



Building an aggregate measure: the case of the European Regional Progress Index: 2020 EU-SPI

Paola Annoni

European Commission - DG for Regional and Urban Policy (DG REGIO)

Policy Development & Economic Analysis Unit

paola.annoni@ec.europa.eu



[#EUSPI](https://twitter.com/EUSPI)

What is the EU-SPI?

An aggregate metric following the same framework of the [Global Social Progress Index](#) by the Social Progress Imperative

Includes exclusively social and environmental indicators (no economic measure)

Measures social outcomes and not inputs

Includes comparable and harmonised indicators for all the European regions

Covers matters that can be directly addressed by policy intervention

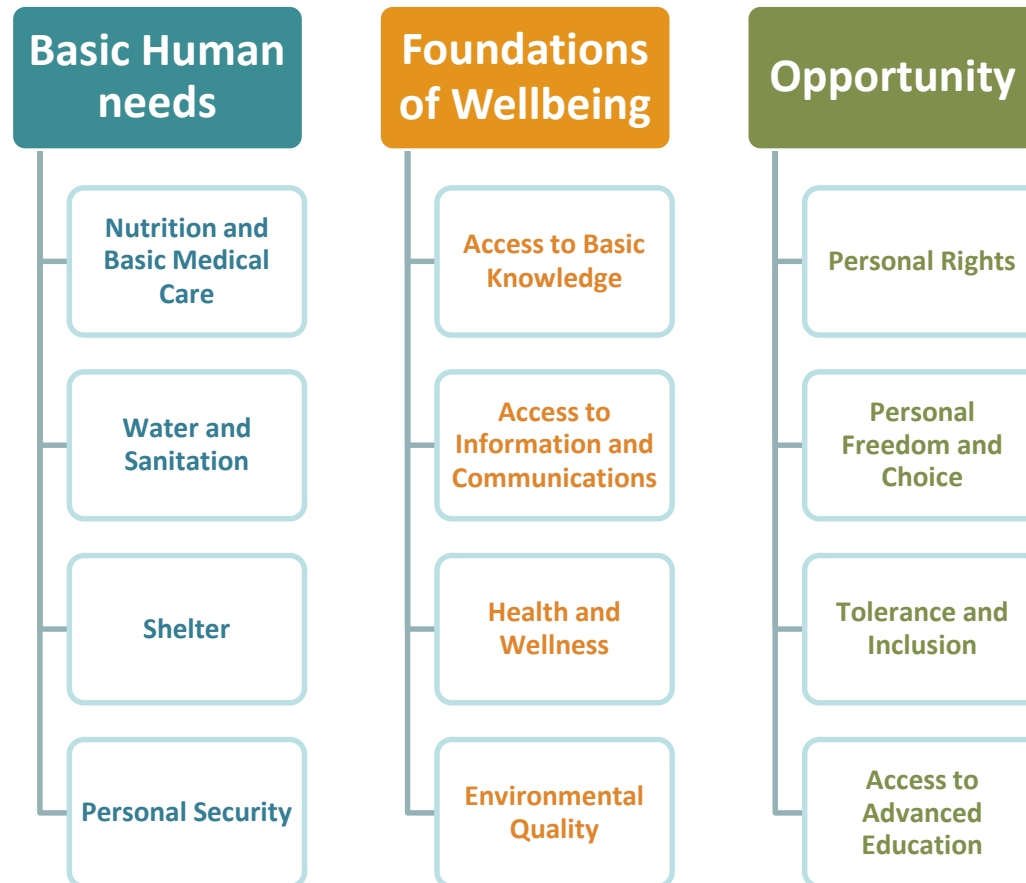
1st edition published in 2016, 2nd on 2 December 2020

FOCUS ON THE *BEYOND GDP* DISCUSSION

(measurable) factors that enable citizens to realise their potential

55 indicators in total elected from a set of around 70 candidates

The European Union Regional Social Progress Index



Basic Human Needs

1. Nutrition and Basic Medical Care

- Mortality rate before 65 (a)
- Infant mortality (a)
- Unmet medical needs (b)
- Insufficient food (b)

2. Water and Sanitation

- Satisfaction with water quality (c)
- Lack of toilet in dwelling (b)
- Uncollected Sewage (d)
- Sewage treatment (d)

3. Shelter

- Burden cost of housing (b)
- Housing quality due to dampness **NEW** (b)
- ~~Housing quality due to darkness **NEW** (b)~~
- Overcrowding (b)
- Adequate heating (b)

4. Personal Security

- ~~Traffic deaths (a)~~
- Crime (b)
- Safety at night (c)
- Money stolen **NEW** (c)
- Assaulted/Mugged **NEW** (c)
- ~~Traffic deaths and injuries **NEW** (a)~~

Foundations of Wellbeing

5. Access to Basic Knowledge

- ~~Pre-primary enrolment rate **NEW** (b)~~
- Upper secondary enrolment rate age 14-18 (a)
- Lower secondary completion rate (a)
- Early school leavers (a)

6. Access to Information and Communications

- Internet at home (a)
- Broadband at home (a)
- Online interaction with public authorities (a)
- Internet access **NEW** (c)

7. Health and Wellness

- Life expectancy (a)
- Self-perceived health status (b)
- Cancer death rate (a)
- Heart disease death rate (a)
- ~~Alcohol consumption death rate **NEW** (a)~~
- Leisure activities **NEW** (b)
- ~~Unmet dental needs (b)~~
- ~~Satisfaction with air quality (c)~~
- Traffic deaths (a)

8. Environmental quality

- ~~Carbon intensity **NEW** (d)~~
- Air pollution NO2 (d)
- Air pollution ozone (d)
- Air pollution pm10 (d)
- Air pollution pm2.5 (d)
- ~~Pollution, grime or other env. problems (b)~~
- ~~Noise (b)~~
- ~~Natura2000 (d)~~

Opportunity

9. Personal Rights

- Trust in the national government (c)
- Trust in the legal system (c)
- Trust in the police (c)
- Active citizenship **NEW** (e)
- Female participation in regional assemblies **NEW** (i)
- Quality of public services (f)

10. Personal Freedom and Choice

- Freedom over life choices (c)
- Job opportunities **NEW** (c)
- Involuntary part-time/temporary employment **NEW** (a)
- ~~Teenage pregnancy (a)~~
- Young people not in education, employment or training NEET (a)
- ~~NEET gender gap **NEW** (j)~~
- ~~Gender pay gap **NEW** (a)~~
- Corruption in public services (f)

11. Tolerance and Inclusion

- Impartiality of public services (f)
- Tolerance towards immigrants (c)
- Tolerance towards minorities (c)
- Tolerance towards homosexuals (c)
- Making friends **NEW** (c)
- ~~Trust in others (e)~~
- ~~Safety net (e)~~
- Volunteering **NEW** (e)
- ~~Attitudes toward people with disabilities (g)~~
- Gender employment gap @ (a)

