Too Much or Not Enough? Assessing Quality with Weighting Data in Cross-sectional Surveys Insights from the European Values Survey

3rd 3MC Conference

Chicago, 28th July 2016

Frédéric GONTHIER

Sciences Po Grenoble, University of Grenoble, France
Associate Professor in Political Science
Head of the Master Program – Opinion, Market and Media Studies
frederic.gonthier@iepg.fr

SCIENCES PO Grenoble

A new look into the issue of data quality improvement

- Data quality is usually addressed by scrutinizing the first steps of the survey life cycle.
  - coverage errors, sampling errors and non-response errors.
- Post-survey adjustements are less used.
  - They can provide insights for cross-national research.
- Weighting can be a relevant indicator
  - Some insights from the *European Values Survey (EVS, 1981-2008)*

The European Values Survey

- Longitudinal survey research program on human values conducted every nine years since 1981 (next wave in fall 2017)
- Pooled dataset of forty-six countries (corresponding to 47 statistical units, with a split between East & West Germany)
  - N= 157,849
- Sampling guidelines and methodological procedures are almost the same (except in 1981, some quota samples were accepted)
  - E.g. representative multi-stage or stratified random sample of the adult population of the country 18 years old and older was used in 2008 (except Armenia and Finland).
  - E.g. in 2008, the net sample size (completed interviews) is about 1500 respondents per country.
- Weight variable (integrated dataset available from Gesis website)
  - Adjusting for gender and age categories' distributions in the national populations.
  - Marginally correcting for the disproportional sample size of the regions in Germany (East and West) and Belgium (Brussels capital region, Flanders and Walloon region).

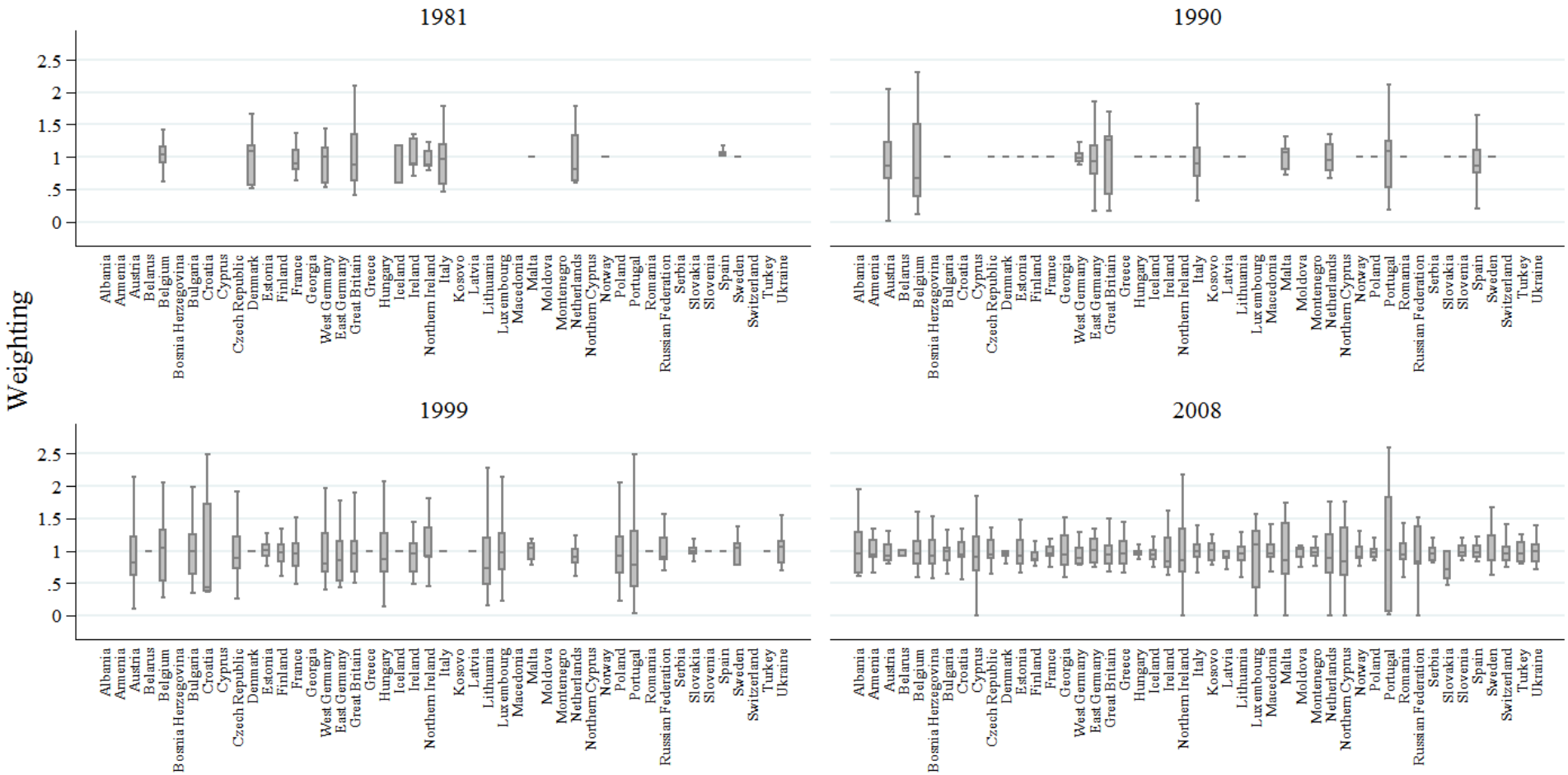
Research question & outline of the presentation

- **Research question : *Has the quality of the EVS data increased since 1981? If so, what are its main drivers?***
- Part I: Data & descriptive results: *Disentangling over- and under-weighting*
- Part II: Methodological design: *Modeling the drivers of data quality*
- Conclusions and discussion

Research question & outline of the presentation

- Research question : *Has the quality of the EVS data increased since 1981? If so, what are its main drivers?*
- **Part I: Data & descriptive results: *Disentangling over- and under-weighting***
- Part II: Methodological design: *Modeling the drivers of data quality*
- Conclusions and discussion

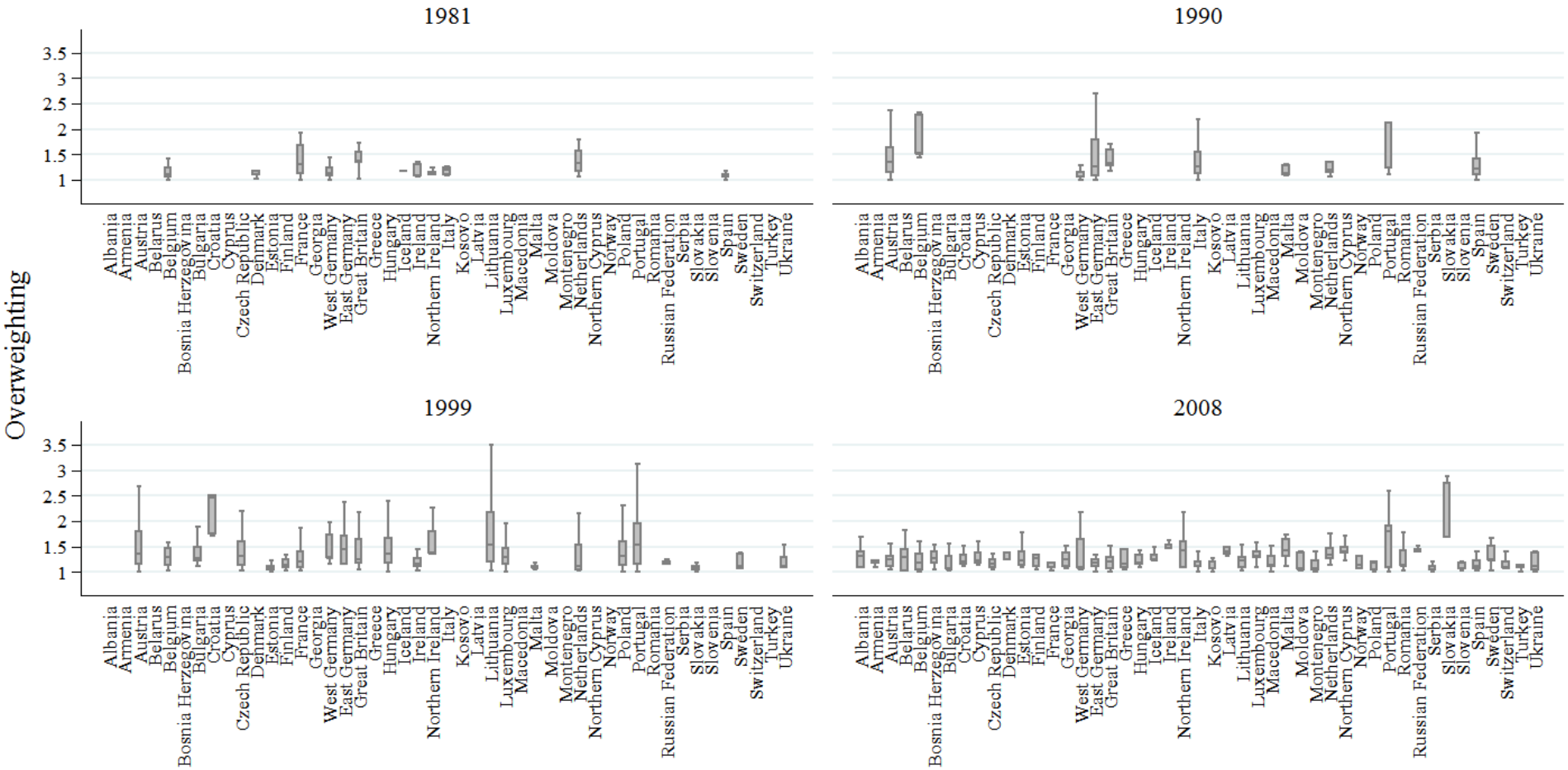
Descriptive overview: Weighting by country



