Adolescent interest in science careers in Europe: Trends between 2006 and 2015, an example of Stata analysis

Joanna Sikora

School of Sociology
Australian National University

Outline

1. Problem: Why study adolescent plans to work in science (STEMM)? Why study the gender gap?
2. Definitional issues: math intensive versus life sciences
3. Data
4. Stata tools
5. Three levels of predictors of STEM career plans
6. Trends in STEM career plans in Europe 2006-2015
7. Challenges of visually presenting complex results

1: Why study STEMM career plans of adolescents?

- Documented historical decrease of interest among youth in science professions (particularly among young women)
- Concerns of government that the future workforce will need quantitative science skill to be competitive in labor market and competent to deal with every day life problems
- Adolescents change their minds, but their overall choices of courses and vocational orientation made at end of compulsory education matter for what happens to them later

Why Europe?

- Consultancy I am doing in 2017 for the European Commission’s Joint Research Centre in Italy.

2. Definitions of STEMM or science

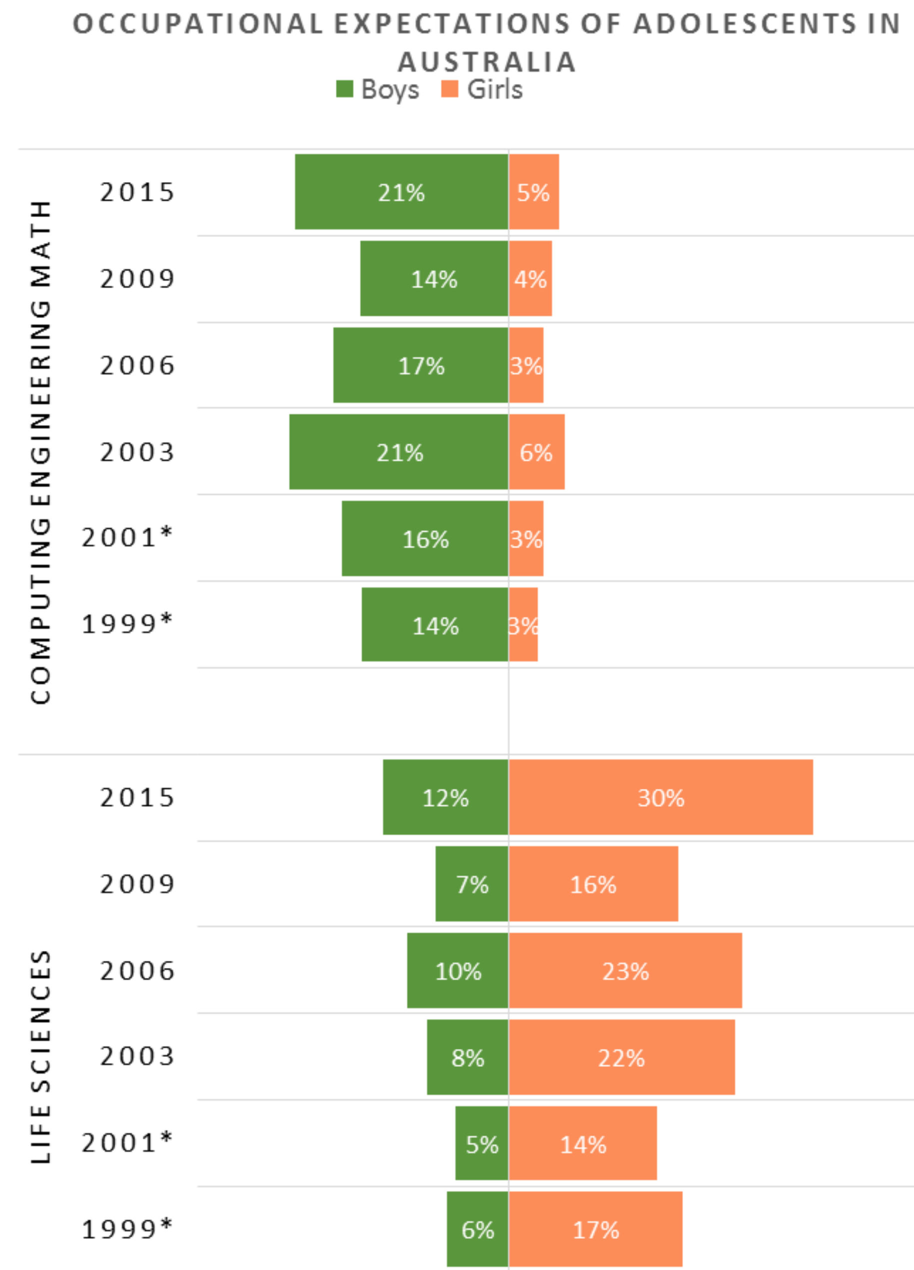
- **Many**
- **Here categories based on the International Standard Classification of Occupations (see ilo.org for ISCO-08 and ISCO-88)**
- **Science occupations involve jobs in ISCO Major 2 and 3 groups i.e. professions, associate professions and a couple of managerial titles**
- **Distinguish two occupational groups in science:**
  1. **Math intensive occupations: engineering, computing, math, physics**
  2. **Life sciences: health, medicine, biology (also nursing and psychology)**

Sikora, J. and **A. Pokropek** (2012), “Gender segregation of adolescent science career plans in 50 countries”, *Science Education,* Vol. 96/2, pp. 234-264, http://dx.doi.org/10.1002/sce.20479.

Australia:

stable pattern of segregation in adolescent occupational expectations

STEMM: Why distinguish between life sciences and math intensive sciences?



