

# The survey fingerprint puzzle of life satisfaction

**Thomas Beuchot<sup>1</sup> & Alberto Prati<sup>2</sup>**

<sup>1</sup> Ecole Normale Supérieure

<sup>2</sup> University College London

October 2023

- ▶ Life satisfaction (LS) is measured via some simple questions such as "Everything considered, how satisfied are you with your life?" [0-10].

# Life satisfaction

- ▶ Life satisfaction (LS) is measured via some simple questions such as "Everything considered, how satisfied are you with your life?" [0-10].
- ▶ It is increasingly proposed as a welfare measure.

# Life satisfaction

- ▶ Life satisfaction (LS) is measured via some simple questions such as "Everything considered, how satisfied are you with your life?" [0-10].
- ▶ It is increasingly proposed as a welfare measure.
- ▶ More than 70% of OECD countries have national frameworks with a wellbeing focus. Among them, 85% use LS.

# Life satisfaction

- ▶ Life satisfaction (LS) is measured via some simple questions such as "Everything considered, how satisfied are you with your life?" [0-10].
- ▶ It is increasingly proposed as a welfare measure.
- ▶ More than 70% of OECD countries have national frameworks with a wellbeing focus. Among them, 85% use LS.
- ▶ LS is measured in (almost) every country in the world every year.

# Different surveys give different answers

- ▶ Who is happier, the Russian or the Kyrgyz?

# Different surveys give different answers

- ▶ Who is happier, the Russian or the Kyrgyz?
- ▶ The answer is not straightforward because different surveys can give different answers → the **survey fingerprint puzzle**.

# Different surveys give different answers

- ▶ Who is happier, the Russian or the Kyrgyz?
- ▶ The answer is not straightforward because different surveys can give different answers → the **survey fingerprint puzzle**.
- ▶ Today's talk is about this puzzle and its potential explanations.



# Different surveys give different answers

- ▶ Who is happier, the Russian or the Kyrgyz?
- ▶ The answer is not straightforward because different surveys can give different answers → the **survey fingerprint puzzle**.
- ▶ Today's talk is about this puzzle and its potential explanations.
- ▶ Another example: the UK.

# Different surveys give different answers

- ▶ Who is happier, the Russian or the Kyrgyz?
- ▶ The answer is not straightforward because different surveys can give different answers → the **survey fingerprint puzzle**.
- ▶ Today's talk is about this puzzle and its potential explanations.
- ▶ Another example: the UK.
- ▶ On a scale from 0 to 10, what was the level of life satisfaction in the UK in the winter 2018-19?

# Different surveys give different answers

- ▶ Who is happier, the Russian or the Kyrgyz?
- ▶ The answer is not straightforward because different surveys can give different answers → the **survey fingerprint puzzle**.
- ▶ Today's talk is about this puzzle and its potential explanations.
- ▶ Another example: the UK.
- ▶ On a scale from 0 to 10, what was the level of life satisfaction in the UK in the winter 2018-19?
- ▶ ...It depends.
  - ▶ 7.28 [C.I. 7.19; 7.37] according to the European Social Survey
  - ▶ 7.71 [C.I. 7.70; 7.72] according to the UK Annual Pop. Survey

# The puzzle

- ▶ How can life satisfaction estimates be so different?

# The puzzle

- ▶ How can life satisfaction estimates be so different?
- ▶ These surveys:
  - ▶ are representative of the same population
  - ▶ use the same psychometric method (question + scale)
  - ▶ were run over the same period.

# The puzzle

- ▶ How can life satisfaction estimates be so different?
- ▶ These surveys:
  - ▶ are representative of the same population
  - ▶ use the same psychometric method (question + scale)
  - ▶ were run over the same period.
- ▶ There is nothing strange in LS suffering from measurement error. All economic variables do, including GDP (Chang & Li, 2018).

# The puzzle

- ▶ How can life satisfaction estimates be so different?
- ▶ These surveys:
  - ▶ are representative of the same population
  - ▶ use the same psychometric method (question + scale)
  - ▶ were run over the same period.
- ▶ There is nothing strange in LS suffering from measurement error. All economic variables do, including GDP (Chang & Li, 2018).
- ▶ But it is important to understand its extent and structure to handle it properly.