12. Access to Advanced Education and LLL

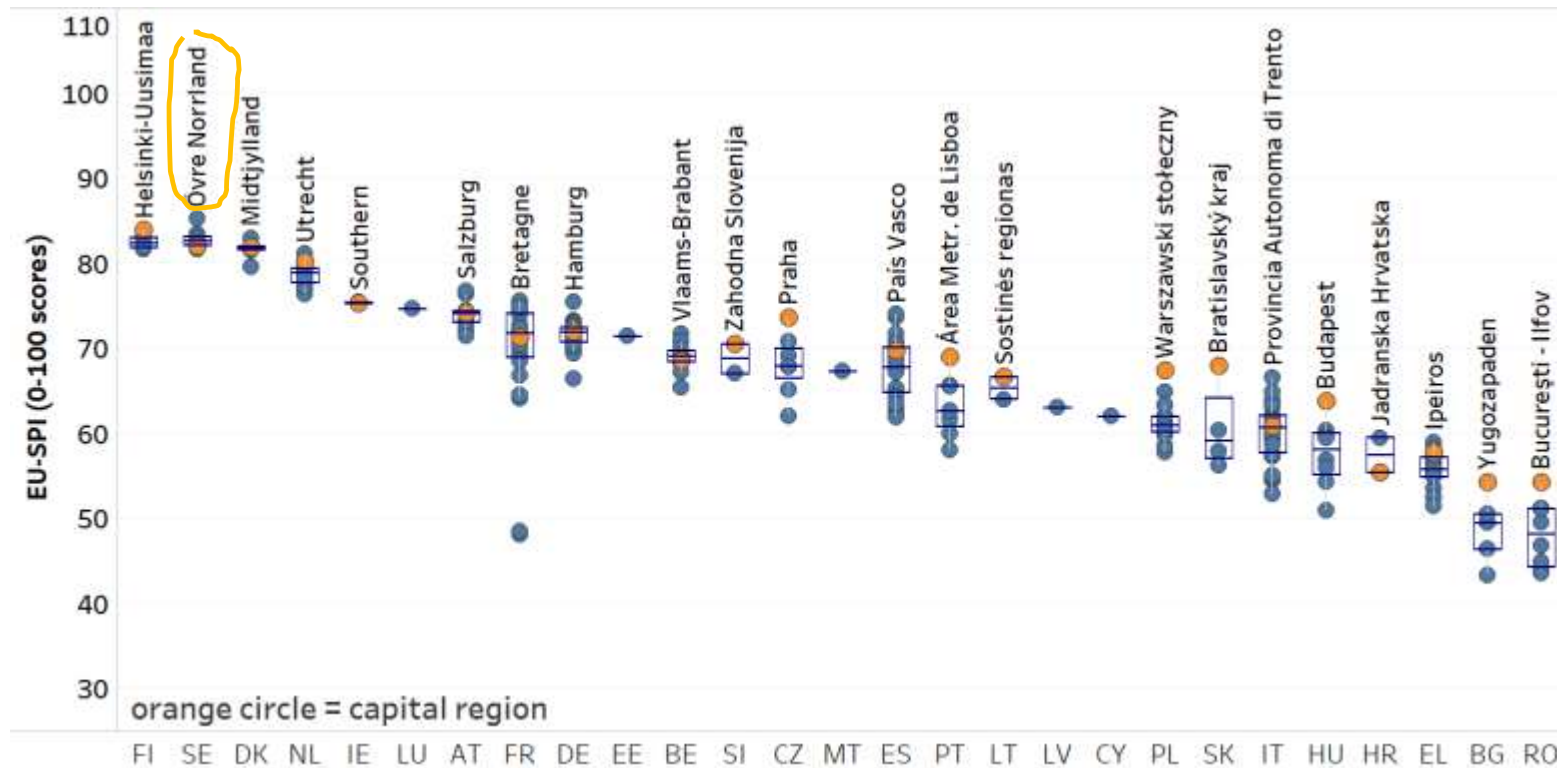
- Tertiary education attainment @ (a)
- Tertiary enrolment (a)
- Lifelong learning (a)
- ~~Accessibility to University **NEW** (h)~~
- Female life-long education and learning **NEW** (a)

SOURCES:

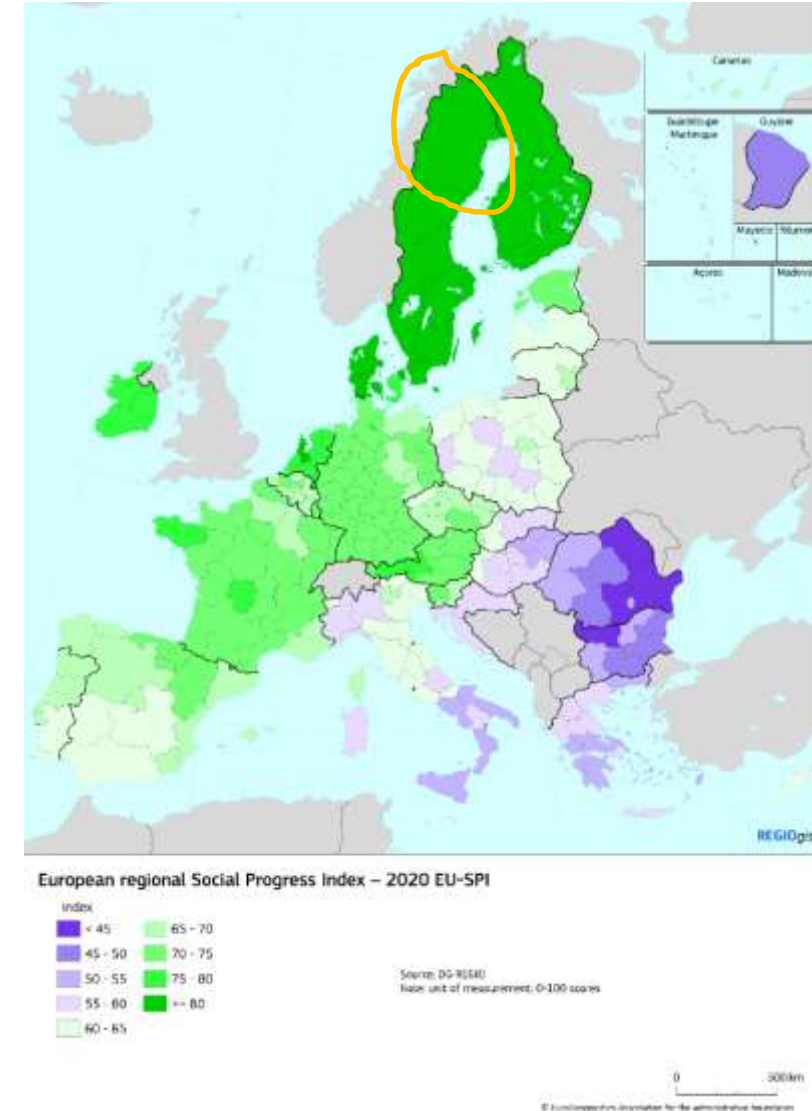
(a) EUROSTAT (b) EU-SILC (c) GALLUP WORLD POLL
 (d) EU Environmental Agency (e) EU-SILC ad-hoc module (Social/cultural participation/material deprivation, 2015)
 (f) Quality of Government Index (EQI 2017) (g) Special Eurobarometer 437, 2015
 (h) REGIO own computations (i) Gender Statistics Database (EU Institute for Gender Equality)

