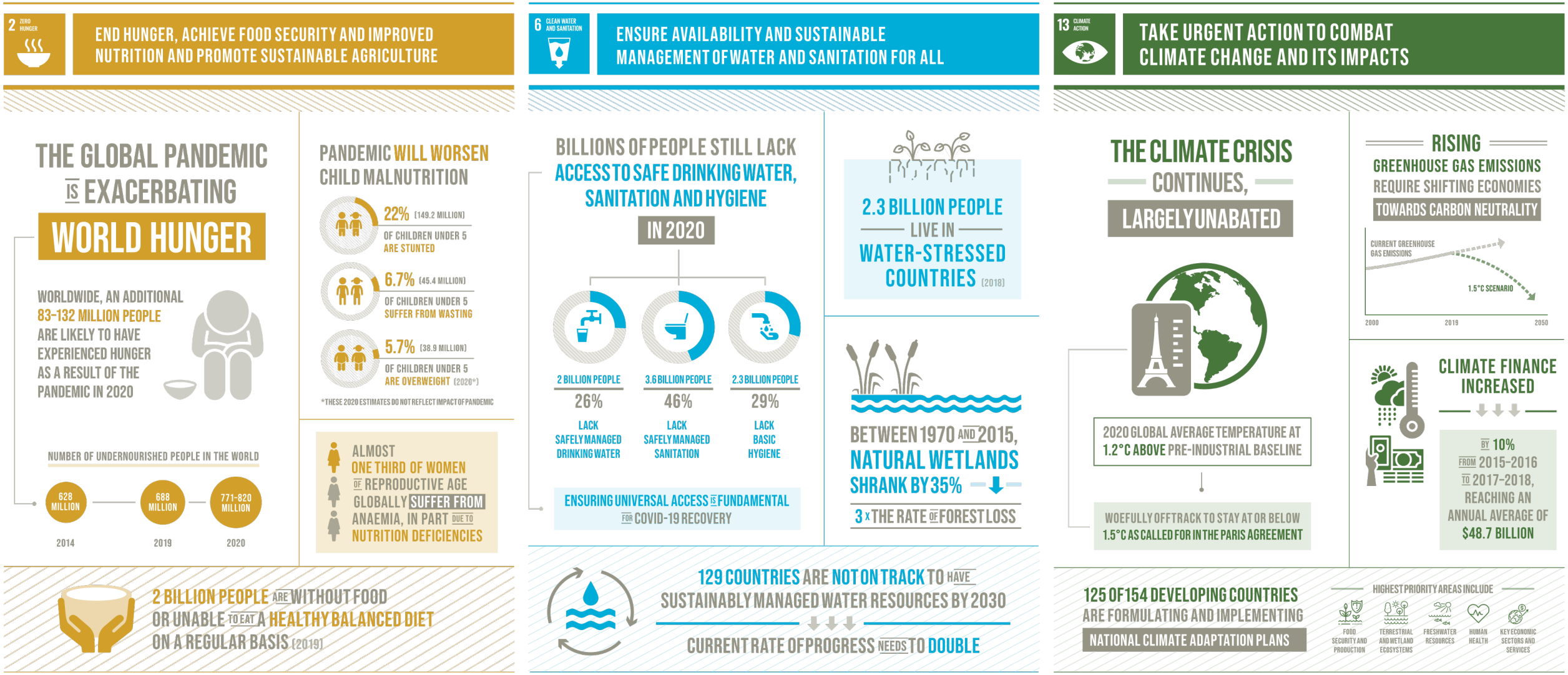
# Countries' research priorities in relation to the Sustainable Development Goals

Hugo Confraria[1,2], Ed Noyons[3], Tommaso Ciarli [1,4]
[1]SPRU, University of Sussex
[2]UECE, ISEG, University of Lisbon
[3]CWTS, Leiden University
[4]UNU-MERIT, Maastricht University