# The puzzle

- ▶ How can life satisfaction estimates be so different?
- ▶ These surveys:
  - ▶ are representative of the same population
  - ▶ use the same psychometric method (question + scale)
  - ▶ were run over the same period.
- ▶ There is nothing strange in LS suffering from measurement error. All economic variables do, including GDP (Chang & Li, 2018).
- ▶ But it is important to understand its extent and structure to handle it properly.
- ▶ To this purpose, we use a **meta-analytical** framework (10 surveys, >2 million respondents).



# Datasets

We compile an exhaustive list of available **representative** surveys which use **comparable** life satisfaction measures from 2000 to today.

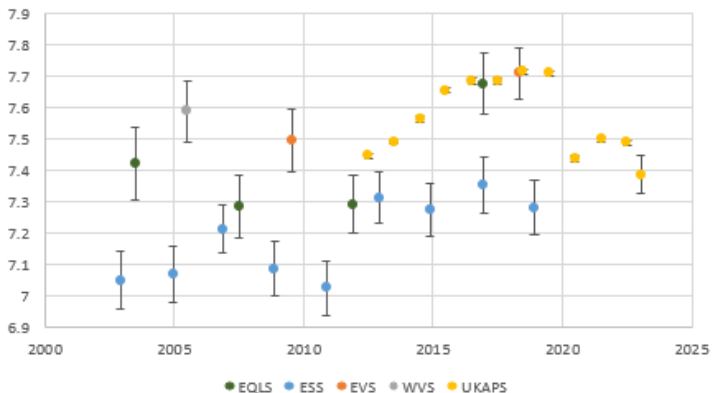
- ▶ Gallup World Poll
- ▶ US Gallup Social Survey
- ▶ PEW Global Attitudes Project
- ▶ European Quality of Life Survey
- ▶ European Social Survey
- ▶ European Value Survey
- ▶ World Value Survey
- ▶ French Statistics on Income and Living Conditions (SRCV)
- ▶ French Consumer Confidence Survey (CAMME)
- ▶ UK Annual Population Survey
- ▶ *Life in transition survey\**

---

<sup>0</sup>\*For country rankings only.

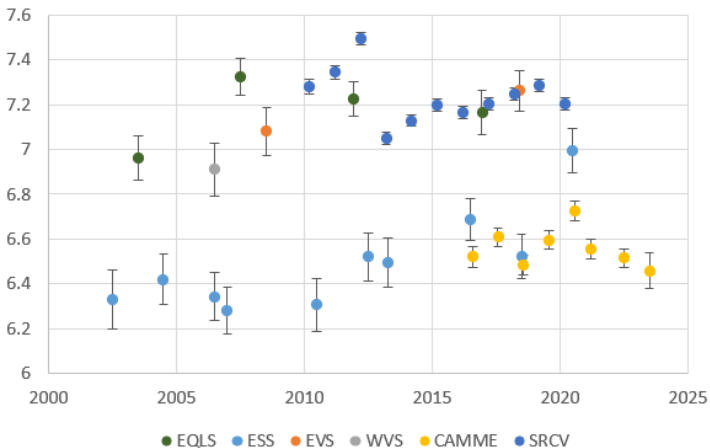
- 1 The puzzle
- 2 A measurement error approach
- 3 Conclusion

Measure: General Life Satisfaction (0-10 or 1-10). Survey-weighted estimated means.