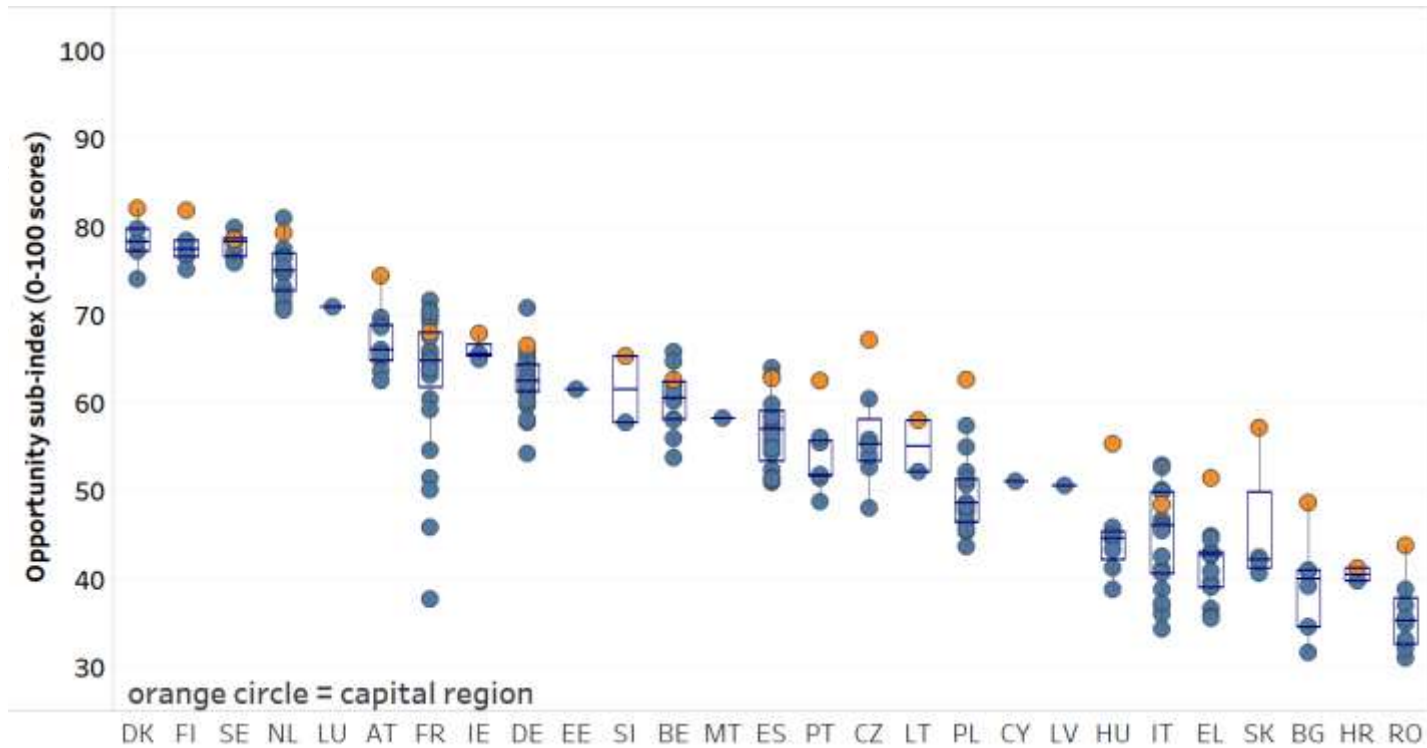
HIGH VARIABILITY

The richest regions not always the top performers



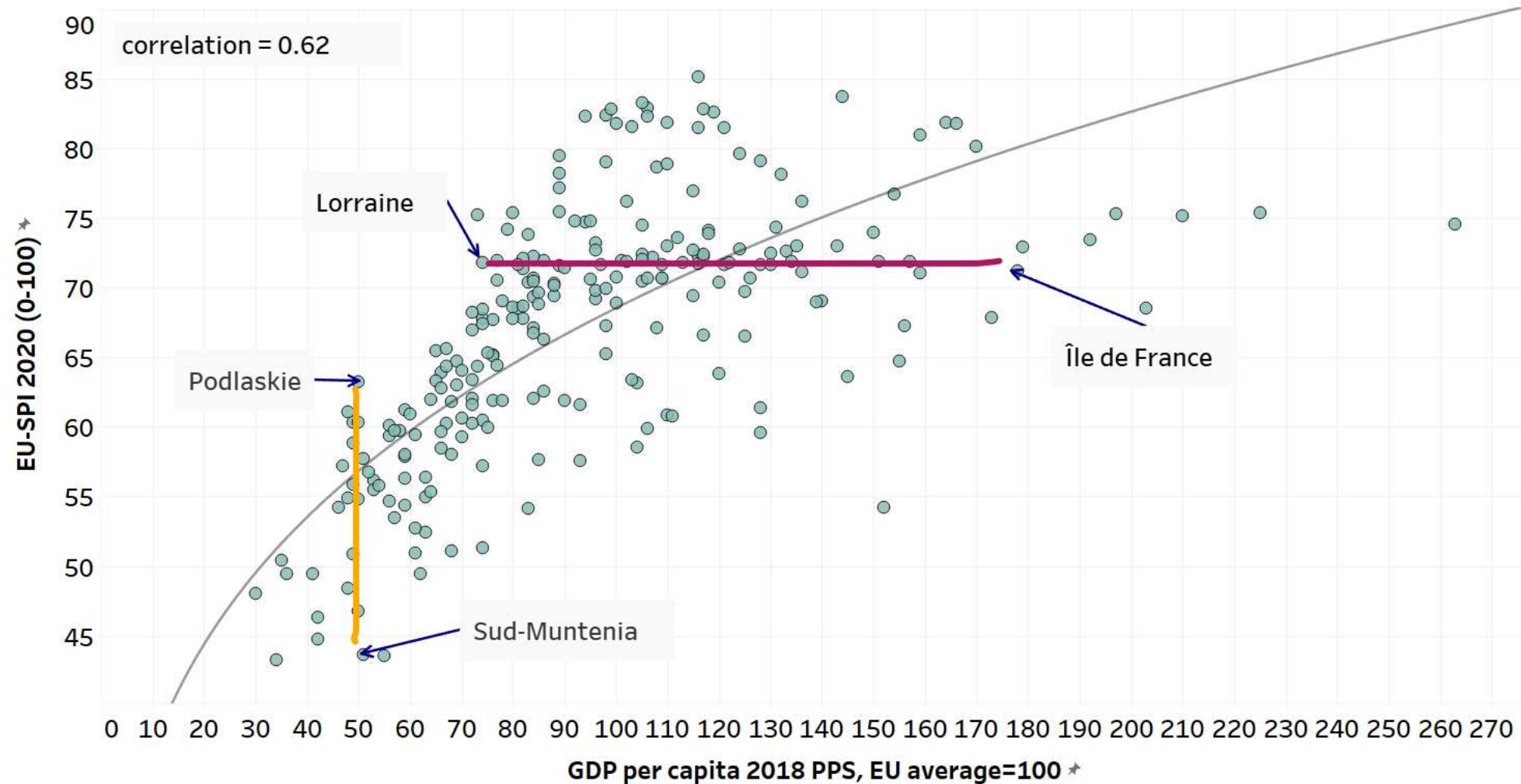
Countries reordered from best to worst according to their national EU-SPI score





GOOD AT BASICS, LESS GOOD AT TRUST AND OPPORTUNITIES

Different GDP, similar social level and vice versa



Flipping the point of view: can EU-SPI explain GDP?

