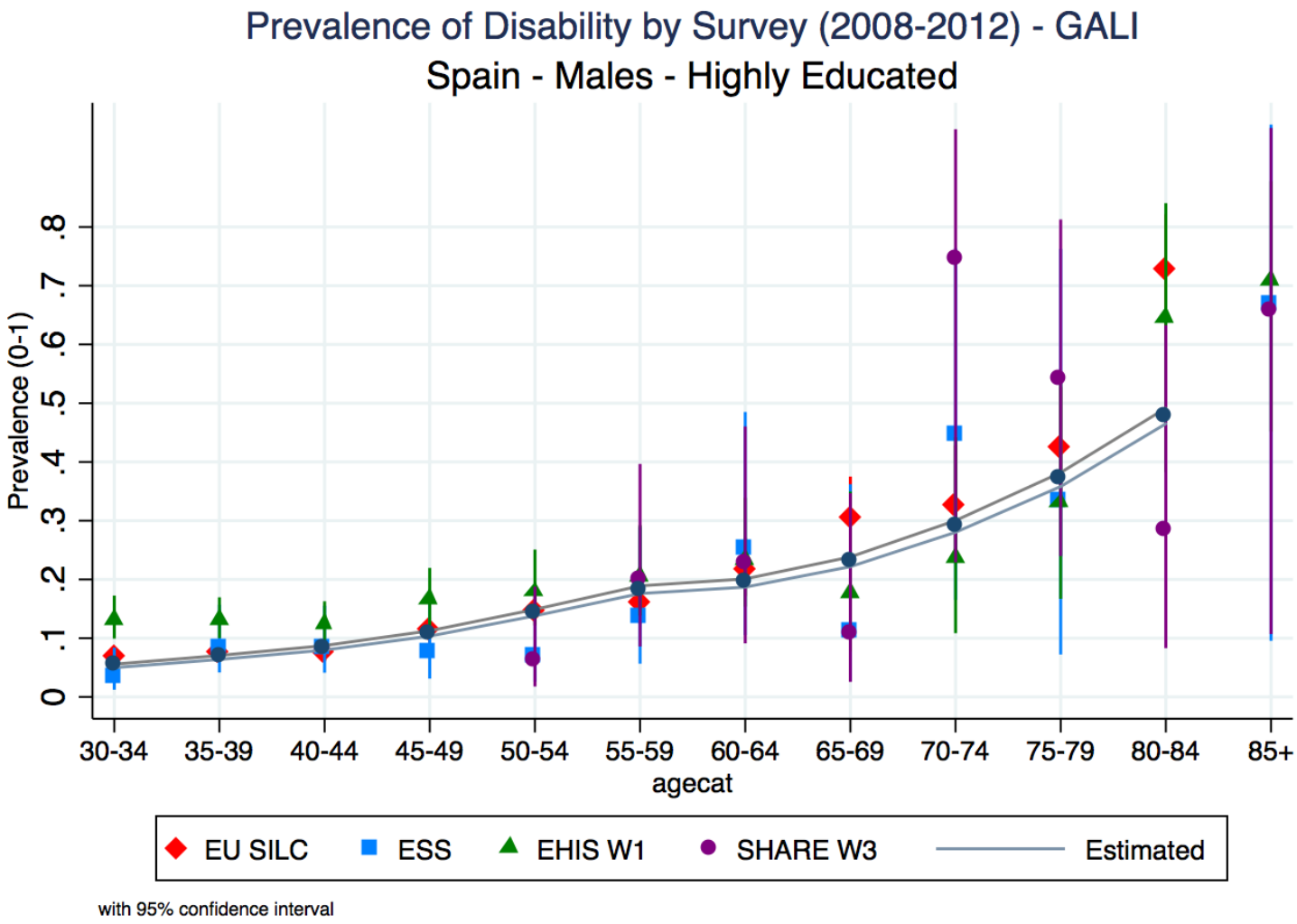
1) Very little data in last age group makes it hard to include it. Further complicated by stratification.
2) Holding other covariates constant, the surveys have statistically significant differences when measuring GALI disability
3) There are non-trivial inequalities in GALI disability across high and 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.