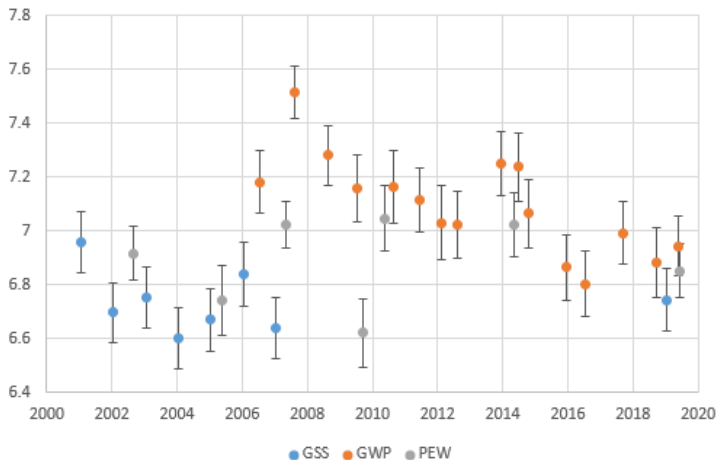


# France

Measure: General Life Satisfaction (0-10 or 1-10). Survey-weighted estimated means.



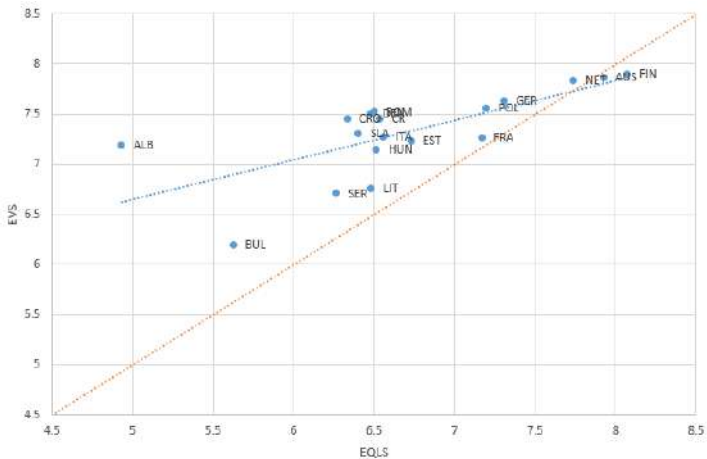
Measure: Cantril Ladder (0-10). Survey-weighted estimated means.



# Europe (EQLS vs EVS)

Measure: LS 1-10. Estimated means.

Period: Aug 16-Mar 17 (EQLS); Jun 17-Jan 19 (EVS).



# The survey fingerprint puzzle

- ▶ High within-survey reliability.
- ▶ Low between-survey reliability.
- ▶ In theory, estimated means should be within the 95% C.I.
- ▶ Often, they are not. Why?

# Outline

1 The puzzle

2 A measurement error approach

3 Conclusion



# How to think about measurement error

Call  $\bar{y}$  the sample average LS. Given two surveys  $p$  and  $q$ , we observe:

$$\bar{y}_p \neq \bar{y}_q \quad (1)$$

# How to think about measurement error

Call  $\bar{y}$  the sample average LS. Given two surveys  $p$  and  $q$ , we observe:

$$\bar{y}_p \neq \bar{y}_q \quad (1)$$

If  $X_i$  is a set of (accurately observed) exogenous determinants of LS, we can decompose LS as follows:

$$\bar{y}_p = \hat{\alpha}_p + \hat{\beta}_p \bar{X}_p \quad (2)$$

$$\bar{y}_q = \hat{\alpha}_q + \hat{\beta}_q \bar{X}_q \quad (3)$$

# How to think about measurement error

Call  $\bar{y}$  the sample average LS. Given two surveys  $p$  and  $q$ , we observe:

$$\bar{y}_p \neq \bar{y}_q \quad (1)$$

If  $X_i$  is a set of (accurately observed) exogenous determinants of LS, we can decompose LS as follows:

$$\bar{y}_p = \hat{\alpha}_p + \hat{\beta}_p \bar{X}_p \quad (2)$$

$$\bar{y}_q = \hat{\alpha}_q + \hat{\beta}_q \bar{X}_q \quad (3)$$

Since  $\bar{y}_p \neq \bar{y}_q$ , at least one of the following must be true:

- ▶  $\bar{X}_p \neq \bar{X}_q \rightarrow$  some groups might be under- or over-sampled.
- ▶  $\hat{\beta}_p \neq \hat{\beta}_q \rightarrow$  relationships are survey-dependent.
- ▶  $\hat{\alpha}_p \neq \hat{\alpha}_q \rightarrow$  there are survey fixed-effects.