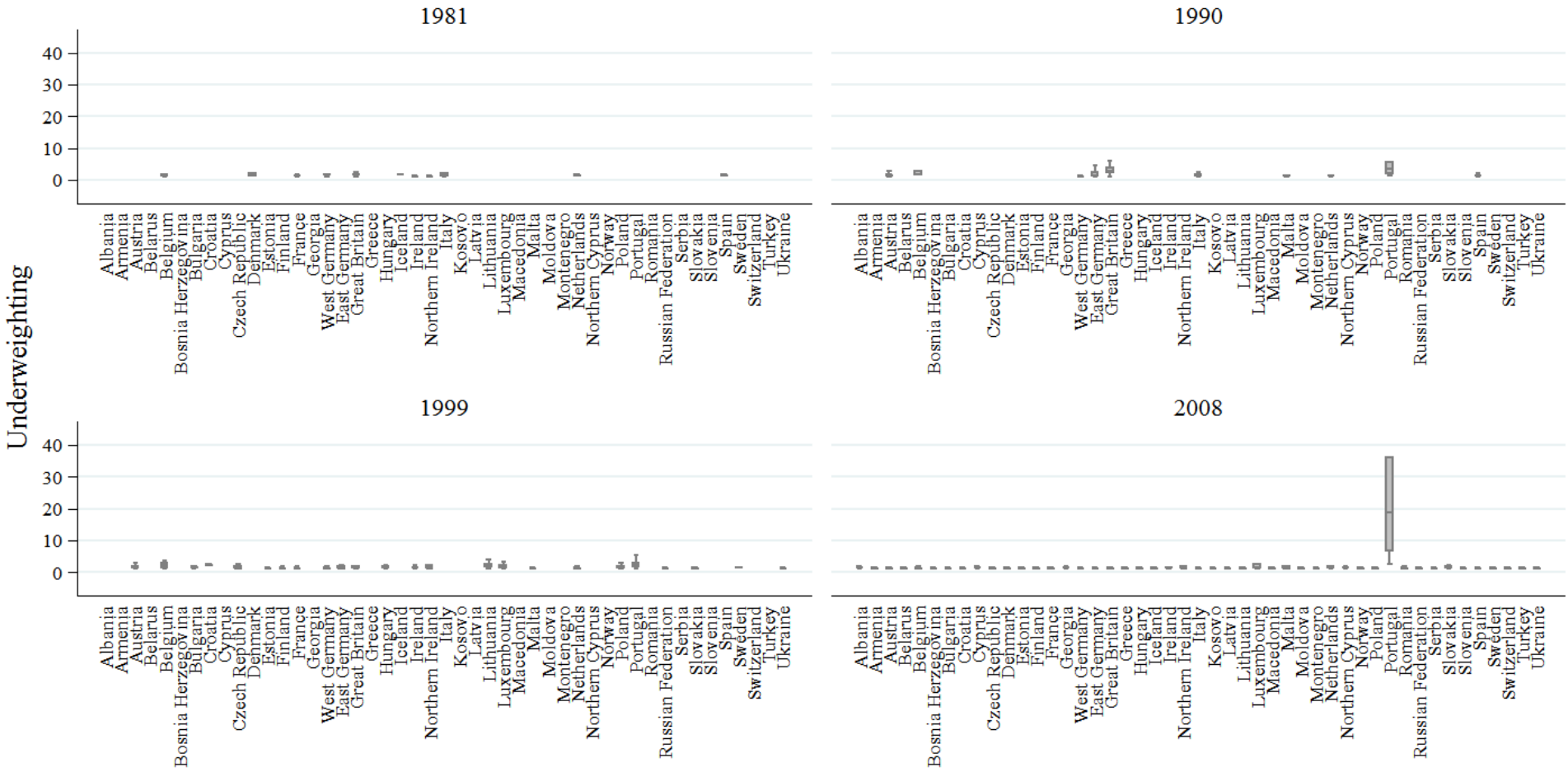
Disentangling over- and under-weighting

- Using the weight variable as a whole may be misleading.
  - Overweighting (i.e., some individuals are under-represented) is a sign of poor quality.
  - Underweighting (i.e., some individuals are over-represented) does not forcefully point to some degraded data.
- New look into data quality with assessing separately the two sides of the distribution of the weight variable.
- Two separate EVS sub-samples.
  1. Weights for the over-represented respondents, theoretically ranging from 1 to +∞ (n=53,532)
  2. Weights for the under-represented respondents, with a rescaled weighting variable from 1 to +∞ (1/weight) to facilitate interpretation (n=71,537)

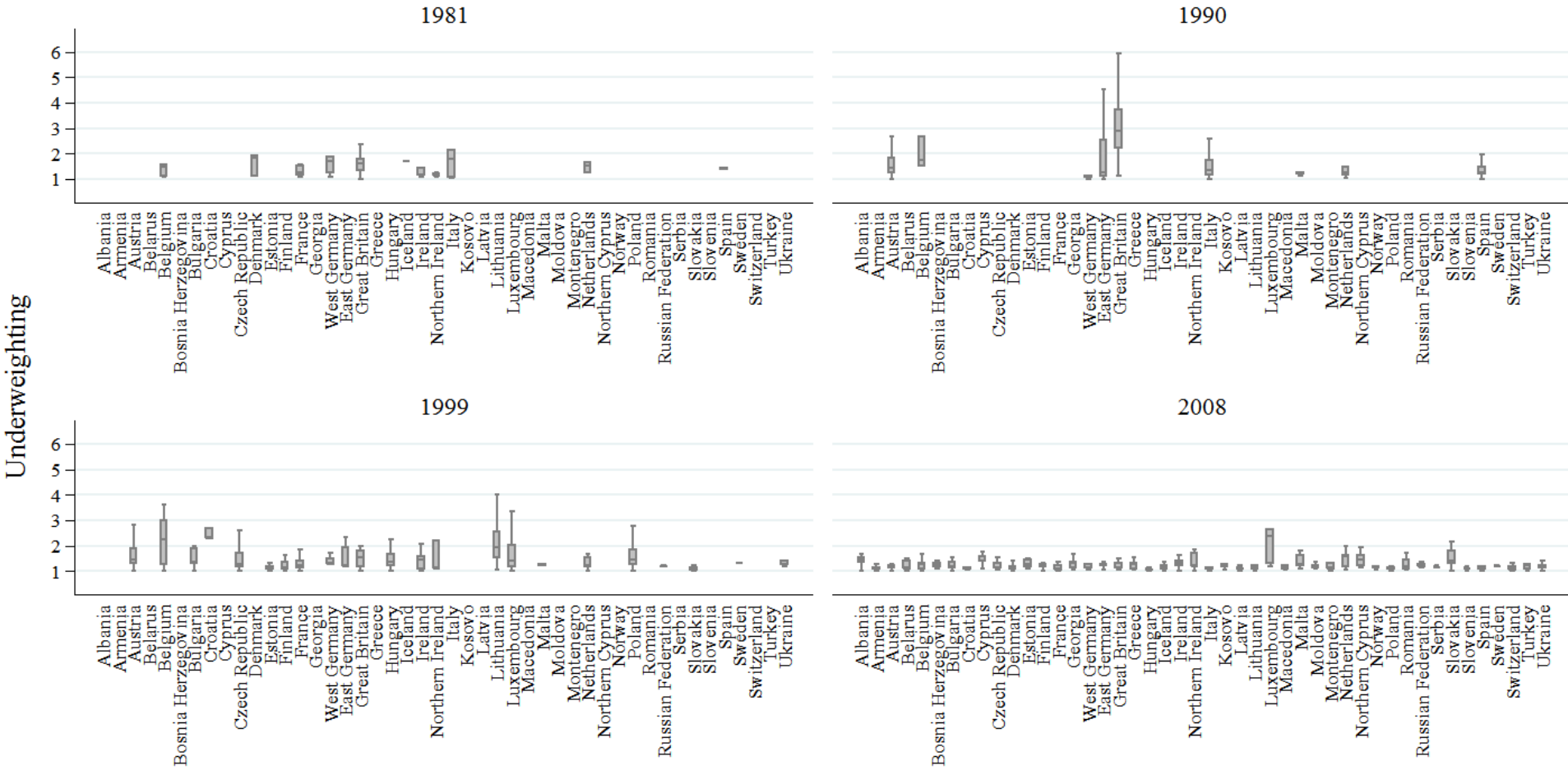
Overweighting by country (the closer to 1, the better)