November 24

CONCORDi 2021- Industrial innovation for competitive sustainability

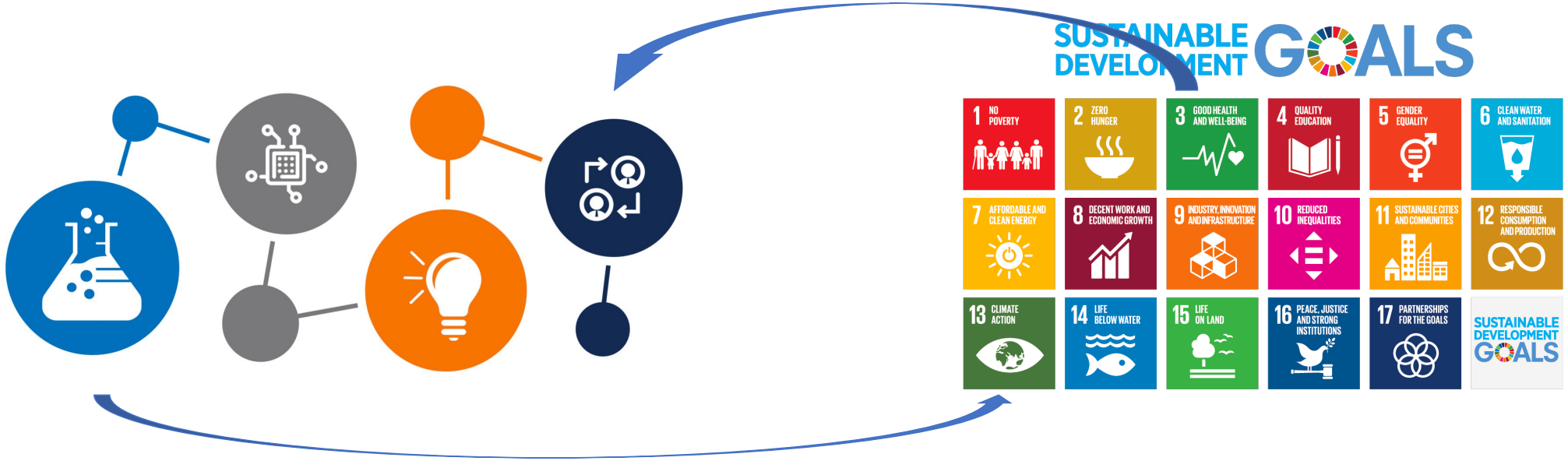
# ‘We are seriously off-track’ UN chief António Guterres




https://unstats.un.org/sdgs/report/2021/Overview/

# In this paper

- We use the SDGs to analyse the extent to which **countries' research priorities align with their greatest SDG challenges**.
- Based on the notion of **building capabilities to address challenges** (Pavitt, 1998; Salter and Martin, 2001), here we make the very general assumption that a misalignment between a country's research priorities and its SDG challenges may reduce the effectiveness of investments in research to address those goals.



# Data and Methods

- **“Revealed” research priorities** of countries related to a certain SDG
  - Web of Science (WoS) publication data
  - 4-step approach: Text mining -> Topic modelling -> Queries -> CWTS clusters -> Loose threshold
  - Specialisation Index: Score between -1 (No research) --- 0 (World average) --- 1 (All research)

- **Societal challenges** of countries in a certain SDG
  - Combination of all available/relevant indicators from: UN SDG database + SDG Index
  - 4-step approach: Data collection -> Normalisation -> Distance to the frontier -> PCA per SDG
  - Societal Challenges: Score between -1 (country at the frontier) --- 1 (major challenge)

# STRINGS Approach: Four Steps

1. Search reports (policy documents, grey literature), scientific publications and other web content that describes specific SDGs (e.g. descriptions, progress and major problems).

2. Select relevant text from the documents and divide it in blocks of entries (paragraphs with an idea) that can be machine-analysed.