# How to think about measurement error

Call  $\bar{y}$  the sample average LS. Given two surveys  $p$  and  $q$ , we observe:

$$\bar{y}_p \neq \bar{y}_q \quad (1)$$

If  $X_i$  is a set of (accurately observed) exogenous determinants of LS, we can decompose LS as follows:

$$\bar{y}_p = \hat{\alpha}_p + \hat{\beta}_p \bar{X}_p \quad (2)$$

$$\bar{y}_q = \hat{\alpha}_q + \hat{\beta}_q \bar{X}_q \quad (3)$$

Since  $\bar{y}_p \neq \bar{y}_q$ , at least one of the following must be true:

- ▶  $\bar{X}_p \neq \bar{X}_q \rightarrow$  some groups might be under- or over-sampled.
- ▶  $\hat{\beta}_p \neq \hat{\beta}_q \rightarrow$  relationships are survey-dependent.
- ▶  $\hat{\alpha}_p \neq \hat{\alpha}_q \rightarrow$  there are survey fixed-effects.

We estimate  $\hat{\alpha}$ ,  $\hat{\beta}$  and  $\bar{X}$  in each of these surveys.

Table: Sample composition ( $\bar{X}$ )

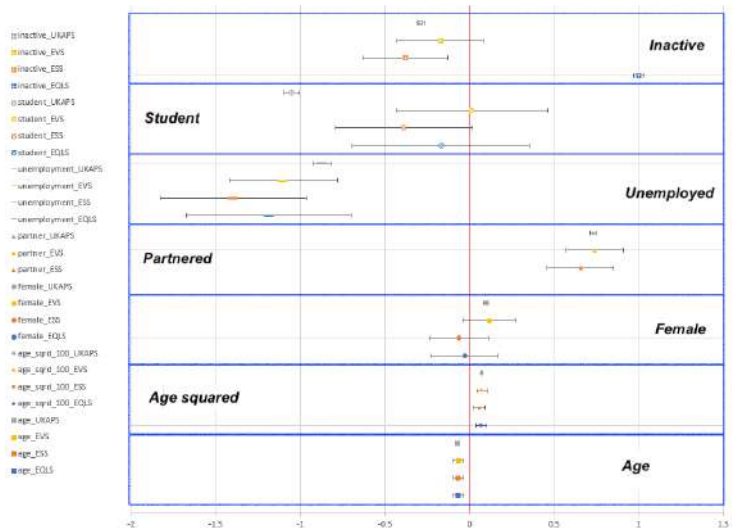
Table: Sample composition ( $\bar{X}$ )

Survey	Female	Age	Empl.	Unempl.	Student	Inactive	Partnered
EQLS	0.51	48.23	0.59	0.04	0.05	0.32	NA
ESS	0.51	47.08	0.56	0.04	<b>0.08</b>	0.32	NA
WVS	0.52	45.91	0.58	<b>0.07</b>	0.03	0.31	0.58
UKAPS	0.51	47.65	0.60	0.03	0.01	0.36	0.51

Means are estimated using probability sampling survey weights.

Regression coefficients ( $\hat{\beta}$ )

## Regression coefficients ( $\hat{\beta}$ )

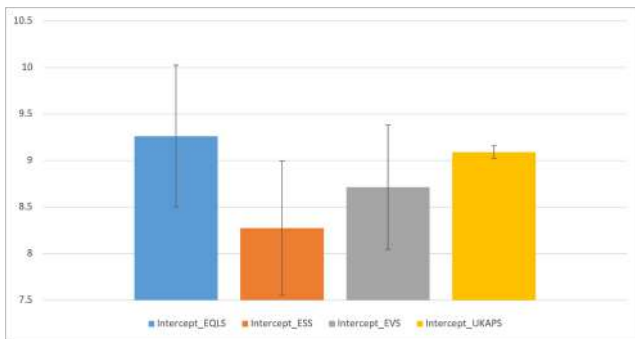




Survey fixed effects ( $\hat{\alpha}$ )