Source: Longitudinal Surveys of Australian Youth

* Denotes the same cohort of students surveyed in Year 10 and 12

Australian National University

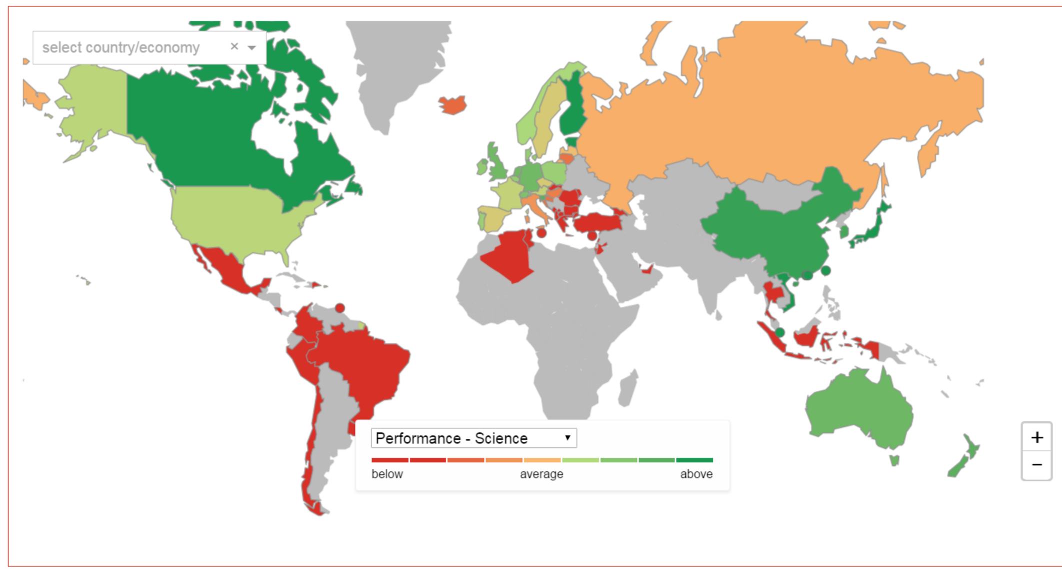
OECD
BETTER POLICIES FOR BETTER LIVES

Data

PISA Programme for International Student Assessment

PISA surveys:
2000 reading
2003 mathematics
2006 science
2009 reading
2012 mathematics
2015 science

Data



https://www.youtube.com/watch?v=q1I9tuScLUA

Occupational expectations:

"What occupation do you expect to work in when you are 30 years of age?"

Verbatim answers coded to the 4 digit level of the International Standard Classification of Occupations ISCO88/ ISCO08

Challenges

- Complex sample design: students clustered in schools
- Weights: replicate weights (BRR weights), to account for complex survey designs in the estimation of sampling variances
- Plausible values: 5 or 10 values representing the likely distribution of a student's proficiency to indicate students' academic performance (multiple imputations)
- Missing data (multiple imputations)
- Presenting complex results in accessible manner

Stata tools used

repest estimates statistics using replicate weights (BRR weights, Jackknife replicate weights,...), thus accounting for complex survey designs in the estimation of sampling variances. It is specially designed to be used with the PISA, PIAAC and TALIS datasets produced by the OECD, but works for ALL and IALS datasets as well. It also allows for analyses with multiply imputed variables (plausible values); where plausible values are included in a *pvvarlist*, the average estimator across plausible values is reported and the imputation error is added to the variance estimator.

Save subset of variables in memory to an Excel file

export excel [varlist] using filename [if] [in] [, export_excel_options]

spmap -- Visualization of spatial data

Three level analyses with interaction terms

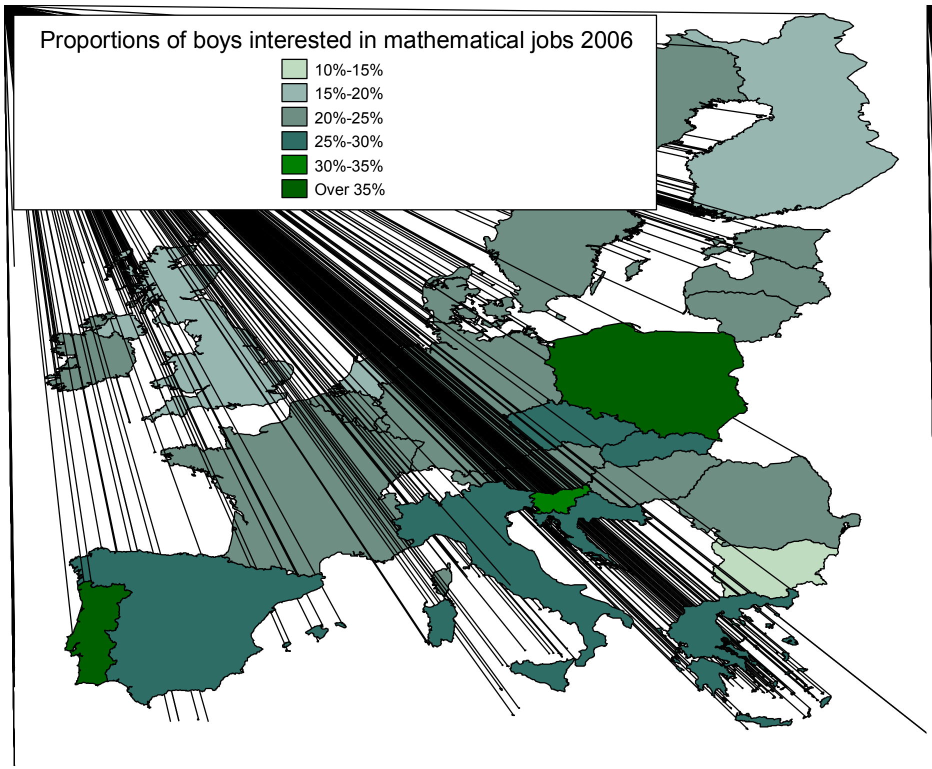
Time trend

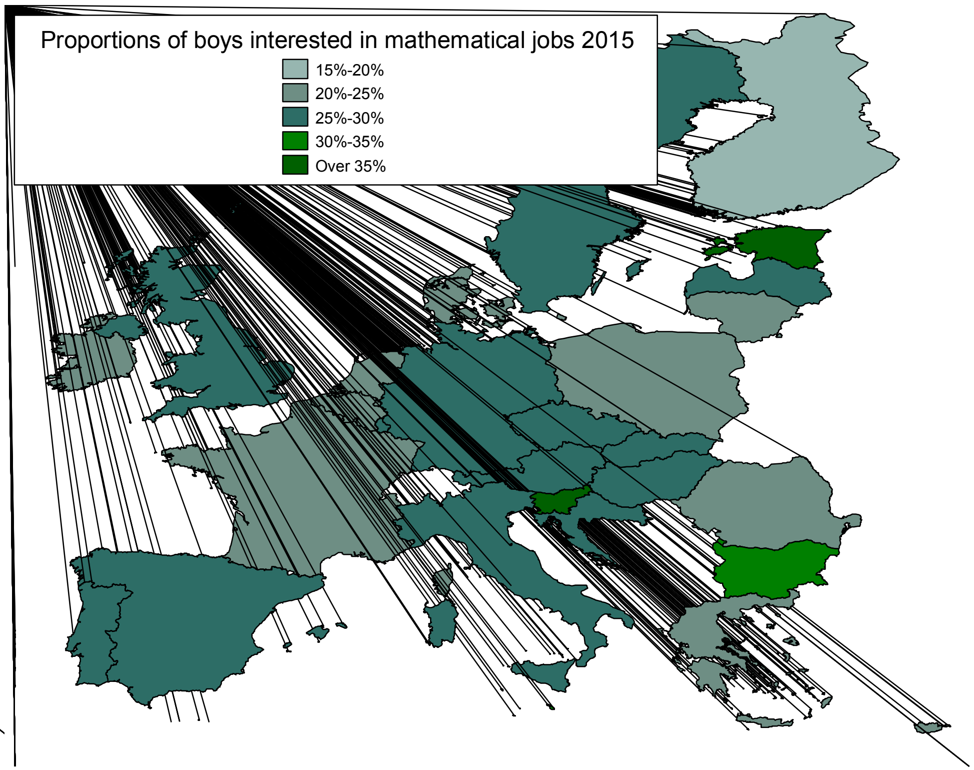
Also focus on two issues:

- Overall interest in STEMM in European countries by gender (% males plus % females

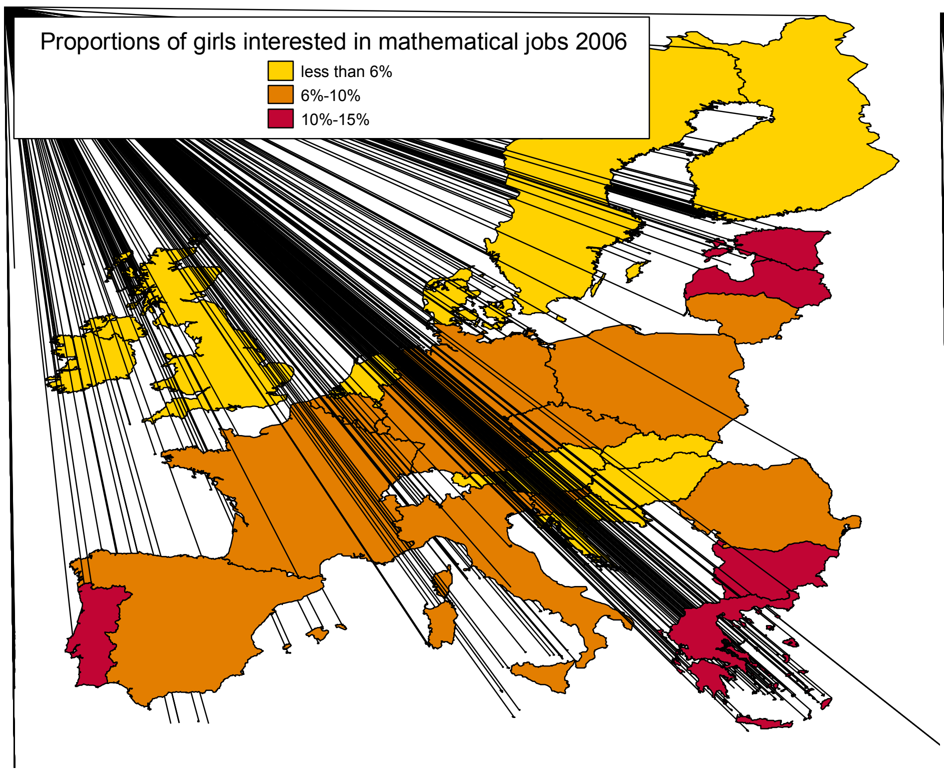
- The gender gap in this interest (% males - % females who want a STEMM job in the future)