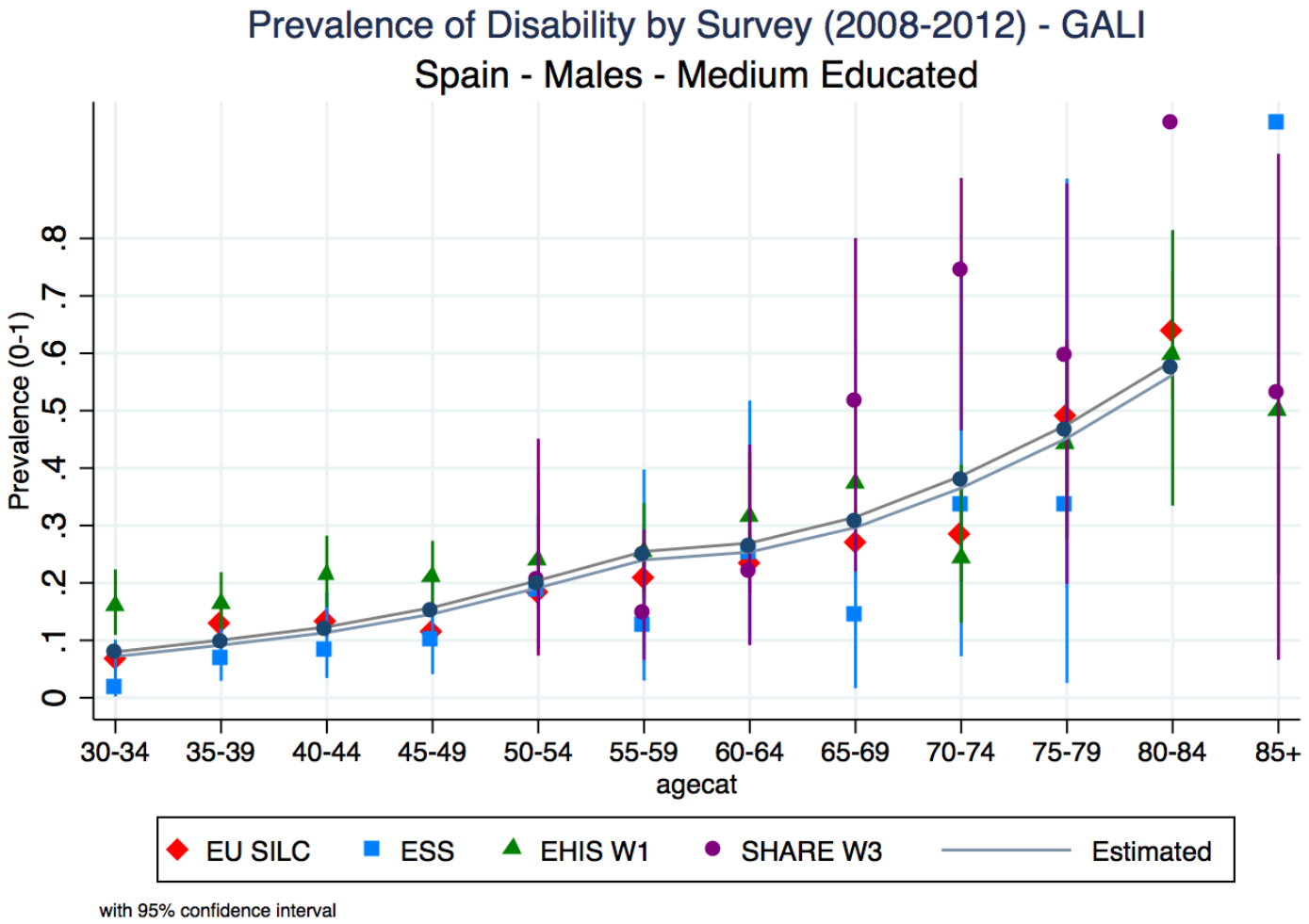
Data References

- *EUSILC UDB 2008 – version 7 of March 2015*
- *EUSILC UDB 2012 – version 1 of January 2016*
- ESS Round 4: European Social Survey Round 4 Data (2008). Data file edition 4.3. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- 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
- Börsch-Supan, A. (2013). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 4. Release version: 5.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w4.500

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)
0
.1
.2
.3
.4
.5
.6
.7
.8
30-34
35-39
40-44
45-49
50-54
55-59
60-64
65-69
70-74
75-79
80-84
85+
agecat
EU SILC
ESS
EHIS W1
SHARE W3
Estimated
with 95% confidence interval


Prevalence of Disability by Survey (2008-2012) - GALI
Spain - Males - Highly Educated
Prevalence (0-1)
0
.1
.2
.3
.4
.5
.6
.7
.8
30-34
35-39
40-44
45-49
50-54
55-59
60-64
65-69
70-74
75-79
80-84
85+
agecat
EU SILC
ESS
EHIS W1
SHARE W3
Estimated
with 95% confidence interval


Prevalence of Disability by Survey (2008-2012) - GALI
Spain - Males - Medium Educated
Prevalence (0-1)
0
.1
.2
.3
.4
.5
.6
.7
.8
30-34
35-39
40-44
45-49
50-54
55-59
60-64
65-69
70-74
75-79
80-84
85+
agecat
EU SILC
ESS
EHIS W1
SHARE W3
Estimated
with 95% confidence interval