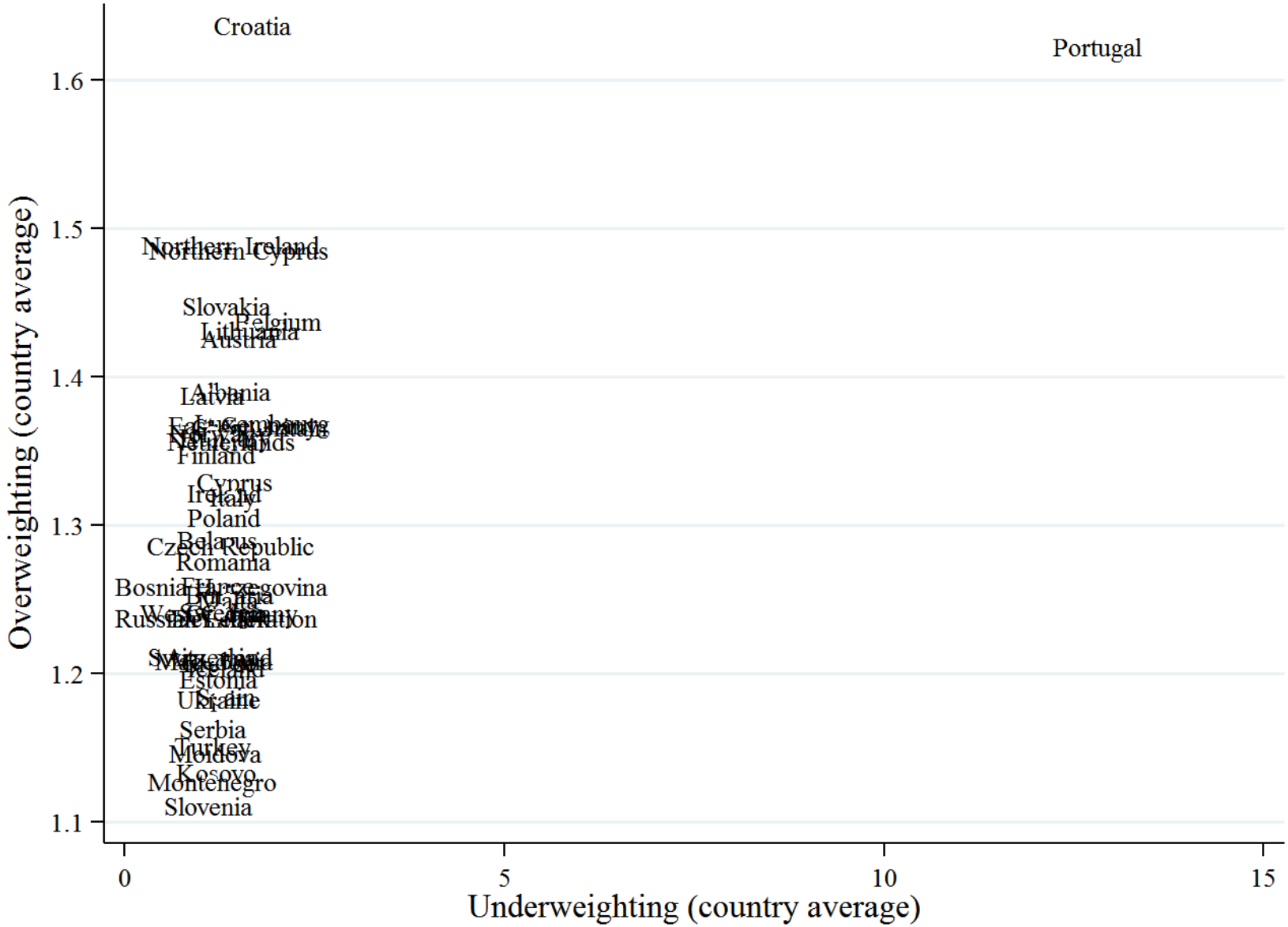
Underweighting by country (the closer to 1, the better)



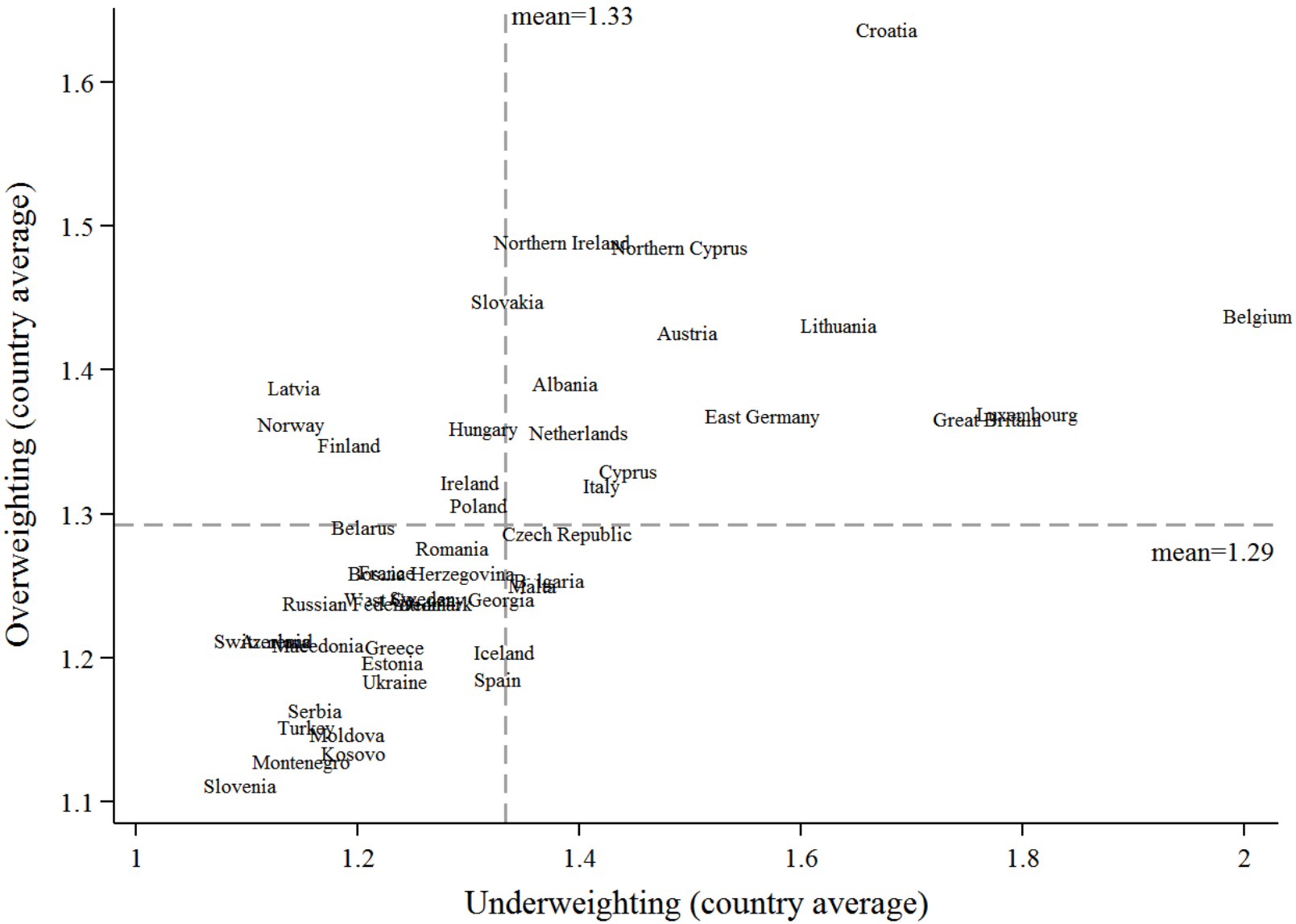
Underweighting by country… without Portugal