# Survey fixed effects

Survey fixed effects ( $\hat{\alpha}$ )



# Why survey fixed effects?

- ▶ We compare the survey-specific constants with the usual suspects from the literature on survey methodology, i.e.:
  - ▶ Seasonal effects
  - ▶ Mode of interview
  - ▶ Wording
  - ▶ Response rates
  - ▶ Survey response rate
  - ▶ Placement in the questionnaire
  - ▶ Quota sampling vs random sampling
  - ▶ Noise in the tails of the LS distribution

# Why survey fixed effects?

- ▶ We compare the survey-specific constants with the usual suspects from the literature on survey methodology, i.e.:
  - ▶ Seasonal effects
  - ▶ Mode of interview
  - ▶ Wording
  - ▶ Response rates
  - ▶ Survey response rate
  - ▶ Placement in the questionnaire
  - ▶ Quota sampling vs random sampling
  - ▶ Noise in the tails of the LS distribution
- ▶ None of these factors comes up as a robust explanation.

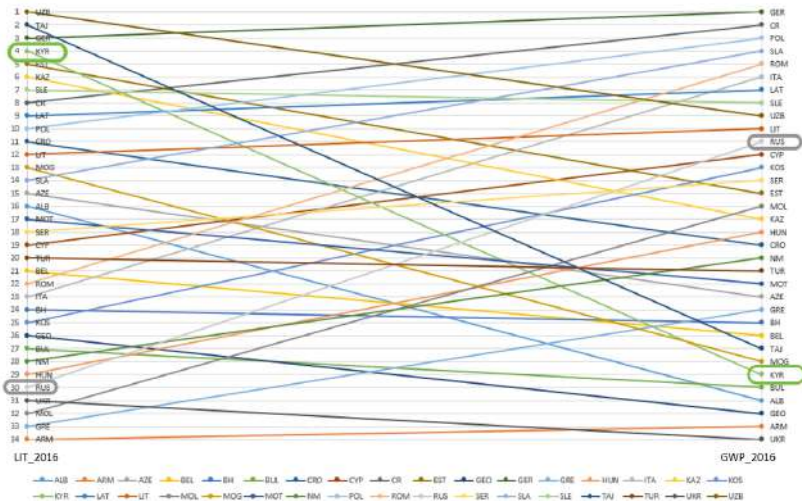
# Why survey fixed effects?

- ▶ We compare the survey-specific constants with the usual suspects from the literature on survey methodology, i.e.:
  - ▶ Seasonal effects
  - ▶ Mode of interview
  - ▶ Wording
  - ▶ Response rates
  - ▶ Survey response rate
  - ▶ Placement in the questionnaire
  - ▶ Quota sampling vs random sampling
  - ▶ Noise in the tails of the LS distribution
- ▶ None of these factors comes up as a robust explanation.
- ▶ Besides, fixed effects should not affect country rankings. But if we look at Kyrgyzstan...

# Who was happier in 2016, the Russian or the Kyrgyz?

# Who was happier in 2016, the Russian or the Kyrgyz?

Measure: Cantril 0-10 (GWP); LS 1-5 (LTS). Country rankings.



- 1 The puzzle
- 2 A measurement error approach
- 3 Conclusion



# Conclusions

- ▶ We document significant differences among estimates of **mean LS** from different surveys.

# Conclusions

- ▶ We document significant differences among estimates of **mean LS** from different surveys.
- ▶ The discrepancy seems to be due to:
  - ▶ Under- and over-**sampling** of some groups
  - ▶ **Survey fixed effects**.
- ▶ Importantly, LS determinants are **reliable** across surveys.
- ▶ Beyond sample composition effects, it is **unclear** why each specific survey leaves its “fingerprint” on LS.