Regression Statistics

Multiple R	0.64
R Square	0.41
Adjusted R Square	0.40
Standard Error	0.78
Observations	240

on z-scores

ANOVA

	df	SS	MS	F	Significance F
Regression	3	96.82	32.27	53.57	0.00
Residual	236	142.18	0.60		
Total	239	239			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	0.00	0.05	0.00	1.00	-0.10	0.10
Basic sub-index	0.09	0.08	1.05	0.29	-0.08	0.25
Foundation sub-index	-0.07	0.11	-0.61	0.55	-0.28	0.15
Opportunity sub-index	0.62	0.11	5.52	0.00	0.40	0.85

Regression Statistics

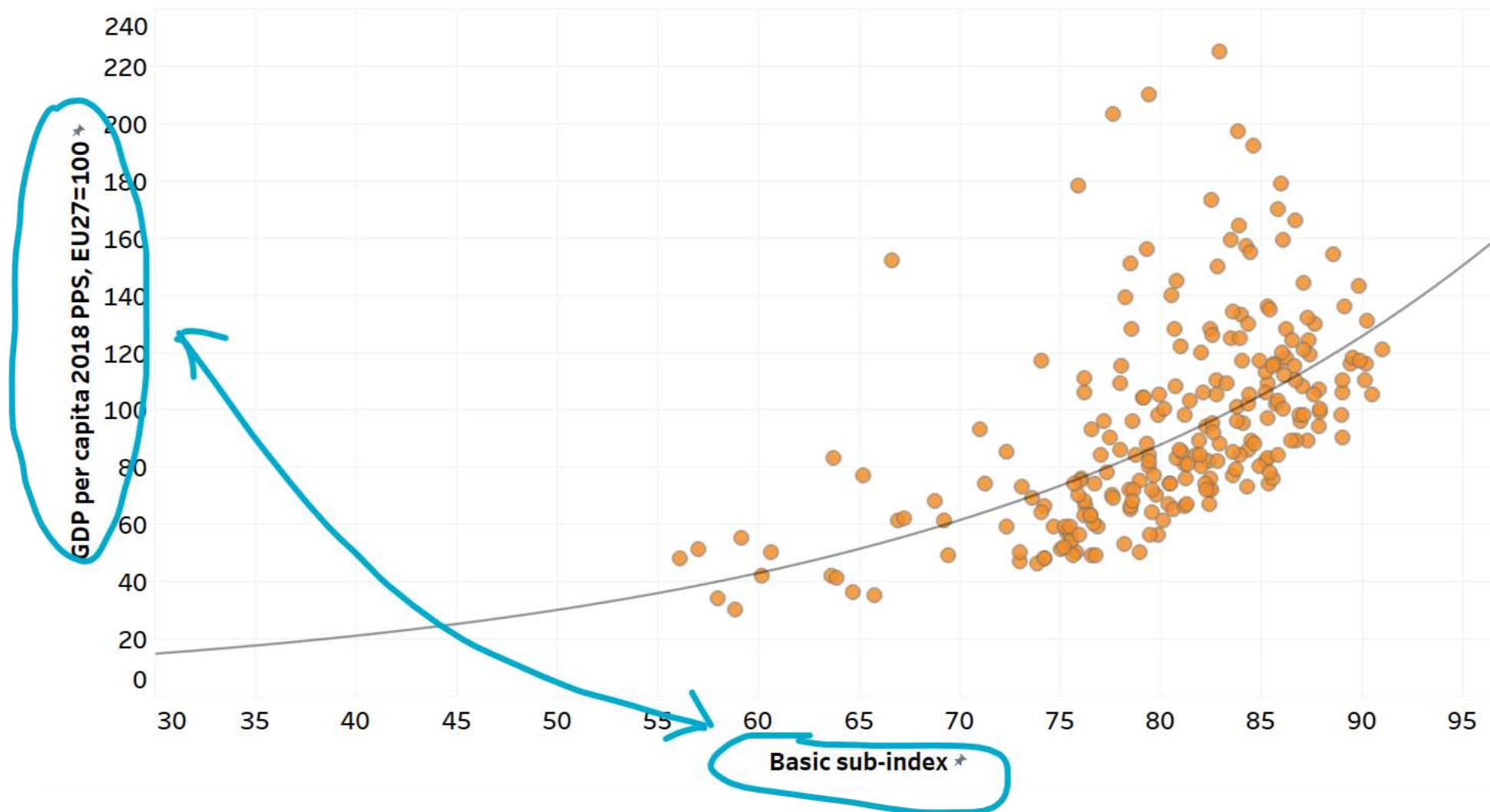
Multiple R	0.69
R Square	0.48
Adjusted R Square	0.47
Standard Error	0.73
Observations	233

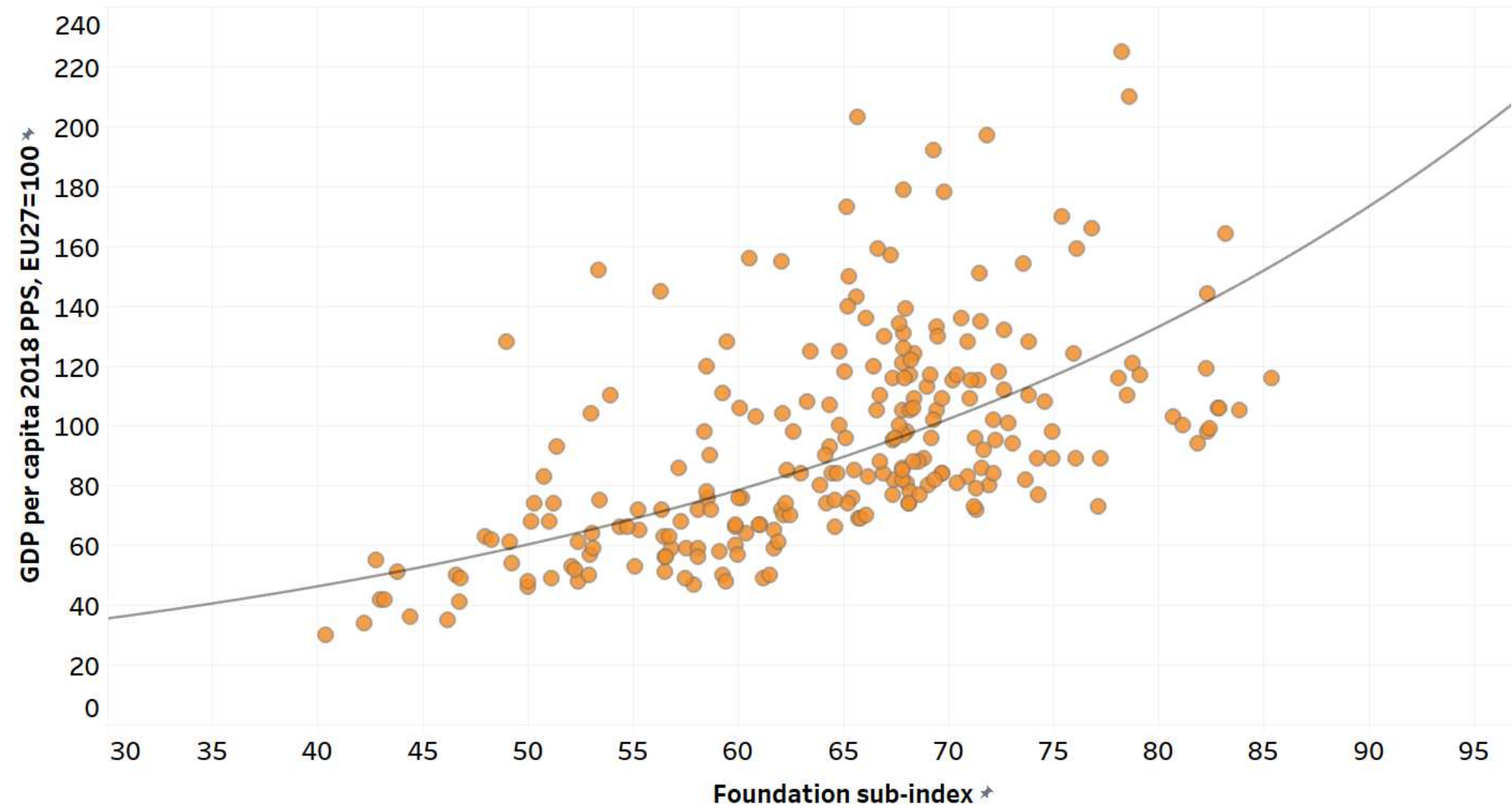
on z-scores,
log(GDP)