Sources:

- SDGs targets and indicators
- SDGs website (Progress & Info)
- UN website
- Wikipedia
- Scientific publications retrieved from Scopus
- Forums
- SDG related reports and policy documents from several sources
- Backwards and forward citations in selected publications

# STRINGS Approach: Four Steps

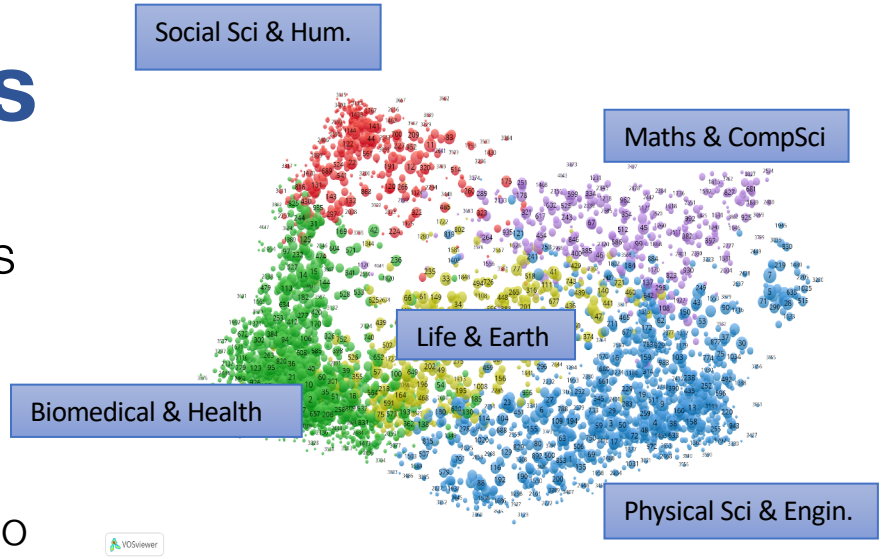
3. Selection of terms and generation of queries

- We combined two topic modelling methodologies: Textrank (Mihalcea and Tarau, 2004) and Vosviewer keywords extraction.
  - Select terms that are relevant because they occur across different text partitions, but not always (e.g. “SDG”, “Sustainability”), and are well connected to other terms
  - Keep most terms with more than one word or most frequent
  - Eliminate duplicates

- We manually assess those terms (min -> three project members)
  - Clearly related to the SDG across contexts (to vague or irrelevant removed)
  - Associated to an SDG, but may need specification
    - "habitat degradation" with "ocean" OR "sea" OR "marine”
    - "fishing resource" with "conservation" OR "sustainable"
  - Manually check all disagreements, and add associated terms where needed
- We performed a robustness check against SIRIS terms
- >1800 terms

| SDG | TERM | COMBINATION "AND" |
|---|---|---|
| (...) | (...) | (...) |
| 14 | unregulated fishing | |
| 14 | water indicator | ocean OR sea OR marine |
| 14 | water quality | ocean OR sea OR marine |
| 14 | wild food source | ocean OR sea OR marine |
| 14 | wild ocean fishing | |
| 15 | affect poor community | biodivers* OR ecosystem OR ecological |
| 15 | aichi biodiversity target* | |
| 15 | animal breed diversity | biodivers* OR ecosystem OR ecological OR sustainab* |
| 15 | animal extinction | |
| 15 | animal species | biodivers* OR ecosystem OR ecological OR sustainab* |
| 15 | animal welfare | biodivers* OR ecosystem OR ecological OR sustainab* |
| 15 | animal welfare program | |
| 15 | animal welfare project | |
| 15 | biodiversity | ecological OR ecosystem |
| 15 | biodiversity conservation | |
| 15 | biodiversity goal | |
| 15 | biodiversity indicator | |
| 15 | biodiversity loss | |
| 15 | biodiversity observation | |
| 15 | biodiversity target | |
| (...) | (...) | (...) |

# STRINGS Approach: Four Steps




4. Search publications in WoS and distribute this "seeds" across publication-level classification (Waltman & van Eck, 2012)