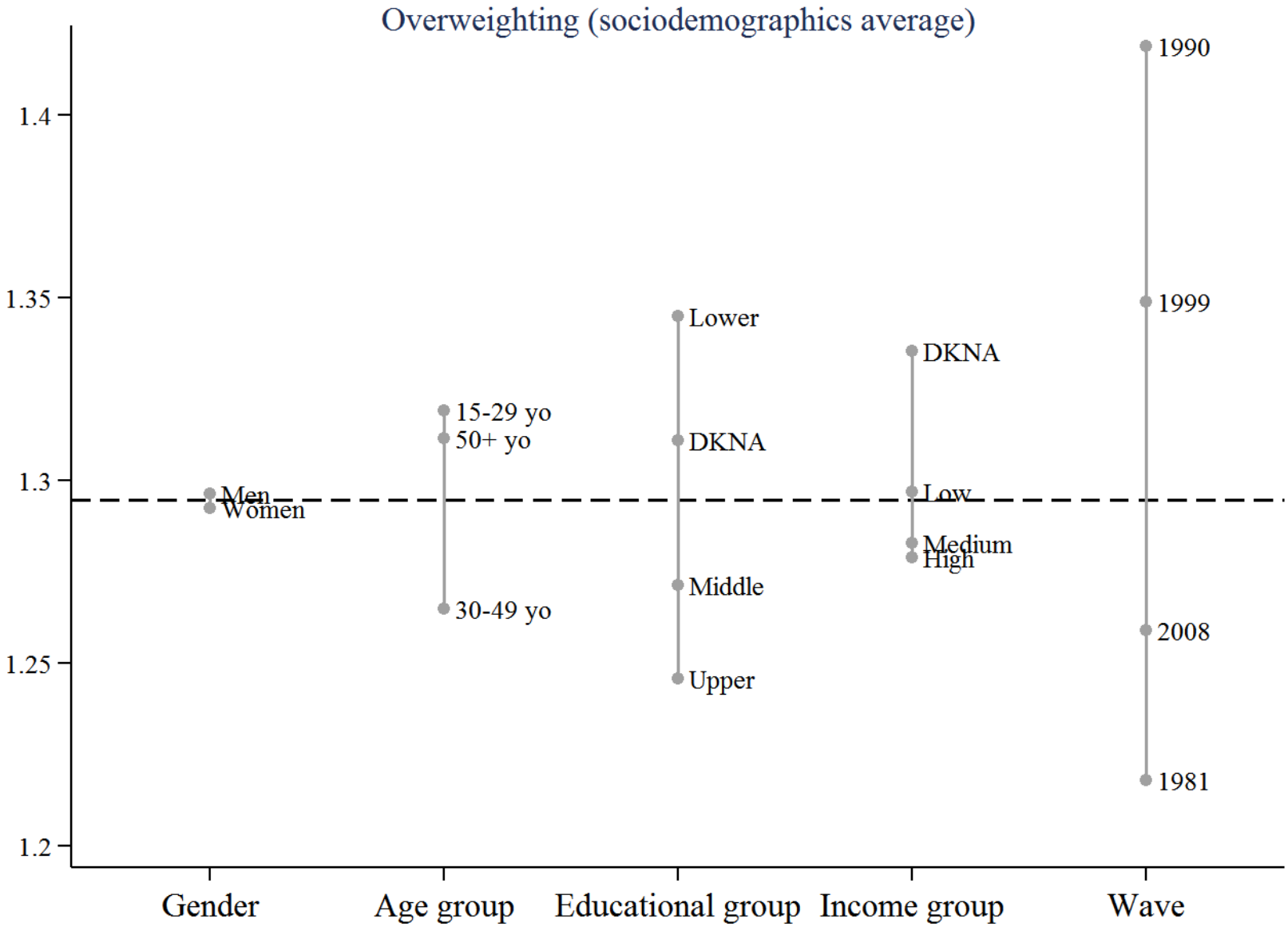
How countries perform (the lower, the better) (1/2)



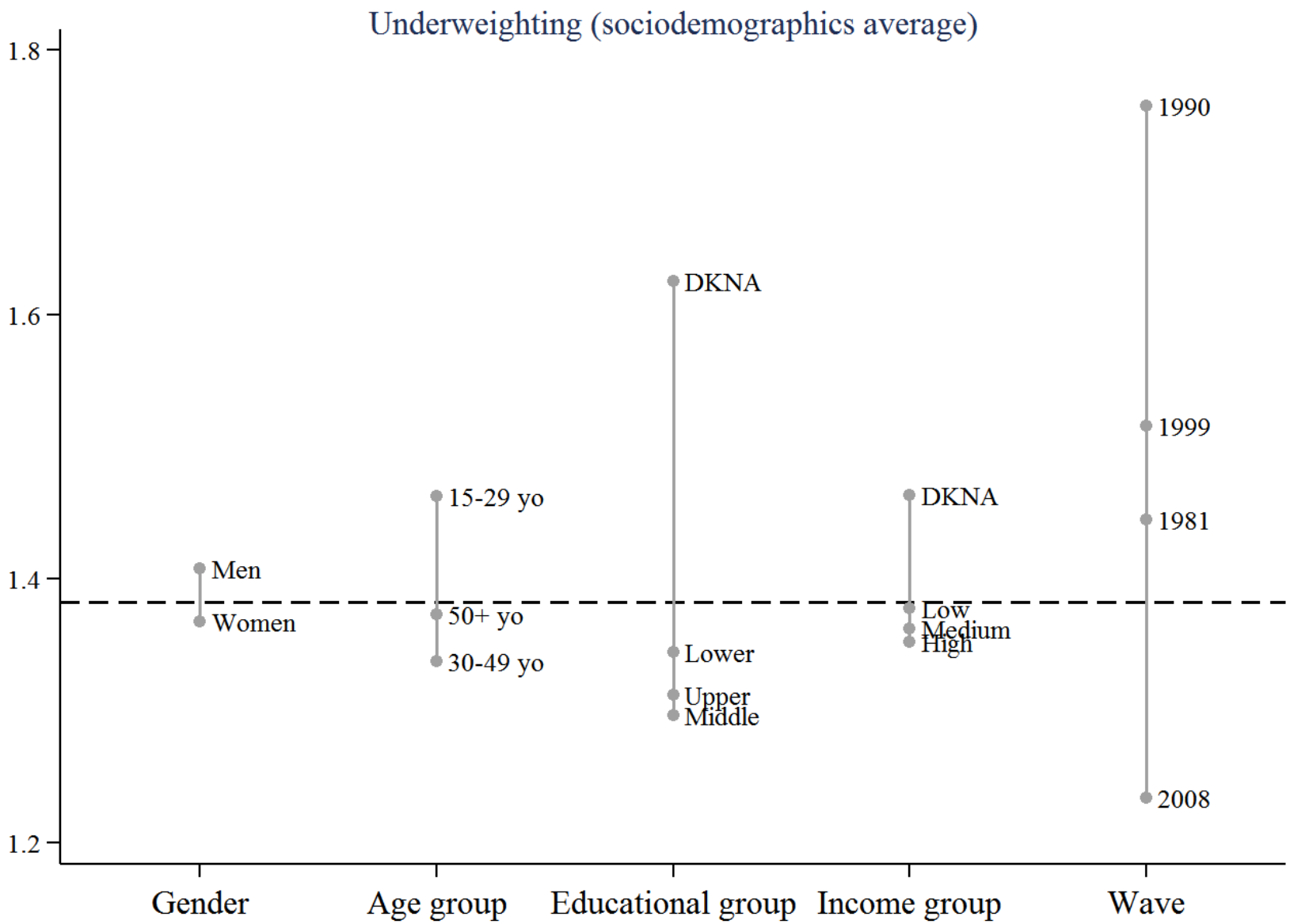
How countries perform... without Portugal (2/2)