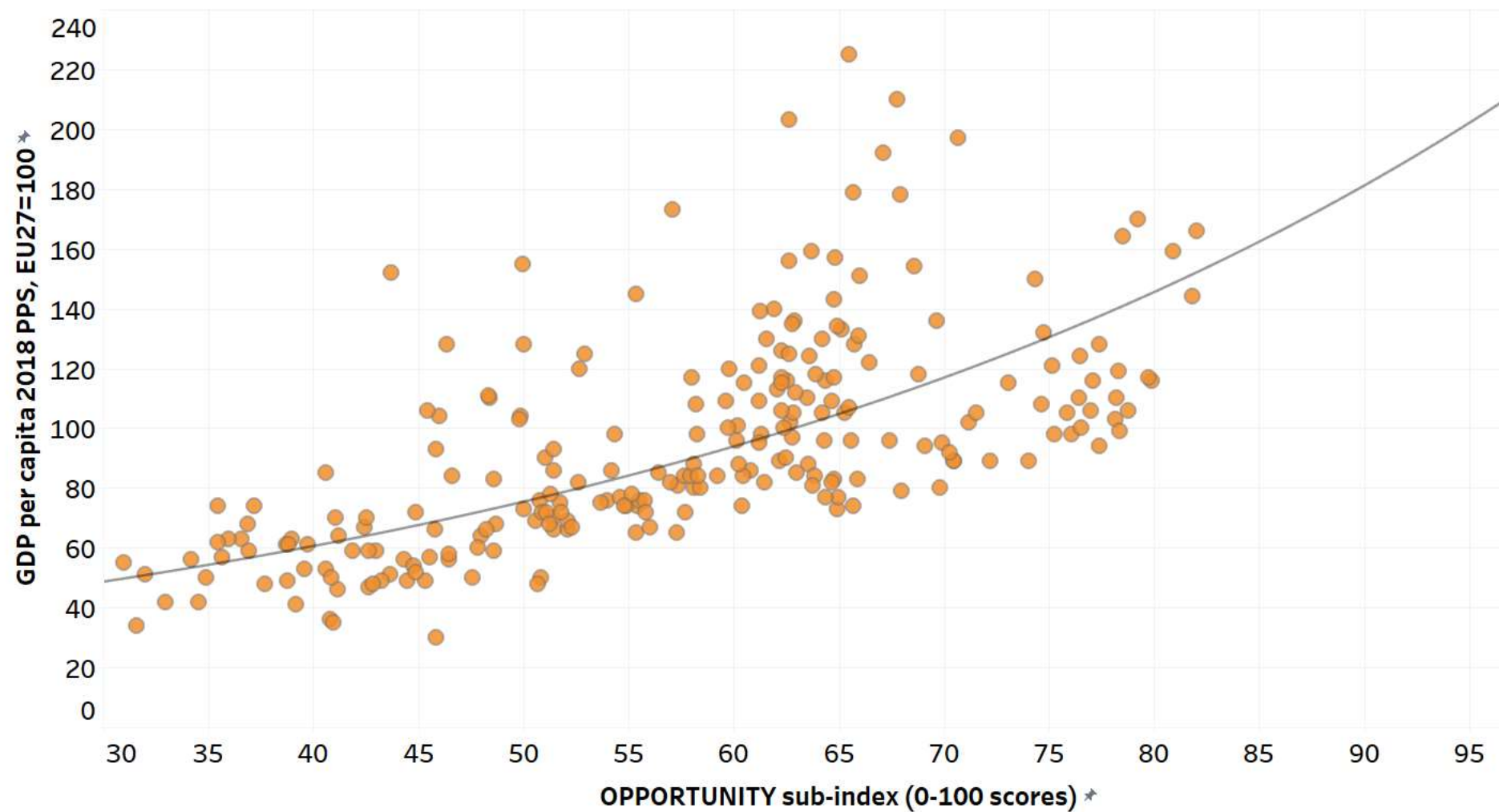
ANOVA

	df	SS	MS	F	Significance F
Regression	3	110.38	36.79	69.27	0.00
Residual	229	121.62	0.53		
Total	232	232			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	0.00	0.05	0.00	1.00	-0.09	0.09
Basic sub-index	0.15	0.08	1.98	0.05	0.00	0.30
Foundation sub-index	-0.03	0.10	-0.34	0.73	-0.23	0.16
Opportunity sub-index	0.60	0.11	5.58	0.00	0.39	0.81







"Trust, the one thing that changes everything" (S.M.R. Covey)



Stephen
M.R.
Covey
(writer,
author of
[*The Speed of Trust*](#))

The people's perspective:

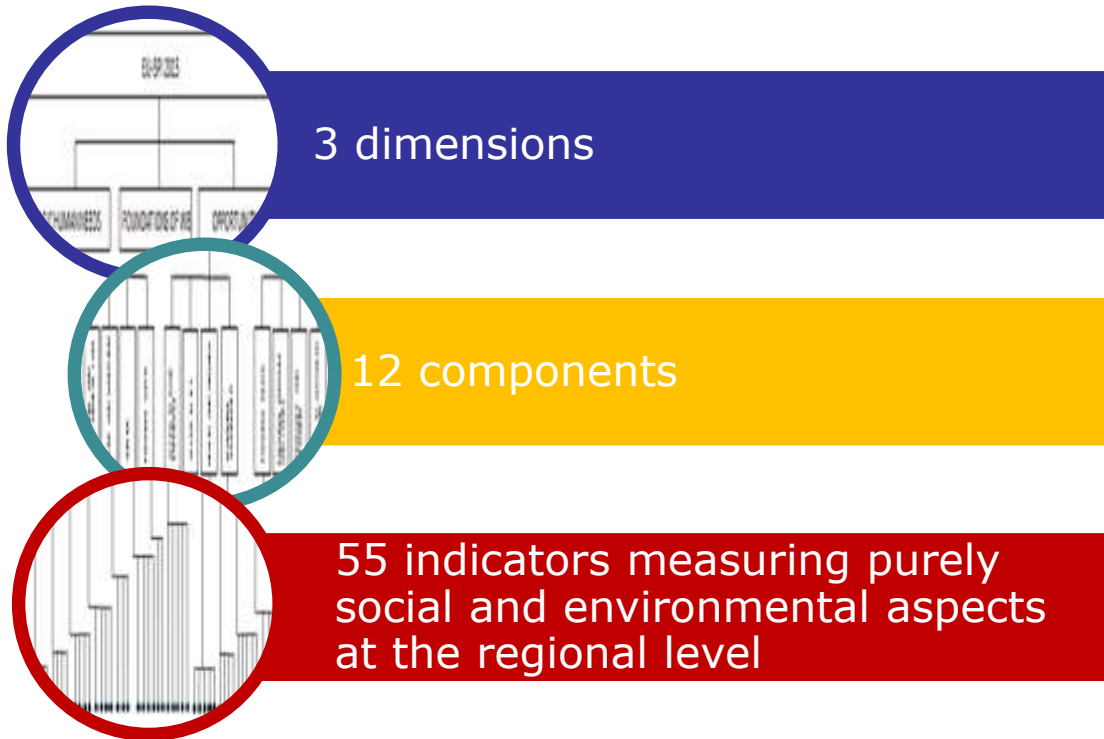
Low trust "affects the way we see our problems and how we try to solve them"

The business perspective:

... "marketplace success is trust monetized. When you outperform your competitors, it's clear that the market has greater trust " in you..."