Thank you for your attention

# Additional material

Table: Sample composition

Survey	Sex	Age	<20k	30k	40k	50k	75k	100k	150k	>150k
GSS	0.51	47.9	<b>0.10</b>	0.13	0.13	0.11	0.16	0.1	0.15	0.13
GWP	0.51	47.6	<b>0.19</b>	0.10	0.10	0.09	0.17	0.11	0.10	0.14
PEW	0.51	47.1	<b>0.20</b>	0.11	0.08	0.08	0.14	0.12	0.14	0.13

Means are estimated using probability sampling survey weights.

# France (2016-18)

Table: Sample composition

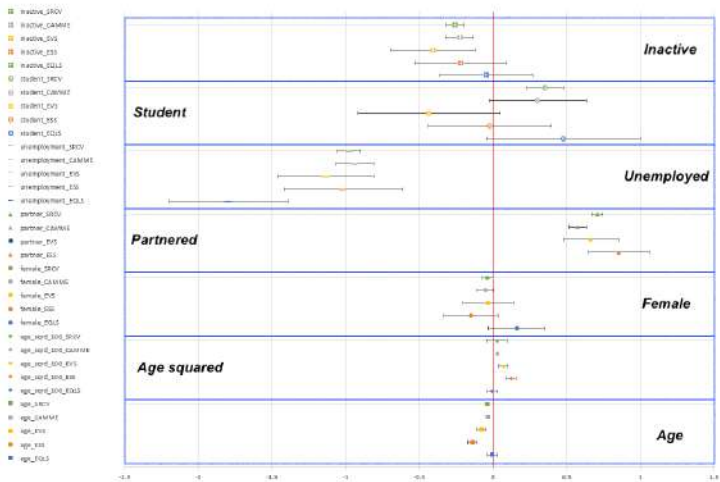
	female	age	empl.	unempl.	student	inactive	partnered
EQLS	0.52	49.2	0.53	0.06	0.05	0.35	NA
ESS	0.52	48.5	0.50	0.06	<b>0.10</b>	0.34	NA
EVS	0.52	50.1	0.50	0.08	0.05	0.37	<b>0.38</b>
CAMME	0.53	53.1	0.60	0.05	<b>0.01</b>	0.34	0.68
SRCV	<b>0.57</b>	52.7	0.49	0.06	0.03	0.41	0.64

Means are estimated using probability sampling survey weights.

# USA (2019)

TBD.png

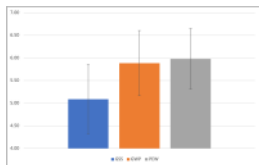
# France (2016-18)



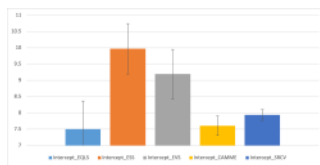


# Survey fixed effects

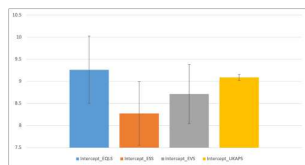
USA



France



UK



## Sample composition adjustment (in progress)

- ▶ LS country rankings are a widely cited statistics (cf. World Happiness Report)

## Sample composition adjustment (in progress)

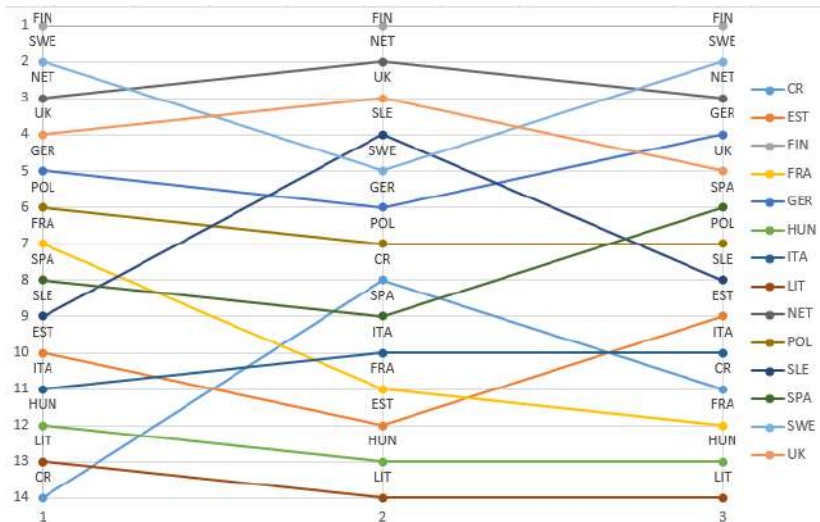
- ▶ LS country rankings are a widely cited statistics (cf. World Happiness Report)
- ▶ Country rankings are unaffected by survey fixed effects, but suffer from sample composition.