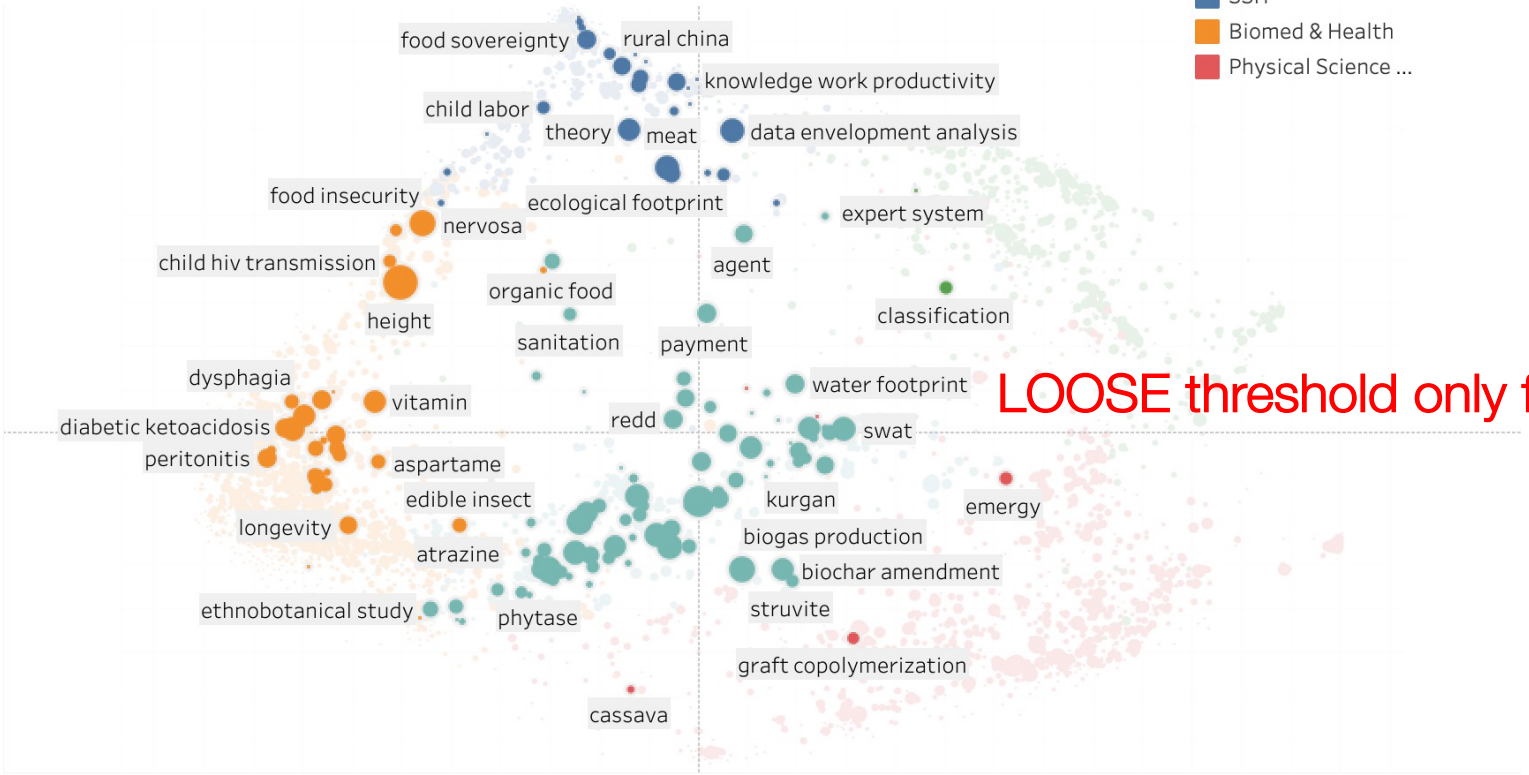
- **Assumption:** Publications that frequently cite/are cited by the publications that contain SDG terms, may also be contributing to the same SDG, even if they do not use SDG terms
- Why research areas/clusters instead of queries?
  - Science is a collaborative effort
  - Less dependent on current jargon
  - Policy language vs scientific language
  - Long term impact

## SDG 2 - End hunger, achieve food security and improved nutrition and promote sustainable agriculture

SDG: 2

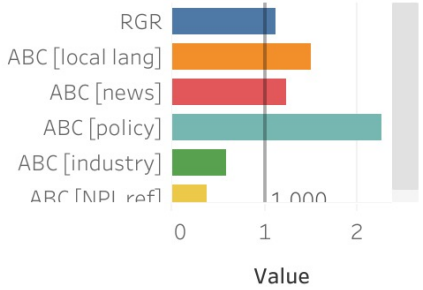
Perimeter: Loose

### Distribution of selected communities (clusters) in SDG 2 (definition: Loose)




Main Discipl
- SSH
- Biomed & Health
- Physical Science ...