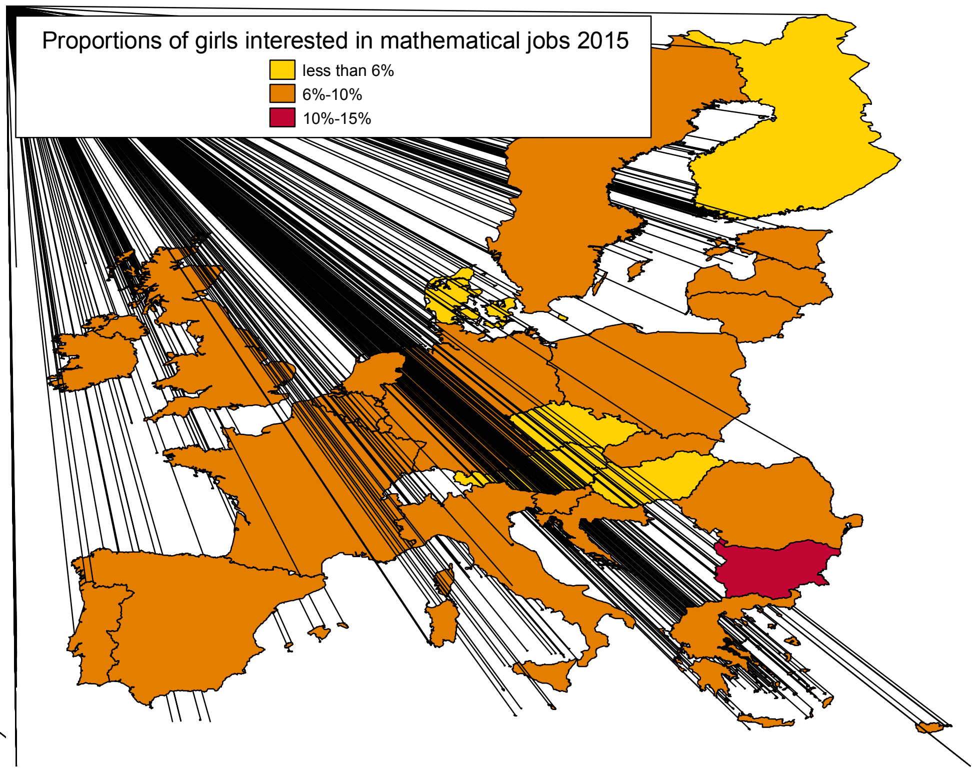
Europe trends for boys: 2006 - 2015



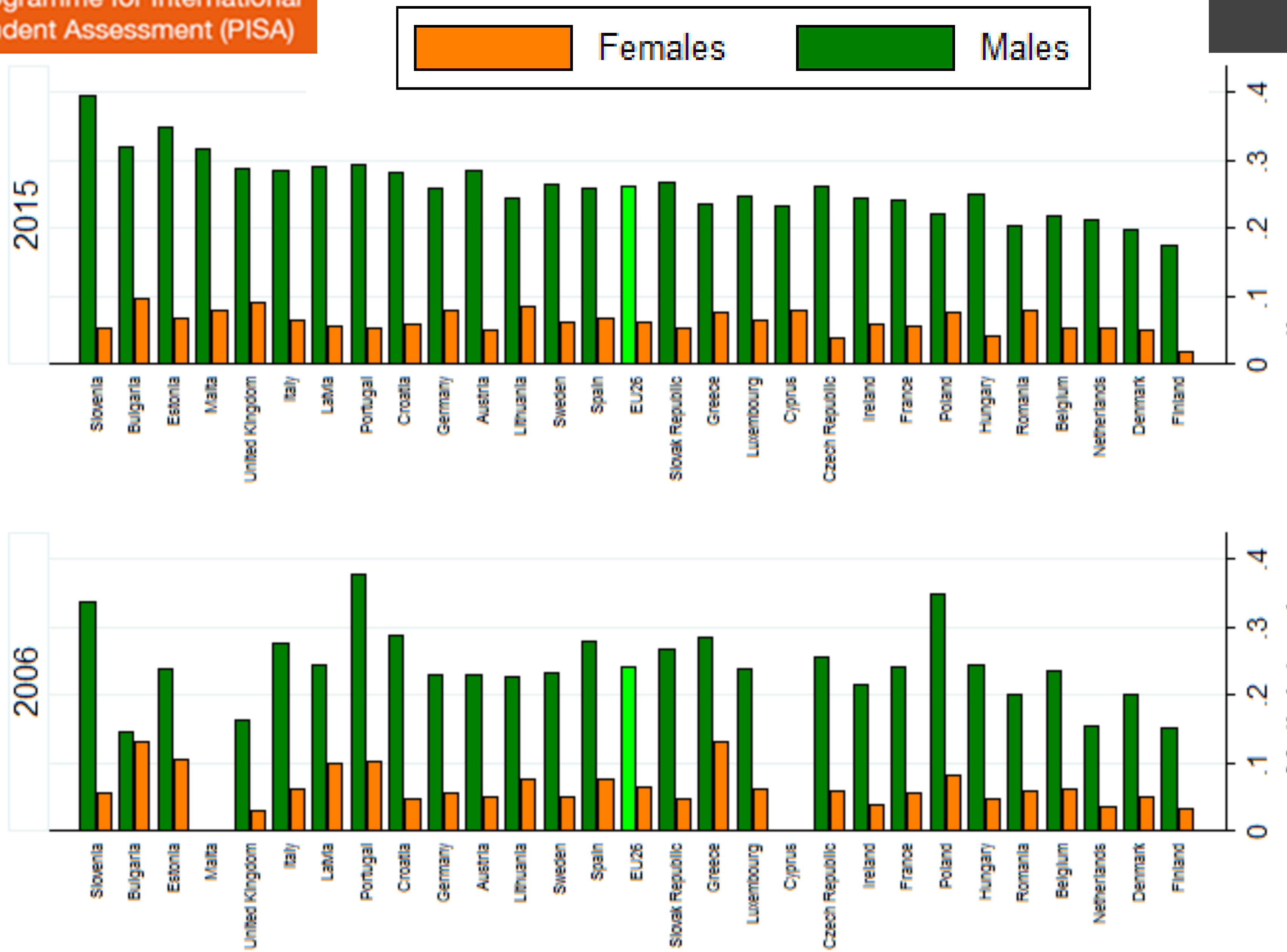


Europe trends for girls: 2006 - 2015

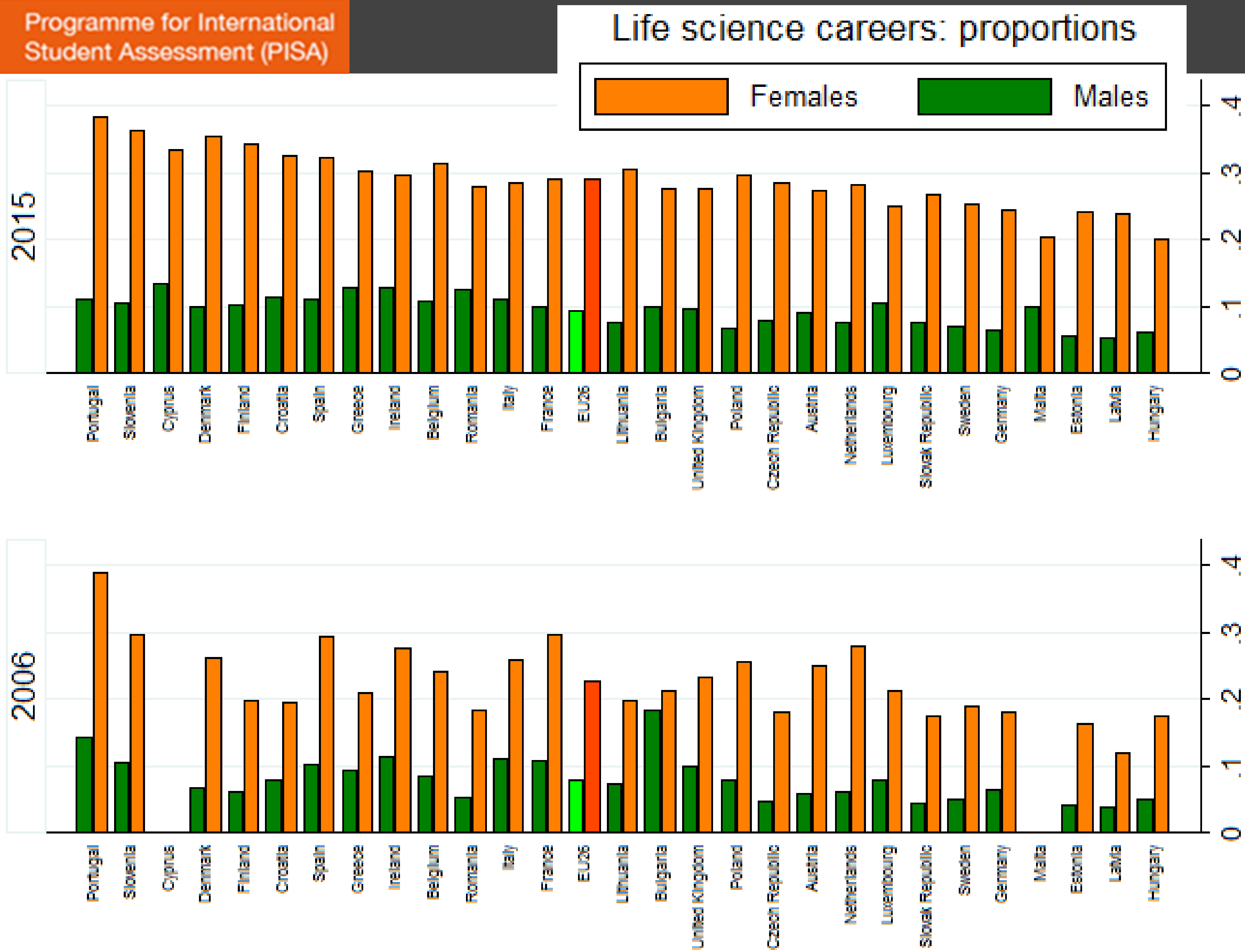