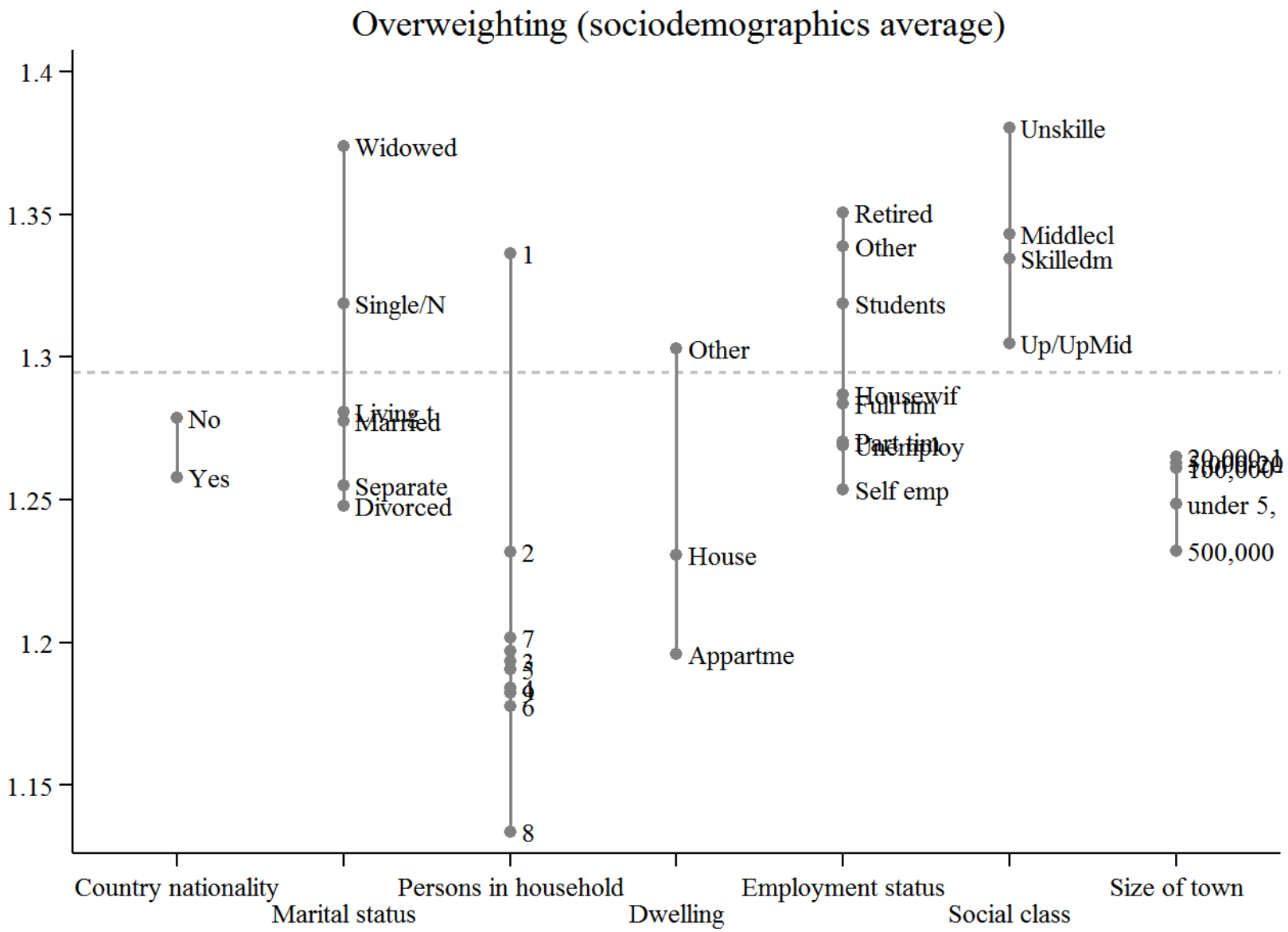
Overweighting: The usual suspects



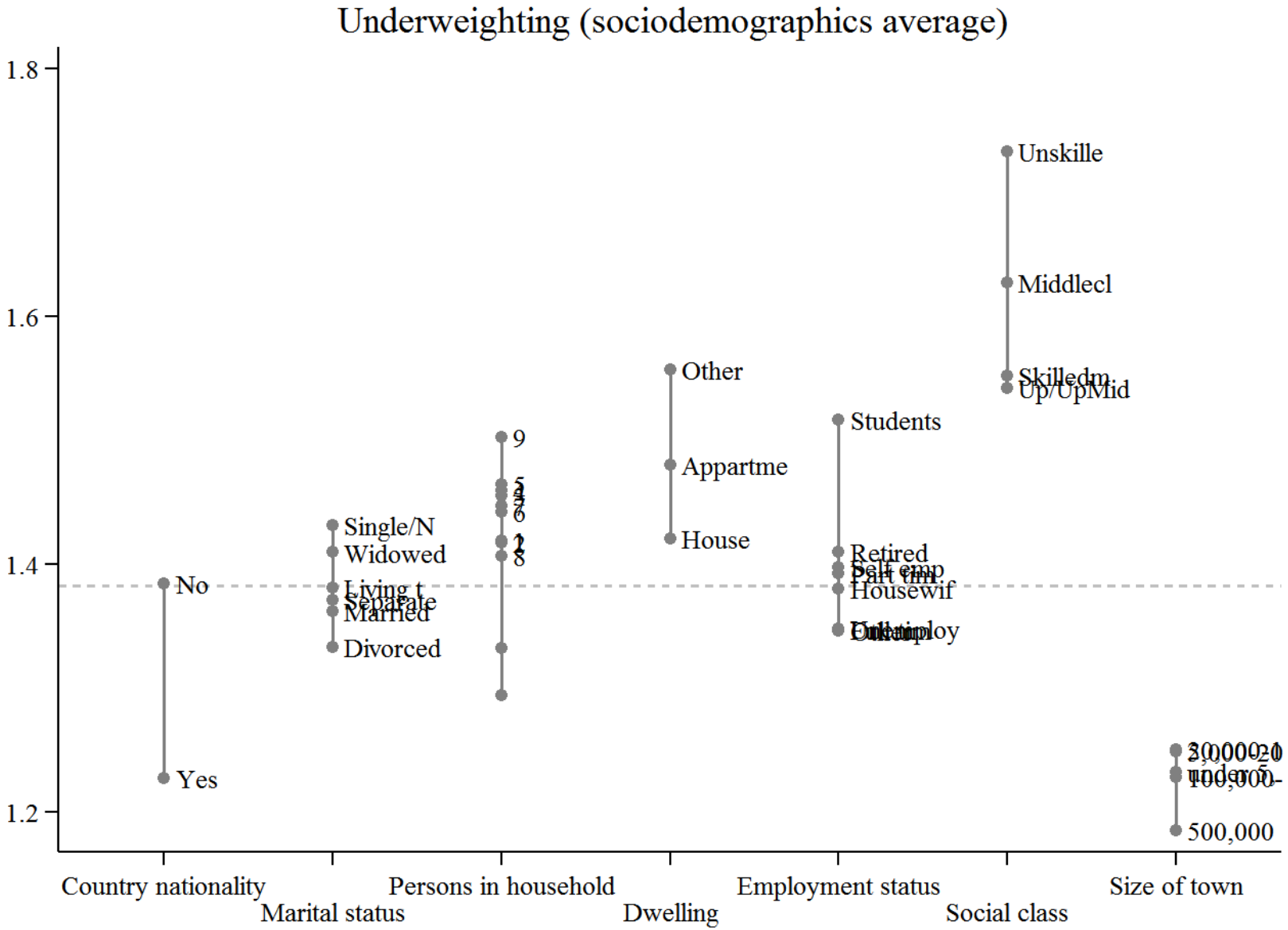
Underweighting: The usual suspects



Overweighting: New suspects



Underweighting: New suspects



Research question & outline of the presentation

- Research question : *Has the quality of the EVS data increased since 1981? If so, what are its main drivers?*
- Part I: Data & descriptive results: *Disentangling over- and under-weighting*
- **Part II: Methodological design: *Modeling the drivers of data quality***
- Conclusions and discussion