### characteristics




LOOSE threshold only for next results

### Categories
(Top in selected clusters)

| Subject Category |
|---|
| Plant Sciences |
| Environmental Sciences |
| Agronomy |
| Economics |
| Soil Science |
| Entomology |
| Food Science & Technology |
| Nutrition & Dietetics |

### Selected communities (clusters) in SDG 2 (definition: Loose)

| Id | Sublabel | weight<Strings> | score<Strings> |
|---|---|---|---|
| 461 | vitamin; india; bangladesh; young child; ethiopia | 1,968 | 0.34 |
| 30 | soil organic carbon; soil microbial community; soil enzyme activity; soil carbon stock; l.. | 1,103 | 0.10 |
| 613 | percutaneous endoscopic gastrostomy; tube; nutritional risk; malnutrition; gastrosto.. | 1,074 | 0.22 |
| 1005 | adoption; conservation agriculture; malawi; crop insurance; agricultural research | 1,002 | 0.25 |
| 1662 | life cycle assessment; greenhouse gas emission; carbon footprint; food waste; lca | 889 | 0.23 |
| 361 | rhizobacteria; growth promotion; bacillus; pgpr; bacillus amyloliquefacien | 660 | 0.10 |
| 2869 | intercropping; integrated nutrient management; legume; cajanus cajan; fenugreek | 507 | 0.62 |
| 895 | soil erosion; rainfall erosivity; water erosion; polyacrylamide; rusle | 490 | 0.12 |
| 26 | height; childhood overweight; food frequency questionnaire; infancy; portion size | 437 | 0.03 |
| 1506 | organic food; functional food; front; health claim; willingness | 370 | 0.11 |

### Keywords
(Top in selected clusters)

| Noun Phrase | |
|---|---|
| effect | 22,462 |
| china | 8,677 |
| impact | 4,849 |
| soil | 4,396 |
| case study | 3,343 |
| growth | 3,147 |
| child | 2,244 |

Targets related to:

- levels of hunger
- malnutrition
- agricultural productivity
- sustainable food production systems
- genetic diversity of seeds, etc.

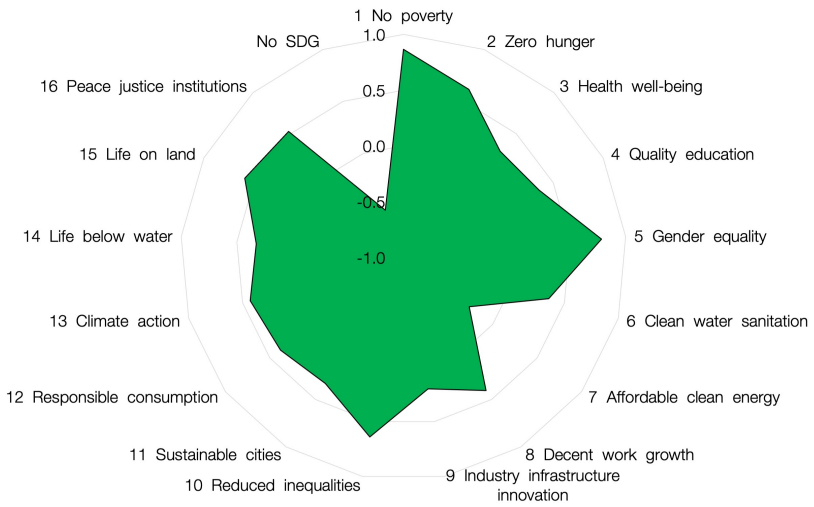
# How to evaluate what countries perform better in what SDGs?