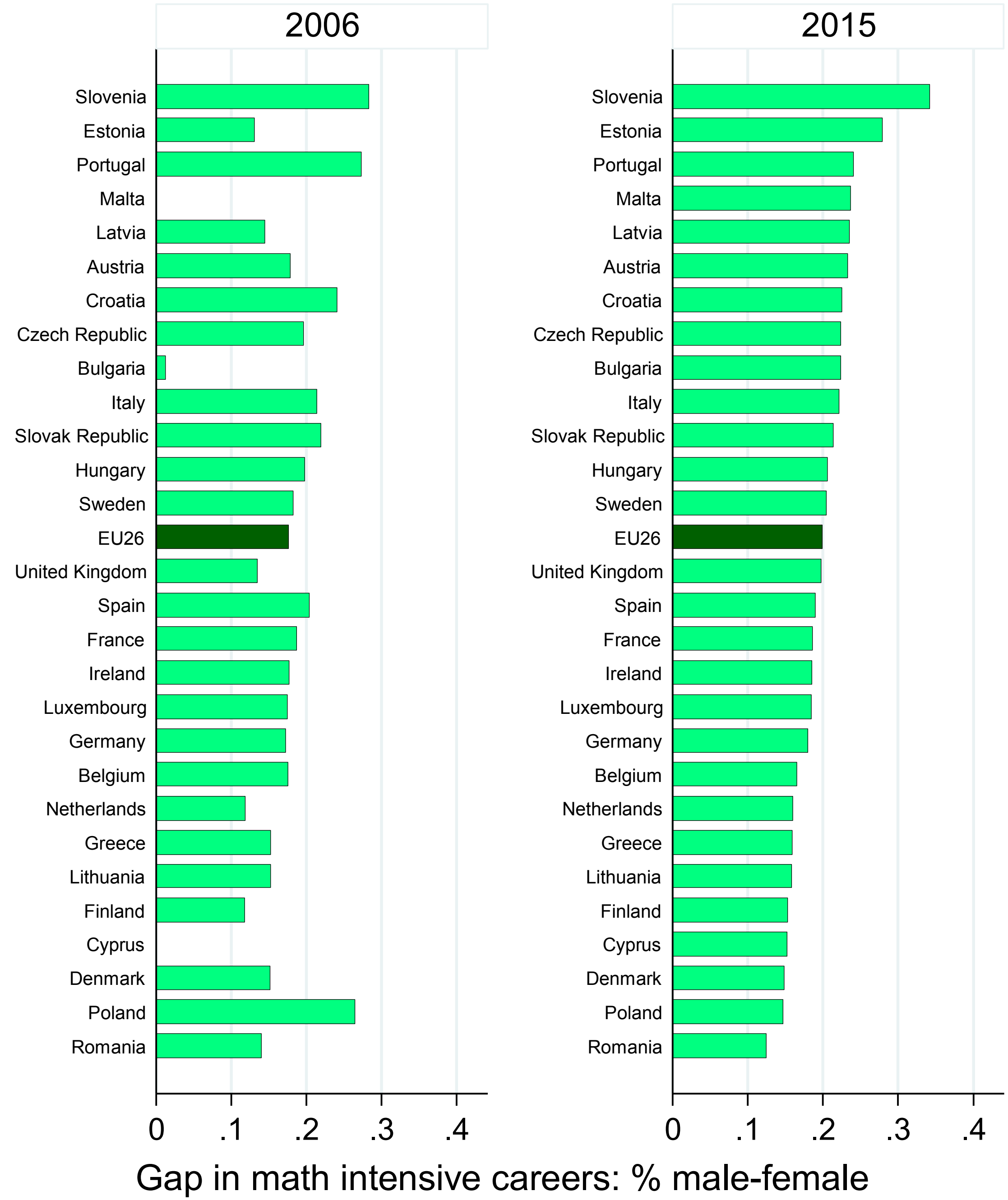
Math intensive careers: proportions



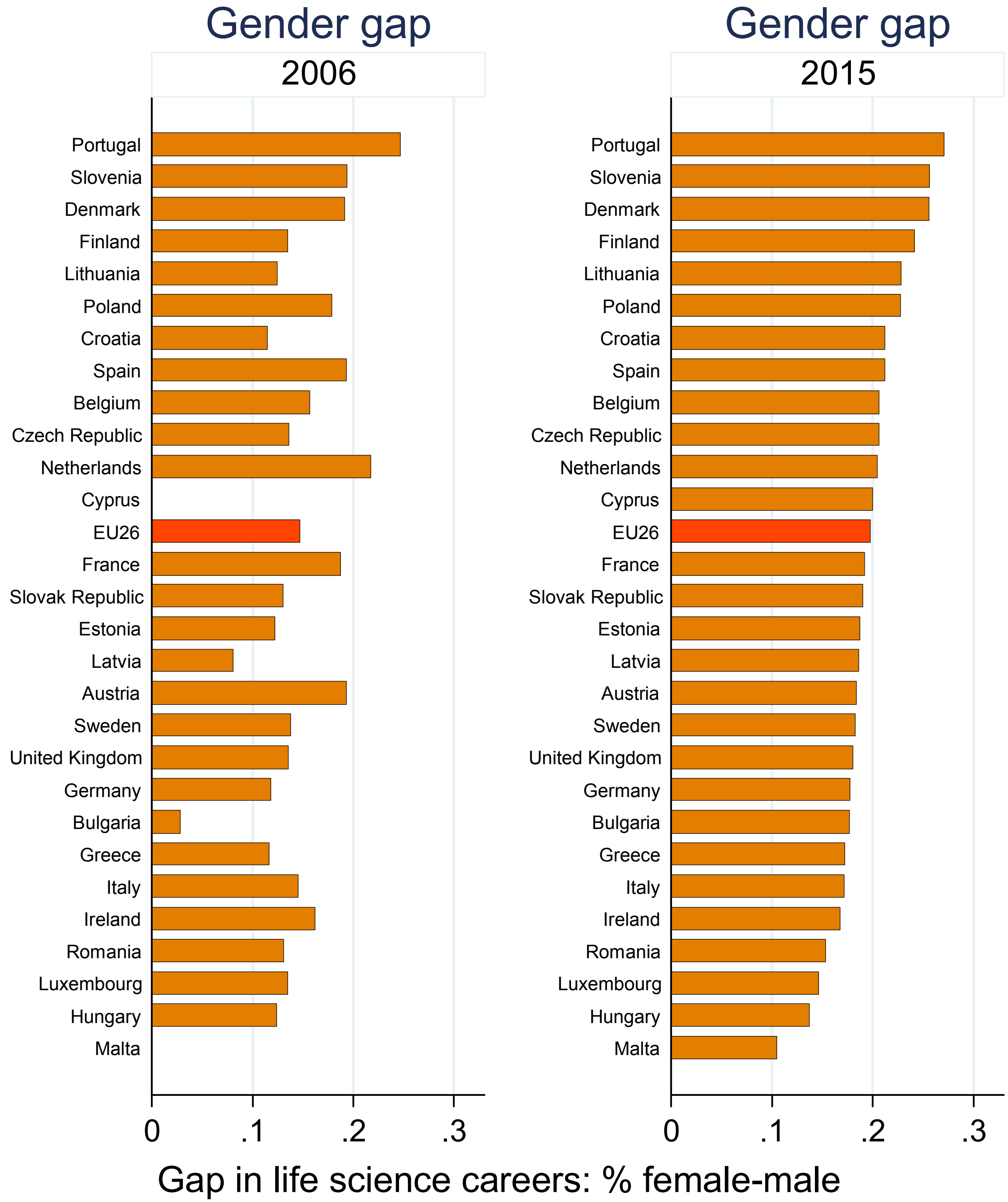
Programme for International Student Assessment (PISA)