Methodological design

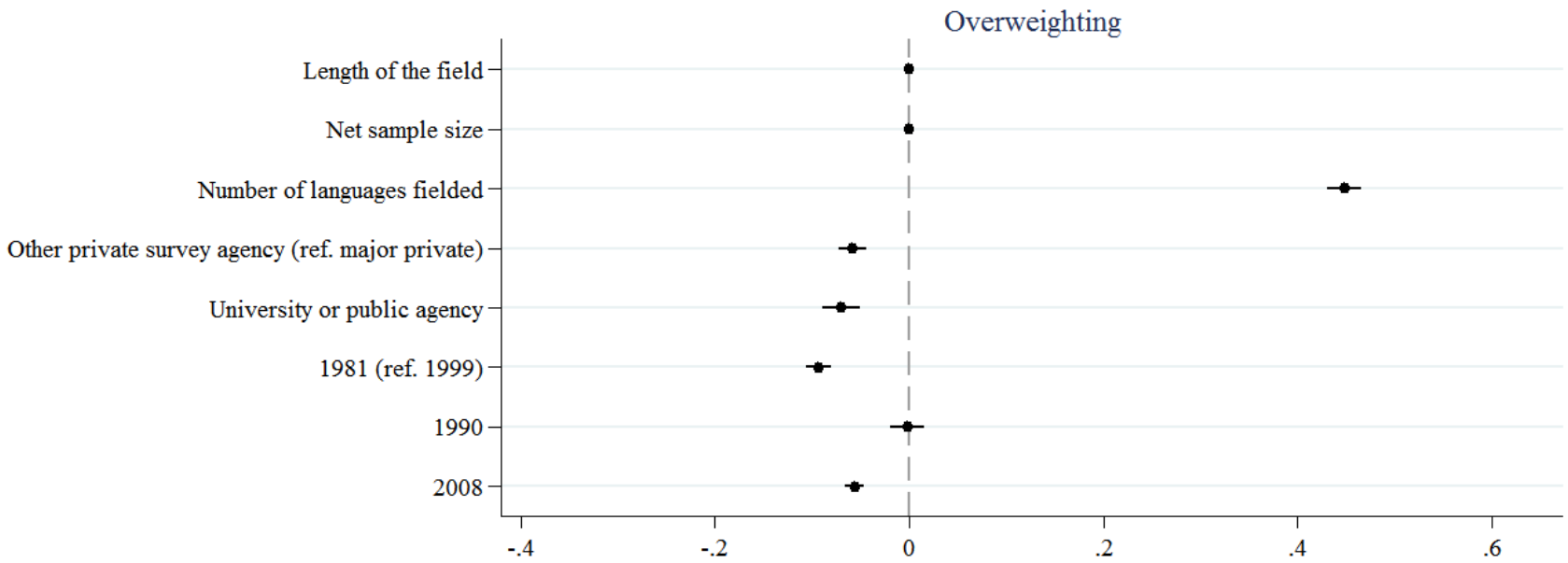
- <u>Dependent variables:</u>
  - Over- & underweighting
- <u>Modeling strategy</u>:
  - Multilevel analysis
    - Two-level model: individuals (level 1) nested in countries (level 2)
    - 46 countries (Portugal excluded)
  - Independent variables (macro-level) from EVS technical reports:
    - 1981-2008:
      - length of the field (days); net sample size; data collector (university; major private company, i.e., Gallup, GfK, Ipsos, TNS; other private company)
    - 2008:
      - number of interviewers, % experienced interviewers, % trained interviewers, drafting written instructions, advance letter to the respondents, incentive, type of mode (PAPI or CAPI), contact rate, refusal rate…

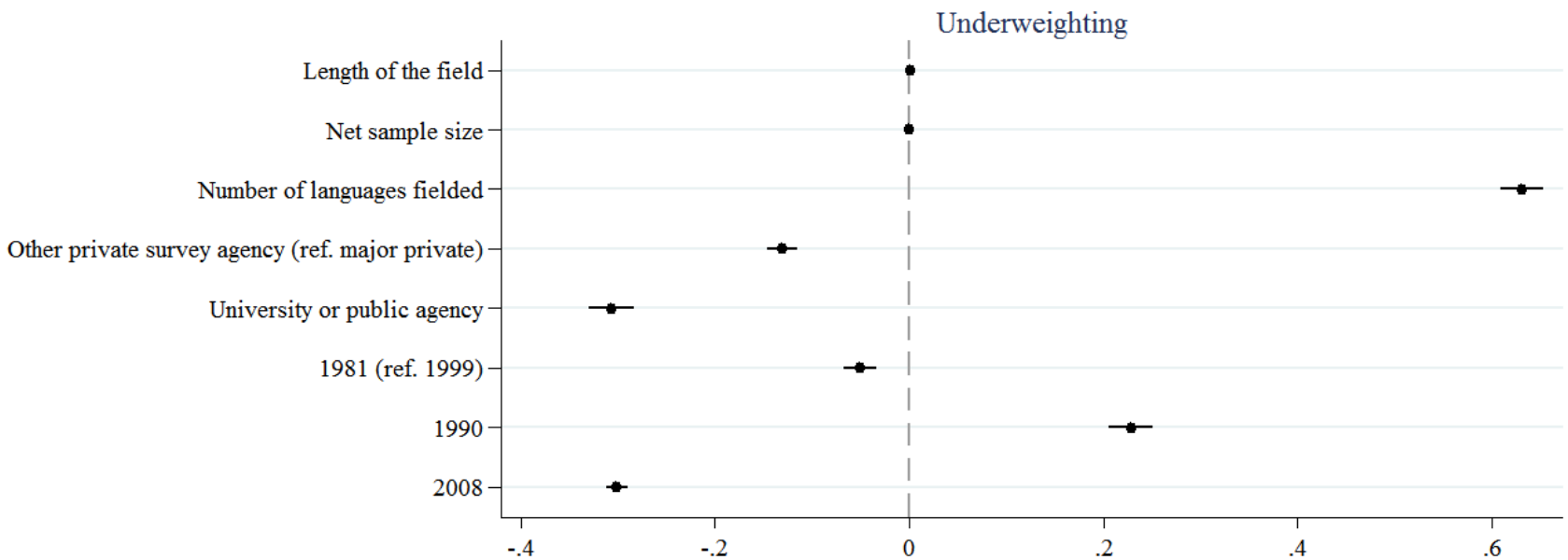
Multilevel analysis of over- and underweighting (2008-2012; 46 countries)

	Null Model Overweighting		Model 1 Overweighting		Null Model Underweighting		Model 1bis Underweighting	
Level 2 variables	Exp(ß)	SE	Exp(ß)	SE	Exp(ß)	SE	Exp(ß)	SE
Length of the field			**-.00071*** **	.0000711			**.00081*** **	.0000832
Net sample size			**-.00012*** **	6.97e-06			**-.00015*** **	8.18e-06
Number of languages fielded			**.44799*** **	.0091001			**.63132*** **	.0113534
Other private compagny (ref. major private)			**-.05834*** **	.0070981			**-.13099*** **	.0079008
Public agency			**-.06943*** **	.0096399			**-.30688*** **	.0118492
1981 (ref. 1999)			**-.09331*** **	.0066826			**-.05029*** **	.0085415
1990			-.00161	.0089322			**.22832*** **	.0112955
2008			**-.05639*** **	.0049119			**-.30095*** **	.0055751
Intercept	1.292354***	.0162659	1.033681***	.0405937	1.33331***	.0286171	.86664***	.0543614
Random effects	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Variance level 1 (individual)	.1355895	.0008443	.1242126	.0007779	.0348237	.0002492	.1249494	.0264439
Variance level 2 (country)	.011996	.0025428	.0689028	.0145915	.0374038	.0078418	4.344952	.9839192
ICC	.0812815		.3567961		.1085809		.3465044	
N level 1 (individual)	51629		51043		69702		68828	
N level 3 (country)	46		46		46		46	
AIC	(43565.3)		38699.42		(115751.1)		96164.04	
BIC	(43591.86)		38796.66		(115778.5)		96264.58	