- We collected all available/relevant indicators from: UN SDG database + SDG Index
  - For the selected **80 indicators** we did a linear transformation (1 - Best and 0 - Worst) in three periods -> 2008-2017; 2008-2012; 2013-2017

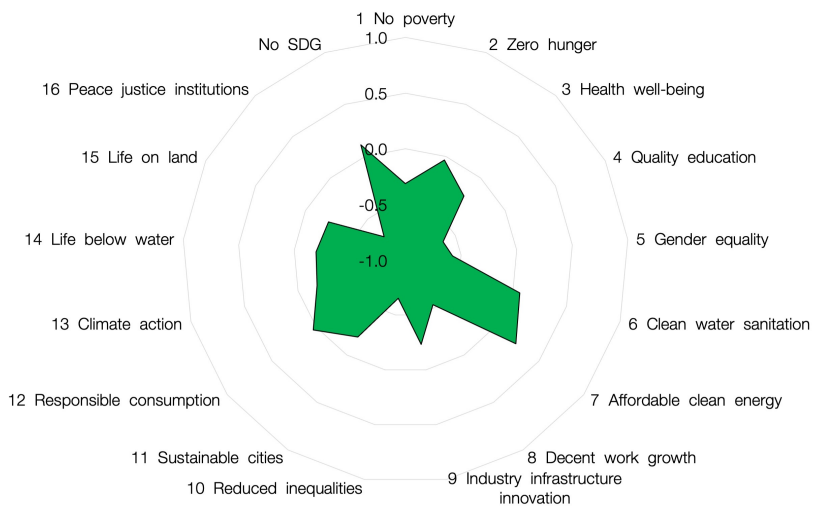
$$N_{ct} = \frac{Worst_t - x_{ct}}{Worst_t - Best_t}$$

  - For each variable, we calculated the **relative distance of each indicator/country to the frontier** of that indicator (top5% - percentile 95), and we changed all values below zero to zero.
  - We calculated **z-scores** for each relative distance to the frontier (top5%).
  - We computed a principle component analysis (**PCA**) (Jackson, 1991) for each SDG with more than one indicator available, and we forced the PCA to give us only one component per SDG.
  - We predicted the scores of all SDGs for all countries and we normalized the results between -1 and 1 (1 = Worst country in a SDG; -1 = Best country).

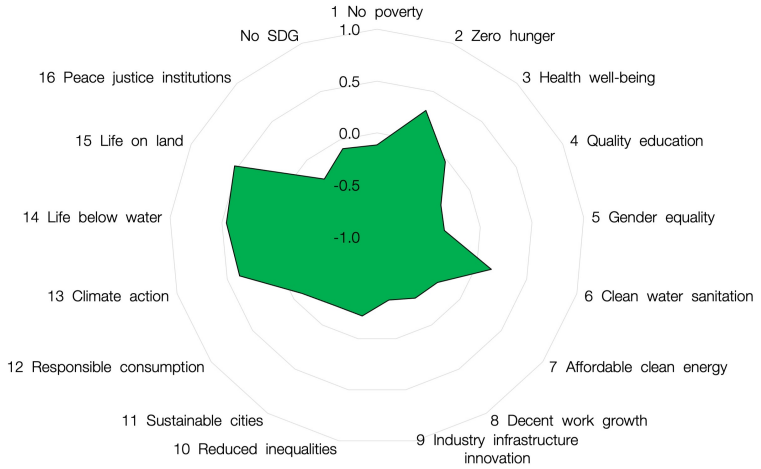
SDGs Research Specialisation in Tanzania - 2015-2019