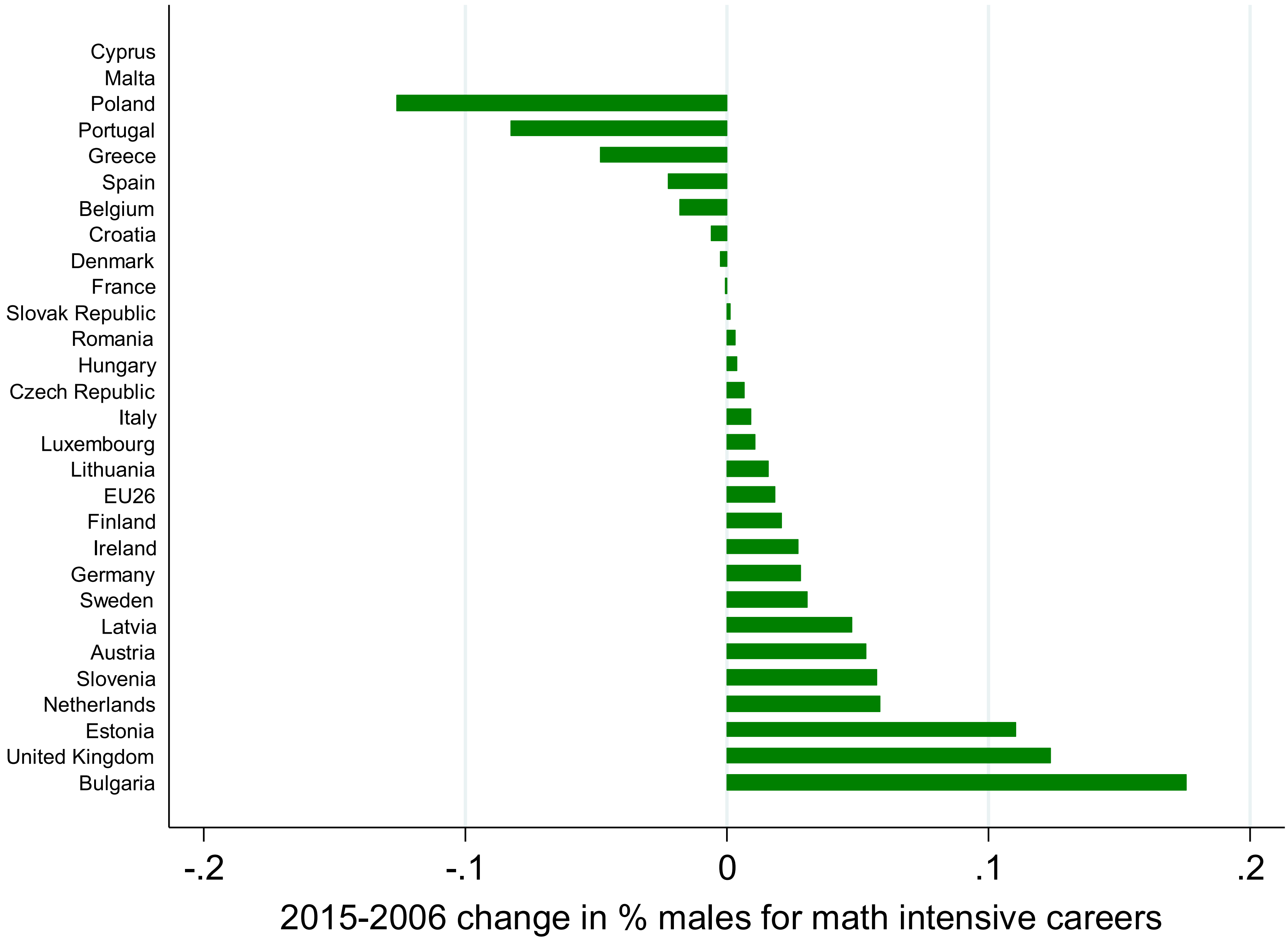


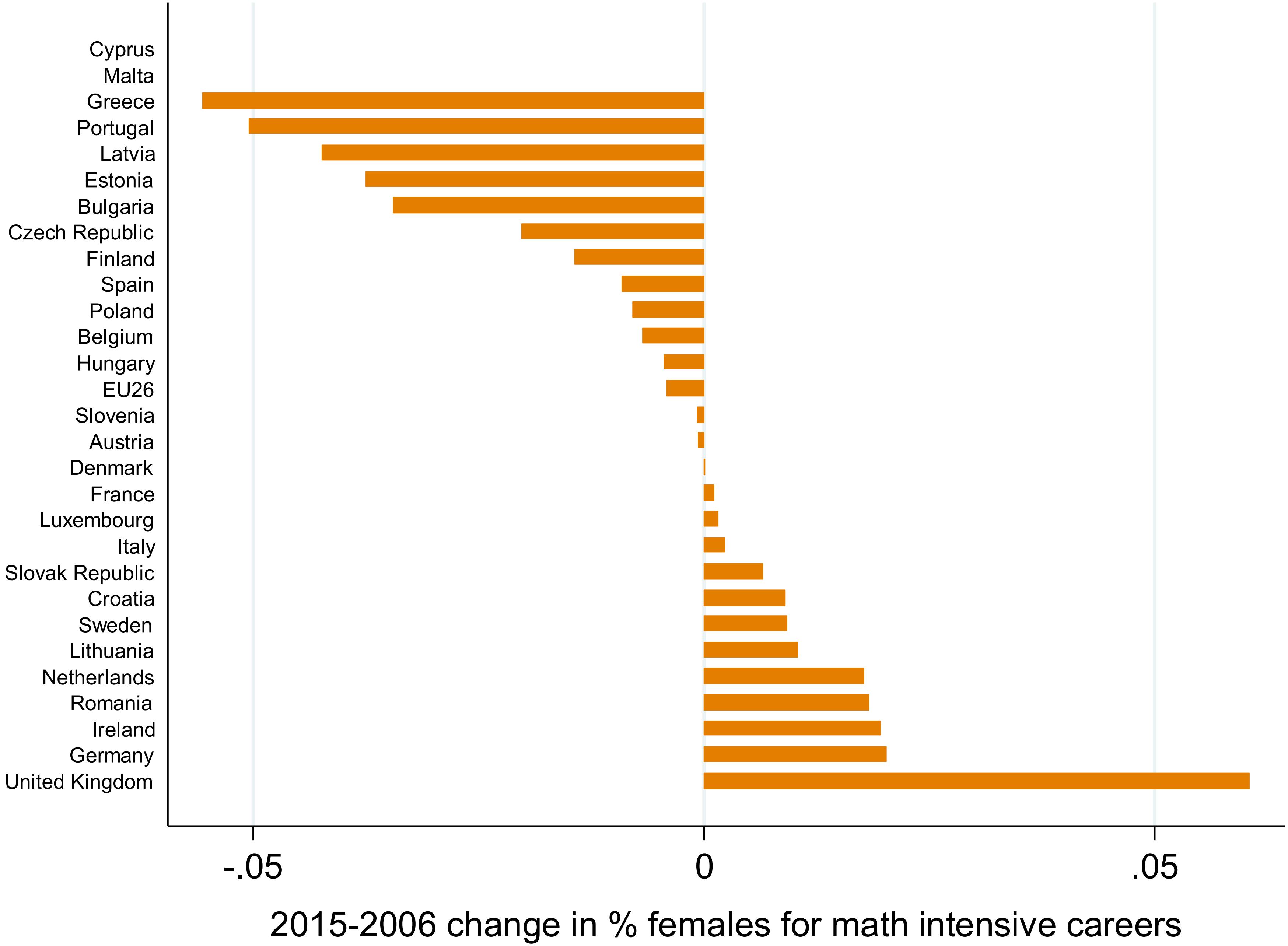
% Male - % Female Over time

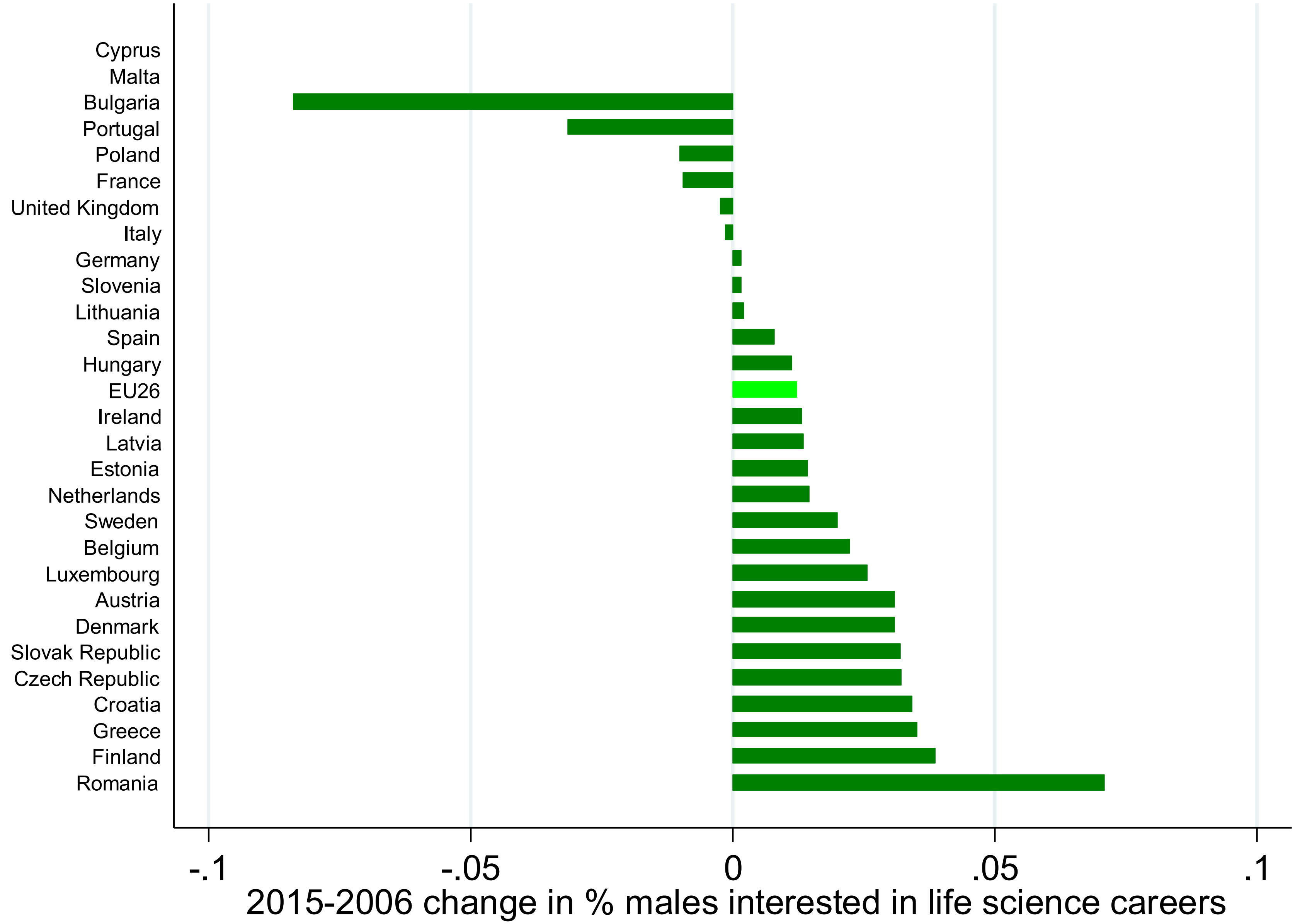


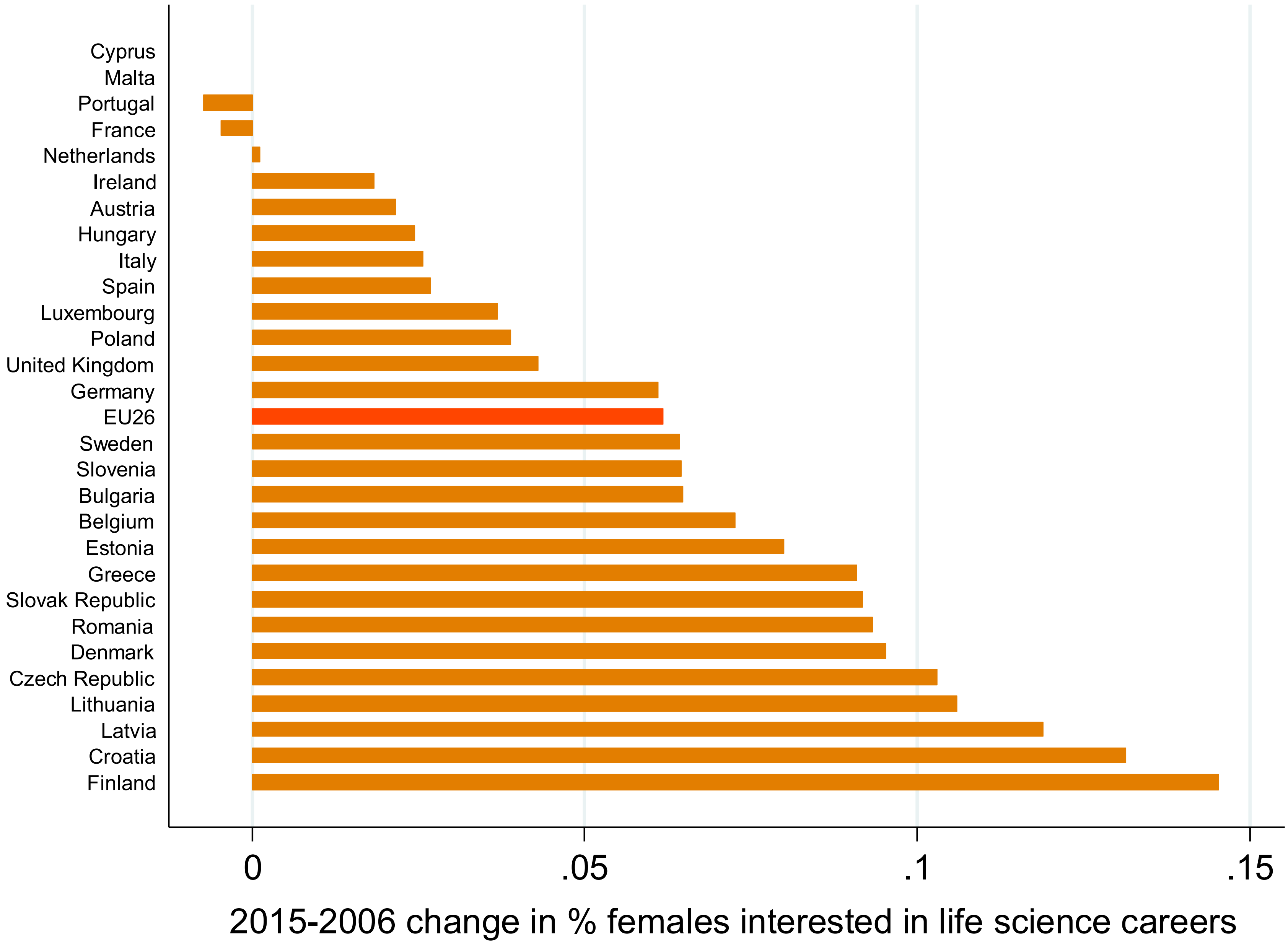
% Female - % Male
Over time




Cyprus
Malta
Poland
Portugal
Greece
Spain
Belgium
Croatia
Denmark
France
Slovak Republic
Romania
Hungary
Czech Republic
Italy
Luxembourg
Lithuania
EU26
Finland
Ireland
Germany
Sweden
Latvia
Austria
Slovenia
Netherlands
Estonia
United Kingdom
Bulgaria
-.2
-.1
0
.1
.2
2015-2006 change in % males for math intensive careers


Cyprus
Malta
Greece
Portugal
Latvia
Estonia
Bulgaria
Czech Republic
Finland
Spain
Poland
Belgium
Hungary
EU26
Slovenia
Austria
Denmark
France
Luxembourg
Italy
Slovak Republic
Croatia
Sweden
Lithuania
Netherlands
Romania
Ireland
Germany
United Kingdom
-.05
0
.05
2015-2006 change in % females for math intensive careers


Cyprus
Malta
Bulgaria
Portugal
Poland
France
United Kingdom