... "consider ride-sharing; person-to-person home sharing; crowd funding; open source technology" ... covid-related recommendations [my note]

Back to “*hard science*”



Main statistical issues to be solved (once basic candidate indicators are collected):

- a. statistical assessment + internal consistency checks
- b. utopian and dystopian values for normalisation
- c. inequality-adverse type of aggregation

Personal freedom and choice component

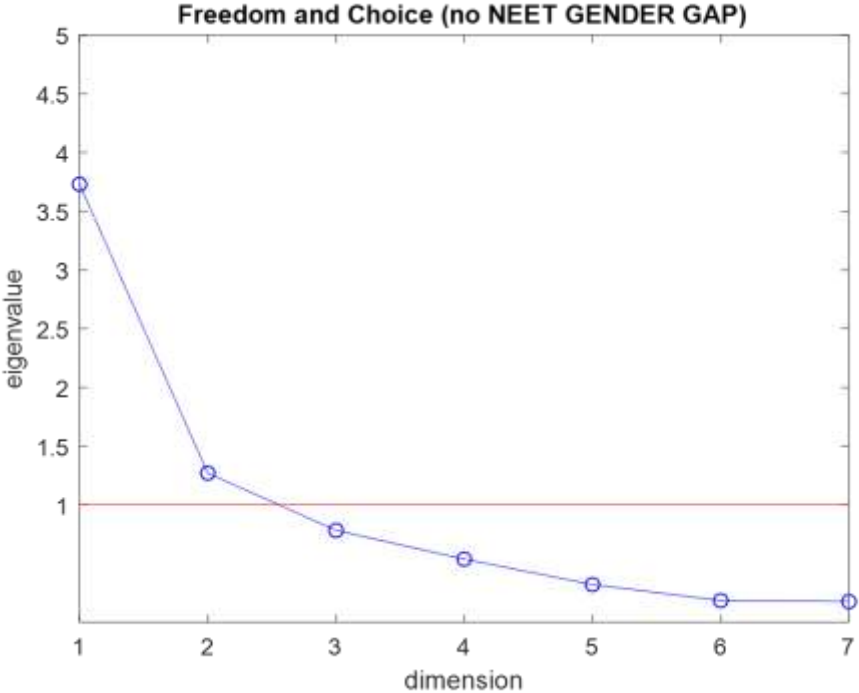
8 candidate indicators

5 eventually selected

Indicator name	Freedom over life choices	Job opportunities	Involuntary part-time/ temporary job	Teenage pregnancy
description of indicator	Share of respondents answering satisfied to the question, "Are you satisfied or dissatisfied with your freedom to choose what you do with your life?"	Share of respondents who think it is a good time to find a job in the city or area where they live	Share of population aged 20-64 in involuntary part-time or temporary job	Ratio between births from mothers 15-19 and the female population of the same age cohort
source	Gallup World Poll (WP134)	Gallup World Poll (WP89)	EUROSTAT (ad-hoc extraction)	EUROSTAT (demo_r_fagec3)
reference year	2020	2020	average 2016-2018	average 2016-2018
orientation	↑	↑	↓	↓
% of missing values	2.50	2.50	2.50	0.00
missing regions	FI20, FRY1-FRY5	FI20, FRY1-FRY5	FI20, FRM0, FRY3, FRY5, RO32, RO42	-
average	0.82	0.47	12.50	1.05
standard deviation	0.09	0.18	8.07	1.13
coefficient of variation	0.11	0.37	0.65	1.07
skewness	-0.72	-0.10	0.74	3.44
maximum value	0.97	0.91	38.79	8.31
region/country corresponding to max	FI1C	CZ01	ES61	FRY5
minimum value	0.59	0.13	1.02	0.19
region/country corresponding to min	EL61	ITG1	SK01	DK01
min value for normalization				
max value for normalization				

Indicator name	Young people not in education, employment or training (NEET)	NEET gender gap	Gender pay gap	Institutions corruption Index
description of indicator	Young people, aged between 15 and 24, neither in employment nor in education and training	Young people, aged between 15 and 24, neither in employment nor in education and training - absolute difference between male and female rates	Gender pay gap in unadjusted form - NACE Rev. 2 activity (earn_gr_gpr2). It is computed as the absolute difference between average gross hourly earnings of female paid employees and of male paid employees as a percentage of average gross hourly earnings of male paid employees. The GPG is calculated on the basis of: - the four-yearly Structure of Earnings Survey (SES)	Perceived level of Corruption. The Index is measured in z-scores
source	EUROSTAT (edat_ifse_22)	EUROSTAT and DG Regio own computations (edat_ifse_22)	EUROSTAT (ad-hoc extraction from Structure of Earnings Survey)	European Quality of Government Index
reference year	average 2016-2018	average 2016-2018	2014	2017
orientation	↓	↓	↓	↑
% of missing values	1.67	15.00	0.00	1.25
missing regions	AT11, DEB2, FI20, FRY5	AT11, AT32-34, BE54, DE14, DE22-27, DE50, DE72-73, DEB2, DECO, DED2, DED4, DED5, ES63, FI20, FRM0, FRY5, ITC2, NL13, NL23, NL34, PL52, PL91, PT15, PT18, PT20, PT30, SK01	-	ES63-64, FRY5
average	10.93	2.48	0.14	0.04
standard deviation	5.68	2.03	0.06	1.02
coefficient of variation	0.52	0.82	0.44	25.26
skewness	1.22	1.74	-0.01	-0.21
maximum value	33.30	11.87	0.28	2.51
region/country corresponding to max	FRY3	RO22	EE	FI20
minimum value	2.50	0.10	0.00	-2.56
region/country corresponding to min	CZ01	NL32	BE31	BG34
min value for normalization				
max value for normalization				