SDGs Research Specialisation in India - 2015-2019

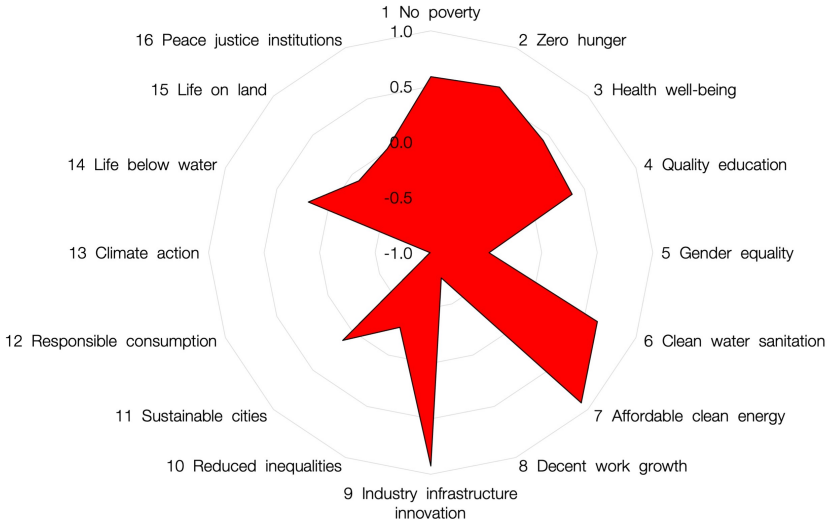



SDGs Research Specialisation in Argentina - 2015-2019



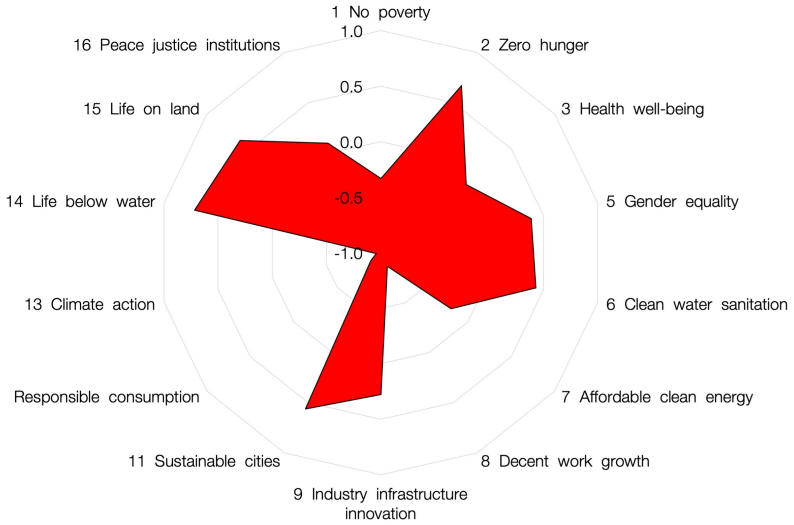

## Lower-middle income

SDGs Score (Challenges) in Tanzania - 2008-2017