Italy
Germany
Slovenia
Lithuania
Spain
Hungary
EU26
Ireland
Latvia
Estonia
Netherlands
Sweden
Belgium
Luxembourg
Austria
Denmark
Slovak Republic
Czech Republic
Croatia
Greece
Finland
Romania
-.1
-.05
0
.05
.1
2015-2006 change in % males interested in life science careers


Cyprus
Malta
Portugal
France
Netherlands
Ireland
Austria
Hungary
Italy
Spain
Luxembourg
Poland
United Kingdom
Germany
EU26
Sweden
Slovenia
Bulgaria
Belgium
Estonia
Greece
Slovak Republic
Romania
Denmark
Czech Republic
Lithuania
Latvia
Croatia
Finland
0
.05
.1
.15
2015-2006 change in % females interested in life science careers

Summary

- **In this kind of complex comparison even the presentation of descriptive statistics poses challenges**
- **Underlying computations and models complex yet results should be accessible to non-technical audiences**
- **The challenge of retaining as many comparative angles as possible in each figure: by gender, by type of science, by year but no clutter!**
- **Later the same challenge to report marginal effects for particular individual student predictors, school characteristics and country level characteristics (use margins with repest, but margins is not easy to use with multiple imputations in this environment (i.e. plausible values in estimations)**
- **So far key our findings:**
- **Large gender occupational expectations gap that favours boys in mathematically intensive occupations and girls in life science occupations persists over time**
- **Yet, over time more adolescent girls in Europe think they will pursue life science careers. Not likely they will take up engineering or computing instead**
- **The gender gap is mostly not explained by student school performance, family background, school characteristics or country features. Some predictors matter but only marginally. This was the case in 2006 and remains the case today…..**