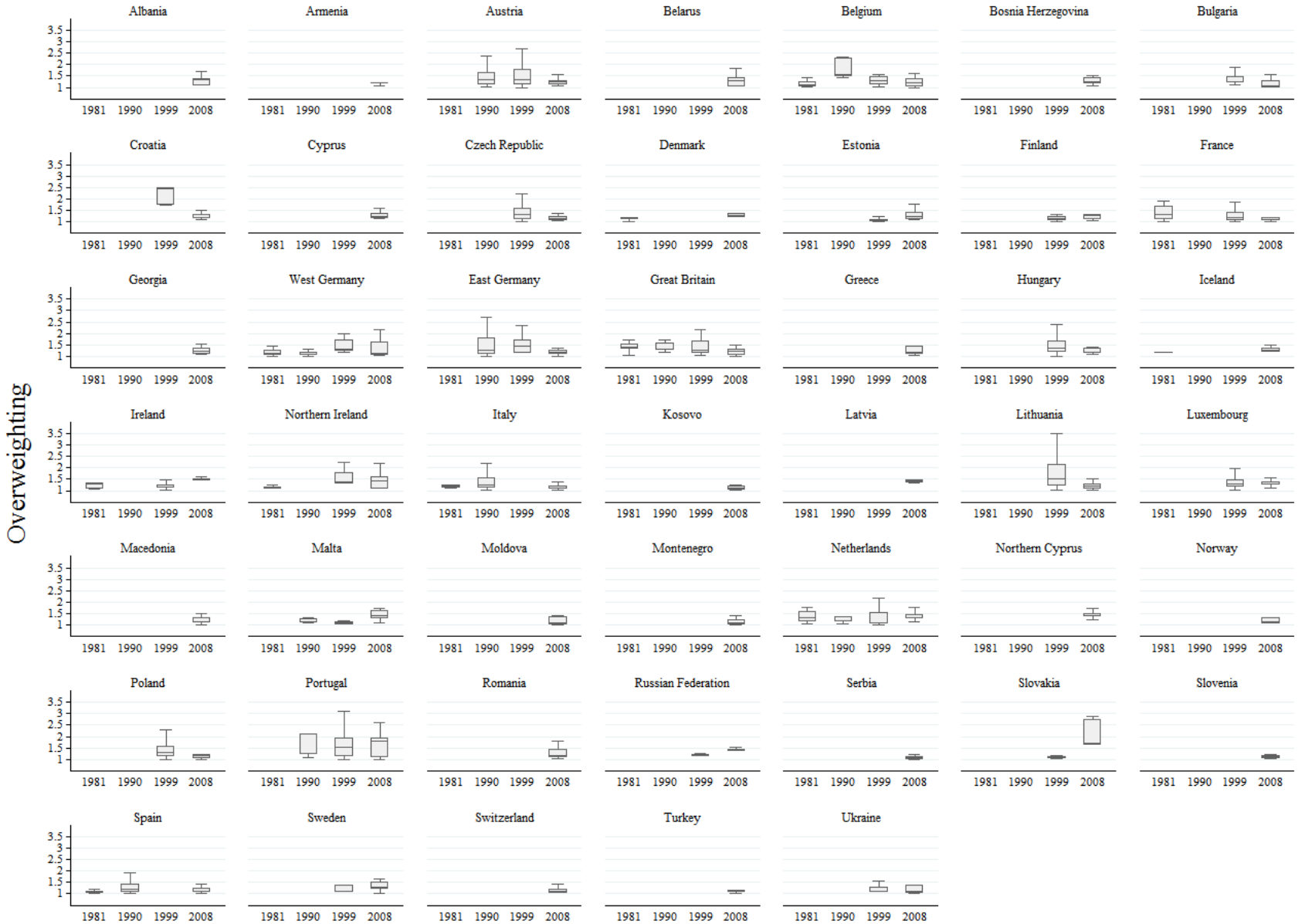
Note : SE=standard error. Sig.: *** p<0.01, ** p<0.05, * p<0.1. ICC=Intraclass Correlation Coefficient. AIC (Akaike Information Criterion) BIC (Bayesian Information Criterion)