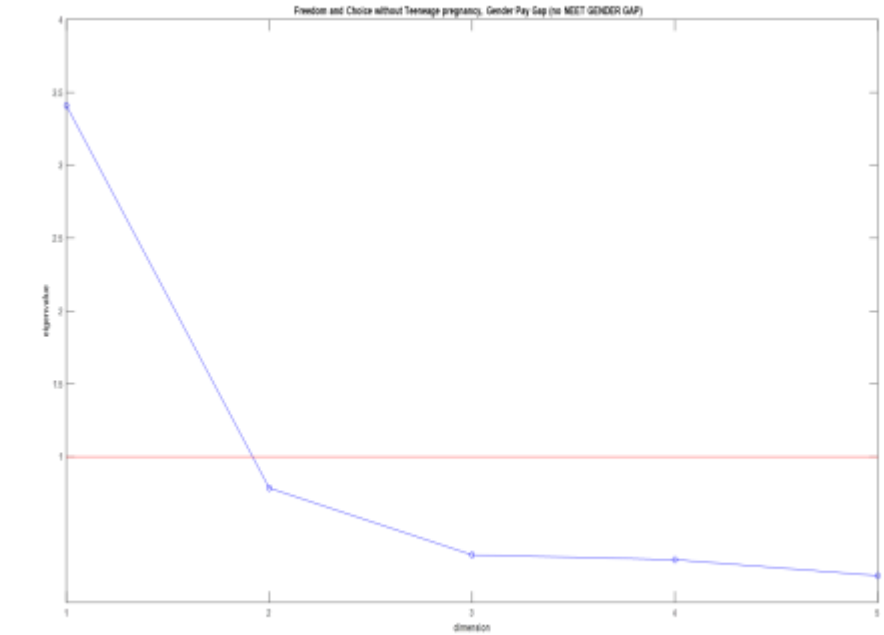
NOTE: Shaded indicators excluded from computation. NEET GENDER GAP excluded because too many missing values. TEENAGE PREGNANCY and GENDER PAY GAP excluded as non fitting (PCA test)



	Component			
	1	2	3	4
Freedom over life choices	-.430	-.015	-.200	.609
Job opportunities	-.466	.049	-.152	-.048
Involuntary part-time/ temporary job	.291	-.617	.030	.527
Teenage pregnancy	.187	.740	-.237	.429
NEET	.466	.050	.120	.299
Gender pay gap	-.274	.175	.914	.213
Institutions corruption Index	-.431	-.191	-.171	.174

ANNONI Paola (REGIO):
UNEXPECTED
SIGN/ORIENTATION

Total Variance Explained		
	Total	% of Variance
1	3.731	53.295
2	1.270	18.144
3	.782	11.173
4	.537	7.675
5	.320	4.576
6	.184	2.633
7	.175	2.503



	Component		
	1	2	3
Freedom over life choices	.458	.335	-.521
Job opportunities	.493	-.024	-.370
Involuntary part-time/ temporary job	-.325	.876	-.108
NEET	-.488	.026	-.092
Institutions corruption Index	.451	.347	.756

Total Variance Explained		
	Total	% of Variance
1	3.410	48.712
2	.785	11.220
3	.327	4.670
4	.294	4.195
5	.184	2.631

Min-max normalization on a 0-100 scale

Boundary choices are (or seem) **straightforward** for some indicators:

Percentage of people trusting the political system:

 0-100 

Uncollected sewage as % of total generated load:

 0-100 

Rather **tricky** for others:

Infant mortality:
(number of deaths aged < 1 year per 1000 births)

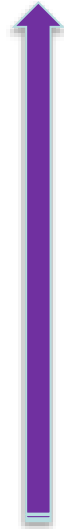
 0-16  Highest since 2008

Air pollution (pm10):
($\mu\text{g}/\text{m}^3$)

 0-40  EU guidelines yearly limit value

Utopian state 

Dystopian state 



Inequality-adverse aggregation function across components and dimensions

Generalized mean of order $\beta = 0.5$

$$I_j = \begin{cases} \left(\frac{1}{q} \sum_{i=1}^q x_i^\beta \right)^{1/\beta} & \beta \neq 0 \\ \left(\prod_{i=1}^q x_i \right)^{1/q} & \text{for } \beta = 0 \text{ (geometric mean)} \end{cases} \quad (\text{equal weighting})$$

q = number of components (or dimensions)

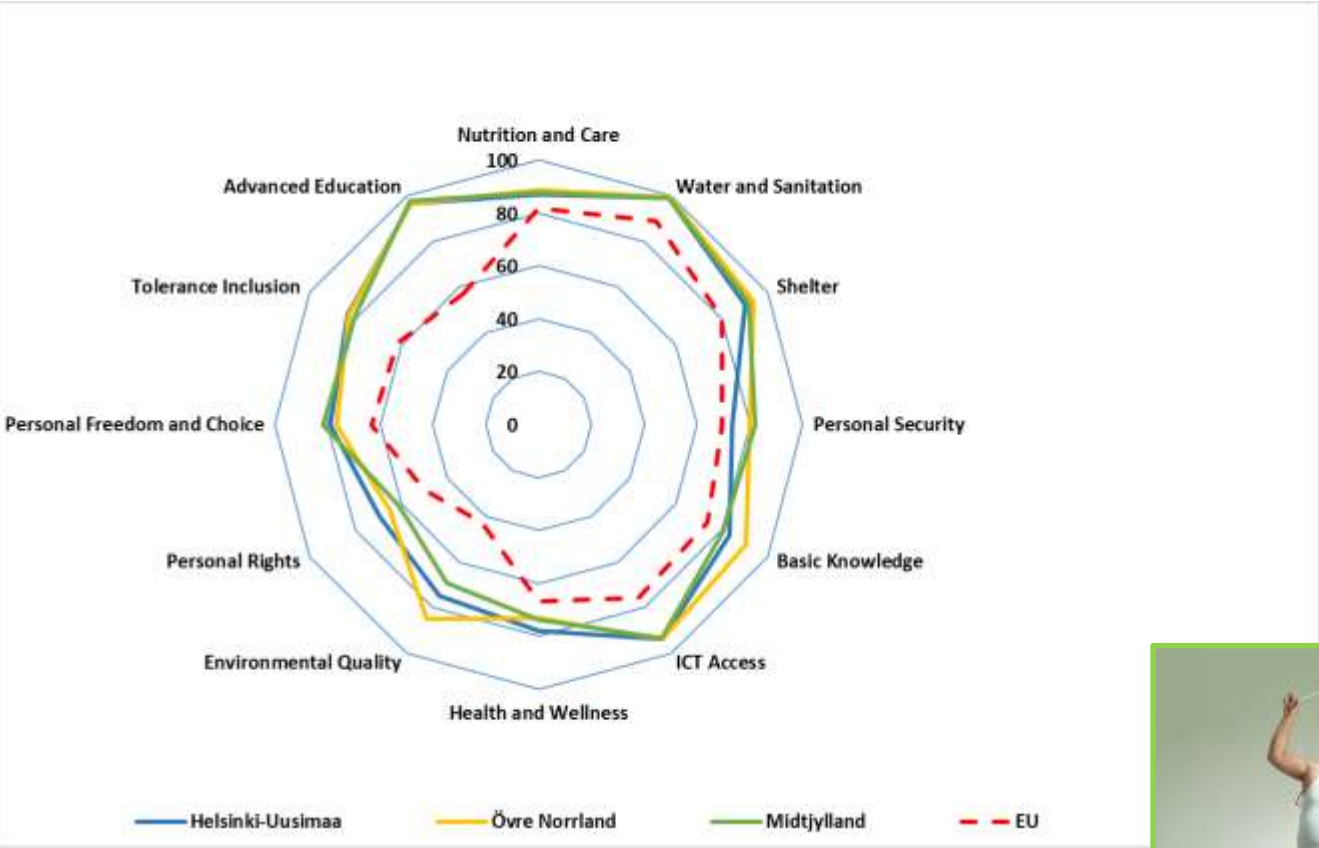
j = region

x_i = value of indicator i observed for region j (reordered to have the higher, the better)

Why:

- ✓ to smooth out compensability effects across components
- ✓ placed in between a geometric and an arithmetic mean
- ✓ penalises regions with low scores in certain components

WHY? BECAUSE THE ROUNDER, THE BETTER



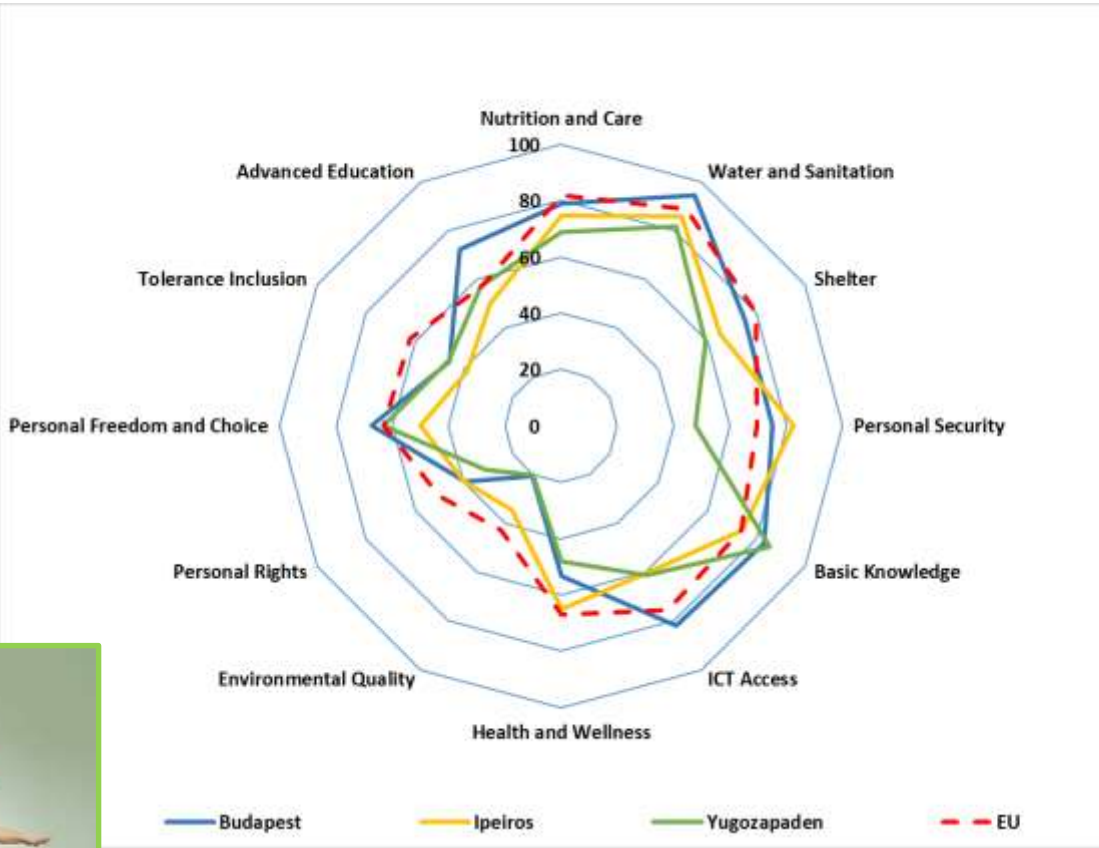
EU-SPI
(0-100)

84

85

83

67

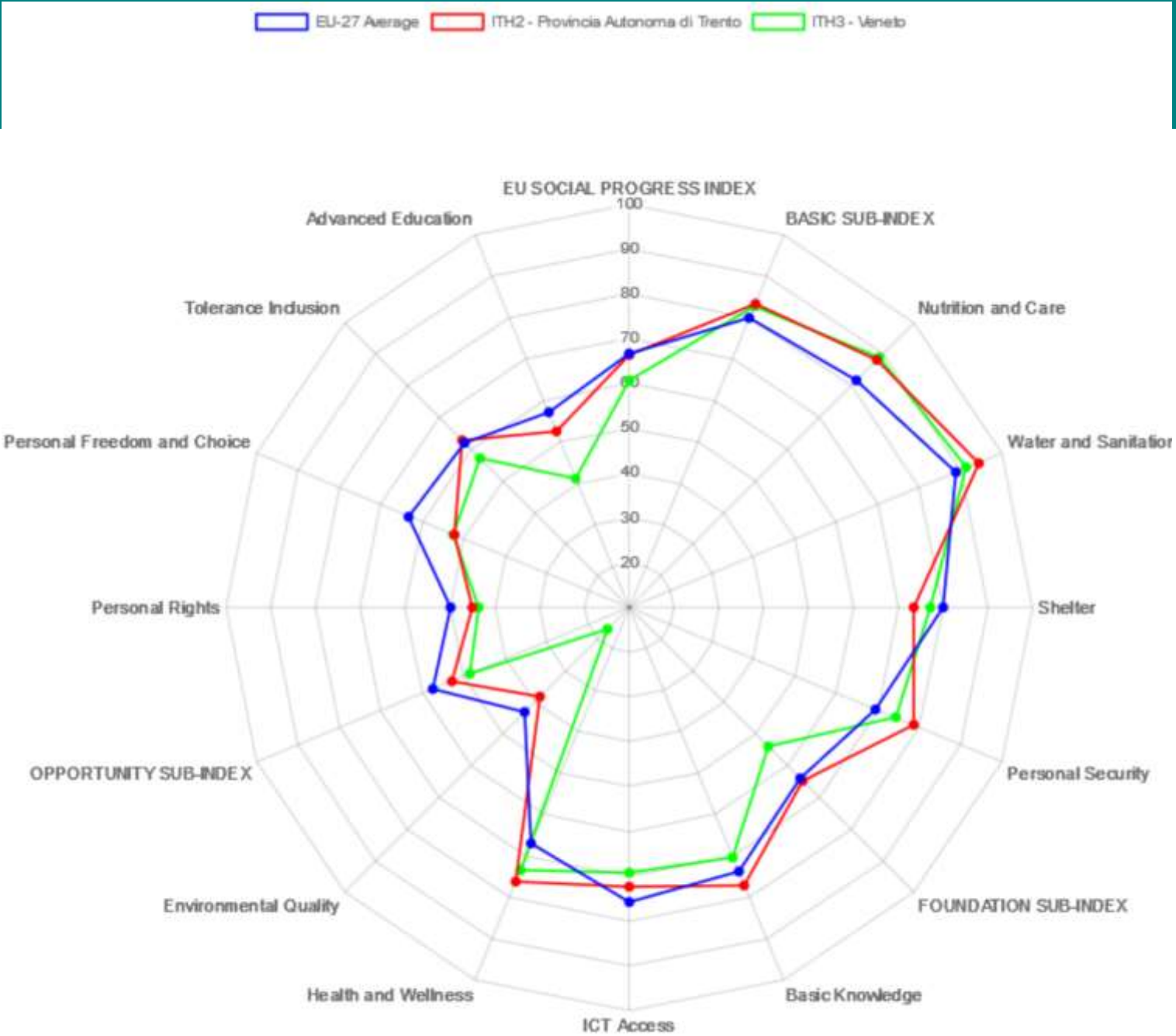


64

59

54

67



EXPLORE OUR DATA:

[2020 EU-SPI Europa page](#)

[Interactive web story on our Open Data Portal](#)

[Working paper](#)

thank
you!