## Sample composition adjustment (in progress)

- ▶ LS country rankings are a widely cited statistics (cf. World Happiness Report)
- ▶ Country rankings are unaffected by survey fixed effects, but suffer from sample composition.
- ▶ How much country rankings vary from a survey to another?

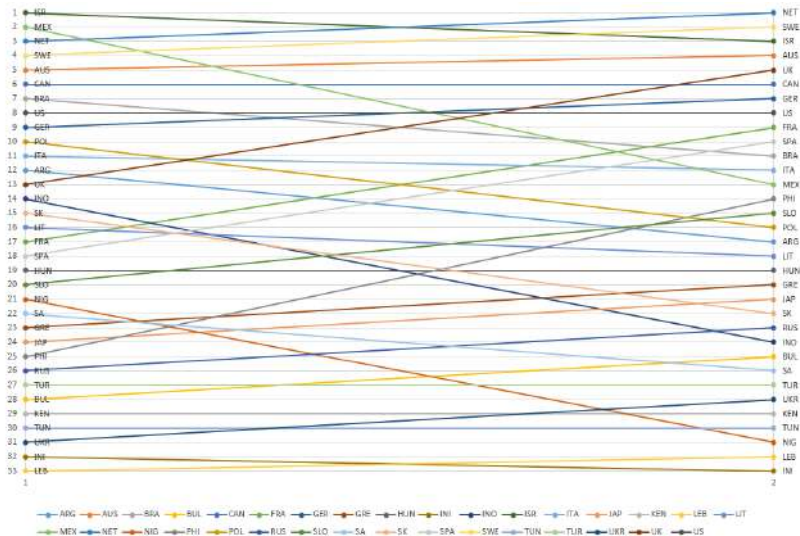
# Europe (2016-18)

Measure: LS (1-10 or 0-10). Country ranking.



# World (2019)

Measure: LS (1-10 or 0-10). Country rankings.



# Sample composition adjustment (in progress)

- ▶ We propose a simple bootstrapping method to adjust sample composition based on national official statistics.

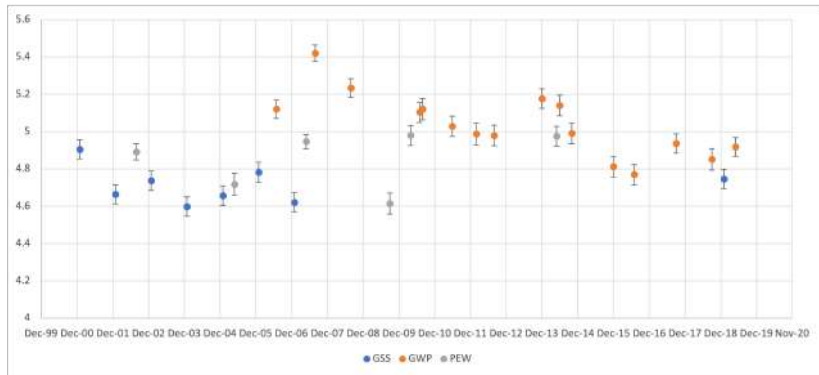
# Sample composition adjustment (in progress)

- ▶ We propose a simple bootstrapping method to adjust sample composition based on national official statistics.
- ▶ We compare country rankings from different surveys with and without the adjustment.



# Winsorization - USA

Measure: General Life Satisfaction (1-7). Mean LS.



back

# World (2019)

Measure: Cantril 0-10. Estimated means.

Period: May-Oct 2019 (PEW); Mar-Dec 2019 (GWP).