Expected and unexpected drivers of data quality

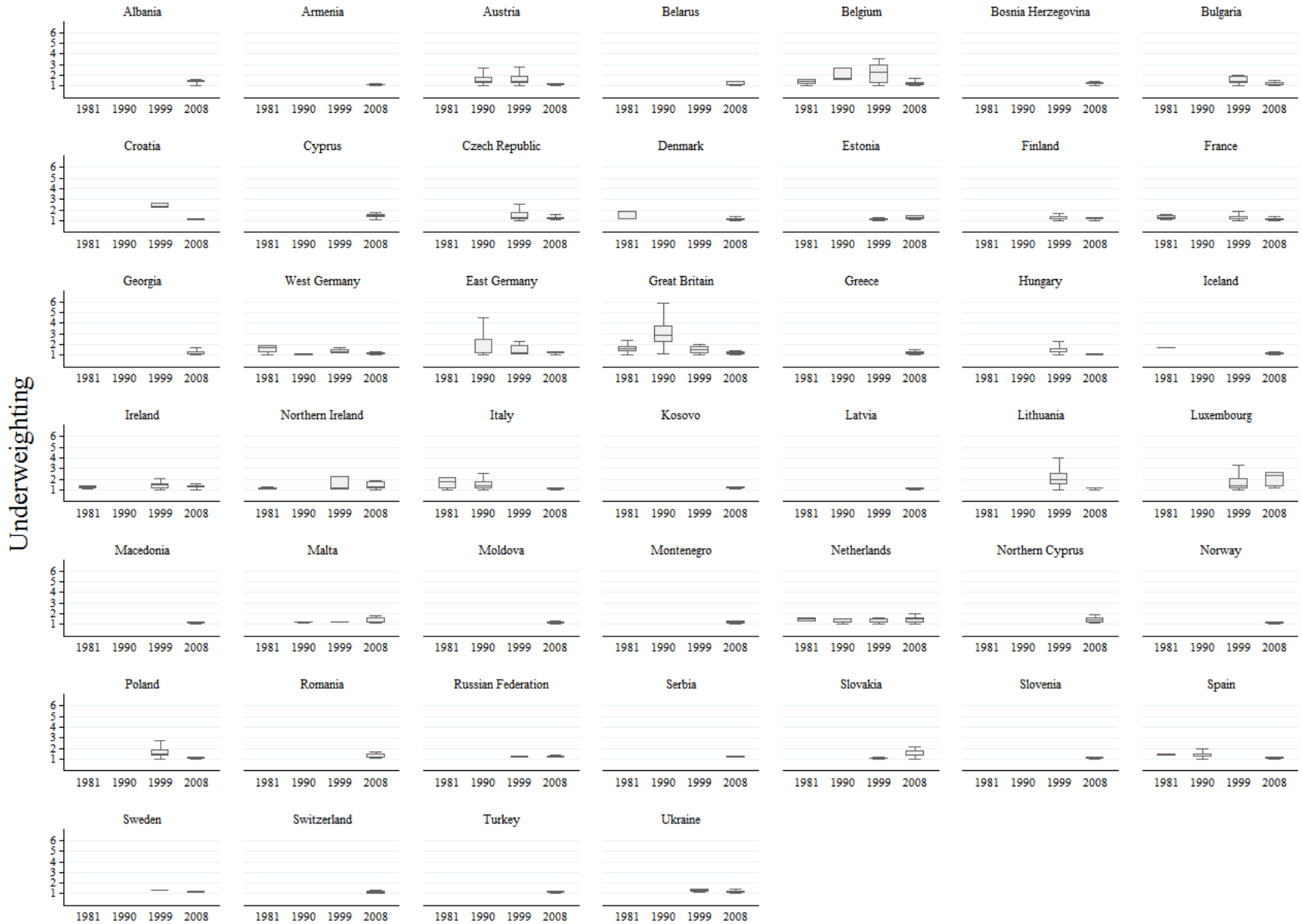




Converging overweighting quality among EVS countries



Converging underweighting quality among EVS countries

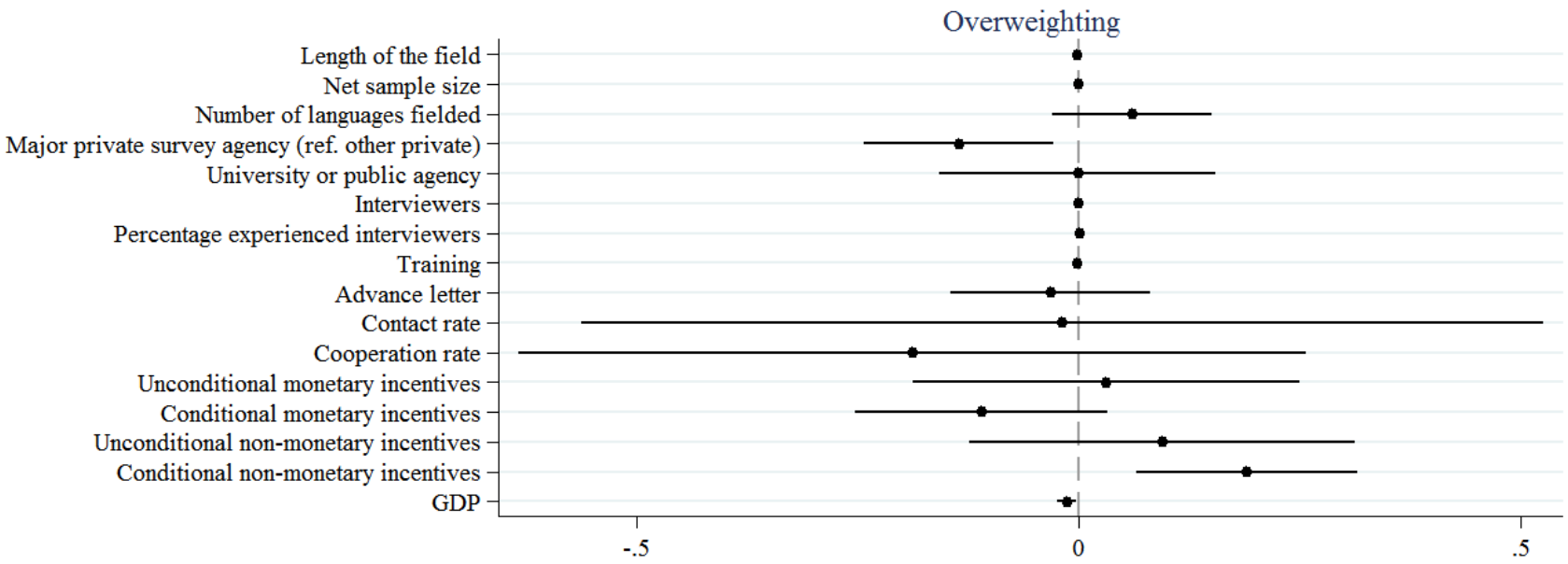


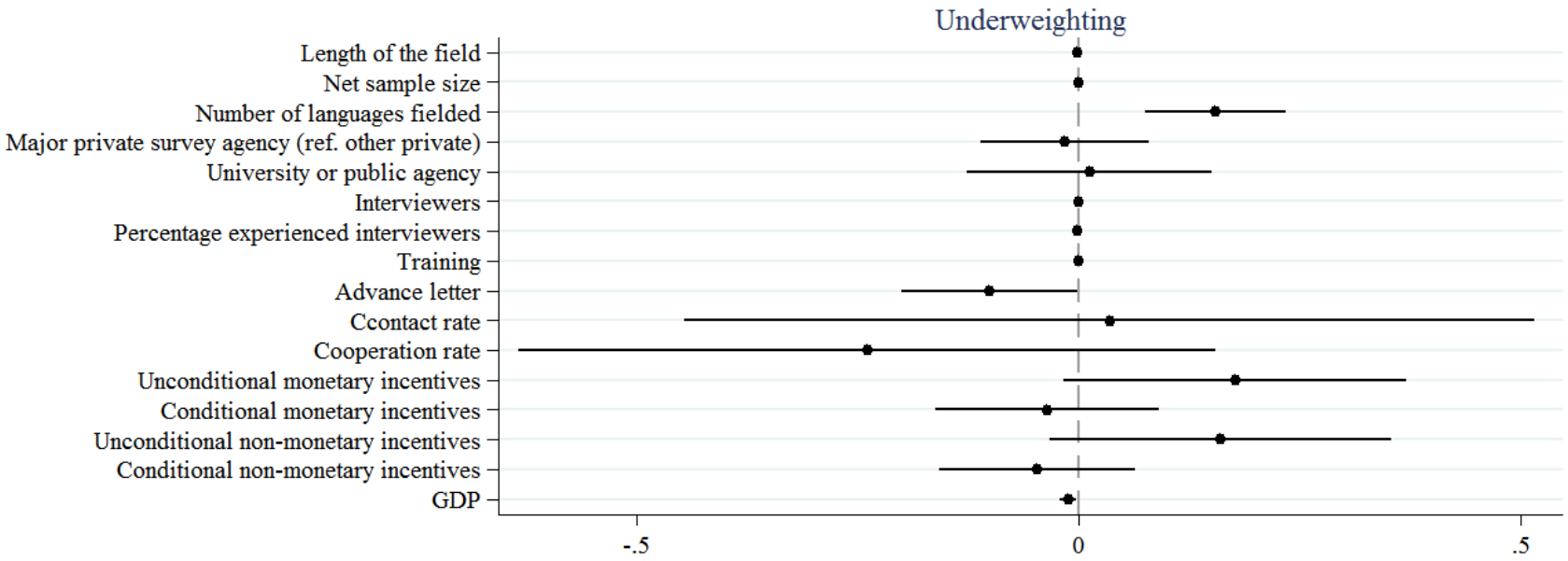
Multilevel analysis of over- and underweighting (2008; 43 countries)

	Model 2 Overweighting		Model 2bis Underweighting	
Level 2 variables	Exp(ß)	SE	Exp(ß)	SE
Length of the field	-.0006633	.0005295	.0004688	.0004688
Net sample size	-.0001087	.000076	-.0000639	.0000673
Number of languages fielded	.0608356	.0459189	**.15545*** **	.0406501
Other private compagny (ref. major private)	**-.13507** **	.0548844	-.0155287	.0485956
Public agency	-.00041	.0797097	.0280569	.0705942
Interviewers	.0004657	.0005041	.0125282	.000446
Percentage experienced interviewers	.0011484	.0011009	-.0005873	.0009749
Training	-.0007906	.0006646	-.00021	.0005886
Advance letter	-.0311358	.0574559	**-.10039* **	.0509069
Contact rate	-.0175435	.2777356	.0352016	.2452595
Cooperation rate	-.1876166	.2272809	-.238296	.2012323
Unconditional monetary incentives	.0319635	.1116253	.1776576	.0988898
Conditional monetary incentives	-.1093997	.0726272	-.0346319	.0642972
Unconditional non monetary incentives	.0950045	.1113056	.1612207	.098641
Conditional non monetary incentives	**.19088*** **	.0641263	-.0464718	.056775
GDP	**-.01271** **	.005211	**-.01144** **	.0046148
Intercept	1.829036***	.3204204	1.784317***	.2859714
Random effects	Estimate	SE	Estimate	SE
Variance level 1 (individual)	.0760328	.0007007	.0366425	.0002725
Variance level 2 (country)	.0164683	.0035914	.0129966	.0028134
ICC level 2 (country x wave)	.1780337		.2618213	
N level 1 (individual)	23589		36208	
N level 3 (country)	43		43	
AIC	6404.516		-16687.75	
BIC	6557.818		-16526.31	

Note: SE=standard error. Sig.: *** p<0.01, ** p<0.05, * p<0.1. ICC=Intraclass Correlation

No obvious methodological driver of data quality in 2008





Research question & outline of the presentation

- Research question : *Has the quality of the EVS data increased since 1981? If so, what are its main drivers?*
- Part I: Data & descriptive results: *Disentangling over- and under-weighting*
- Part II: Methodological design: *Modeling the drivers of data quality*
- **Conclusions and discussion**

Conclusions (1/2)

- The main drivers of data quality (as regards gender & age)
  - + Large sample sizes, survey fielded by academics and national private agencies tend to deflate both under- and over-representation (i.e. over- & underweighting)
  - + Length of the field reduces under-representation (i.e. overweighting)
  - + Advance letters lower over-representation (i.e. underweighting)
  - – The number of languages fielded increases both under- and over-representation (i.e. over- & underweighting)
  - – Conditional non monetary incentives have a detrimental effect on over-weighting (i.e. they increase underrepresentation)

Conclusions (2/2)

- No obvious methodological driver in 2008
  - As far as gender and age are concerned, the quality of collected data has increased over time.
  - EVS countries seem to have reached similar quality levels.
  - Survey attitudes and traditions, in terms of ease of contact and willingness to cooperate, do not seem to be that influential.
- Yet, the effect of GDP on over- & under-weighting suggests that substantial differences remain between countries.
  - Survey climate may have an impact on data quality, but less through response propensity than through the quality of "national infrastructures" for data collection.


That's all Folks!