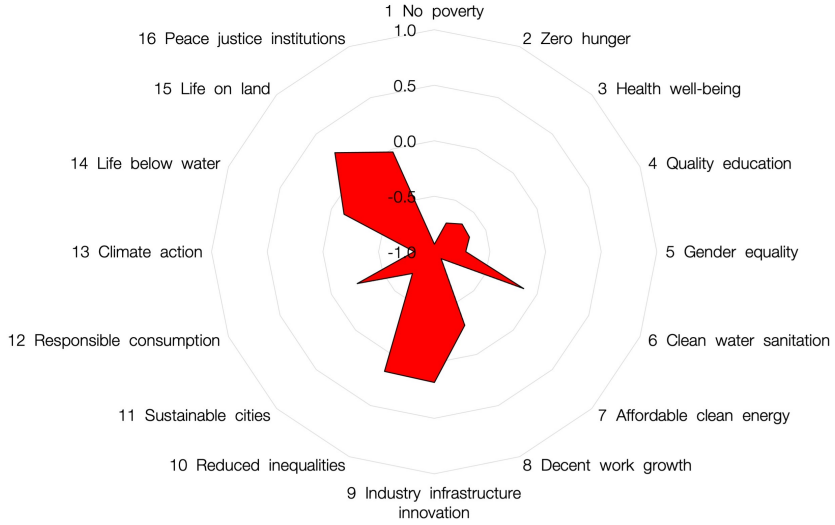

## Lower-middle income

SDGs Score (Challenges) in India - 2008-2017

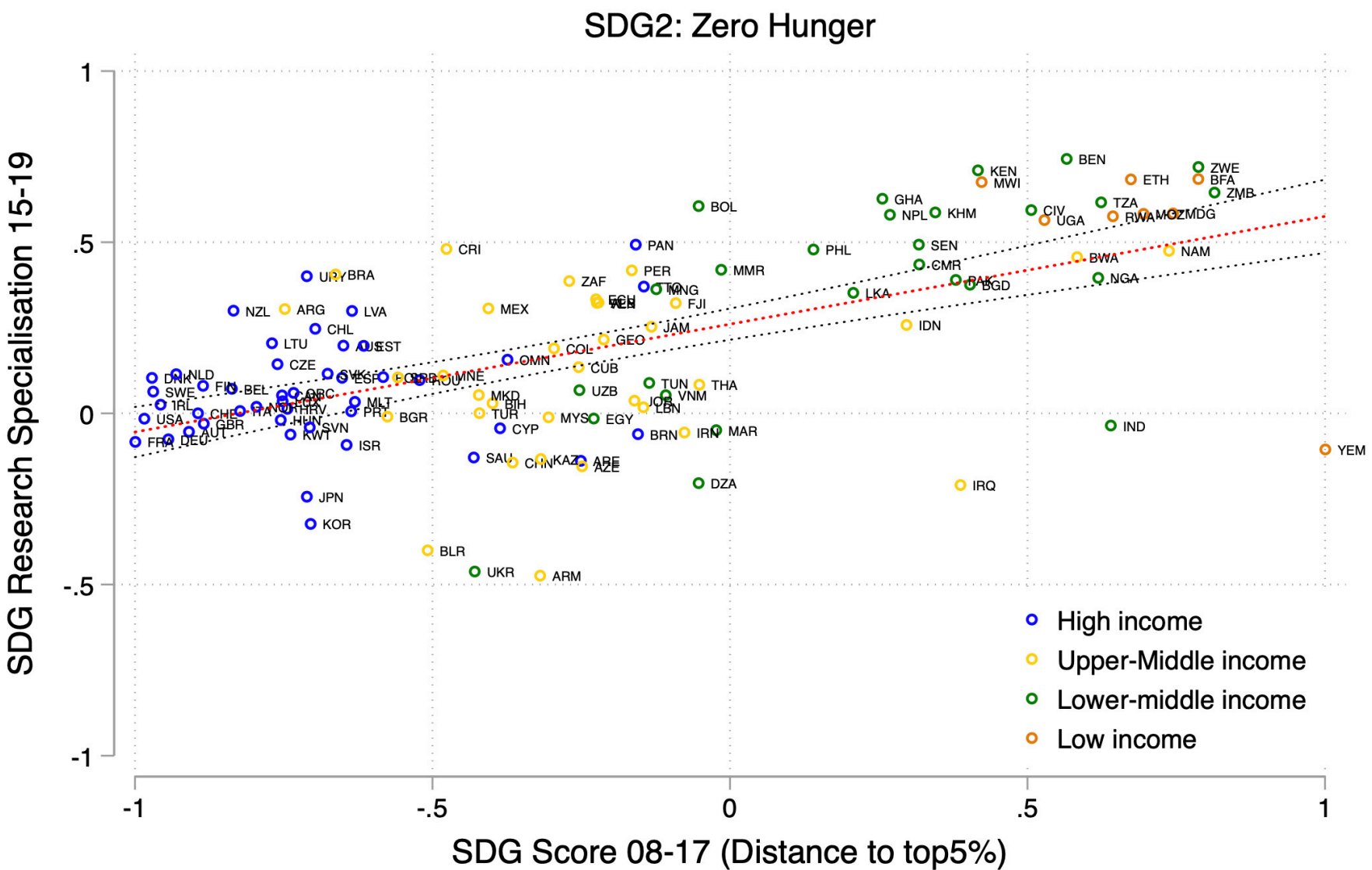



## Upper-middle income

SDGs Score (Challenges) in Argentina - 2008-2017



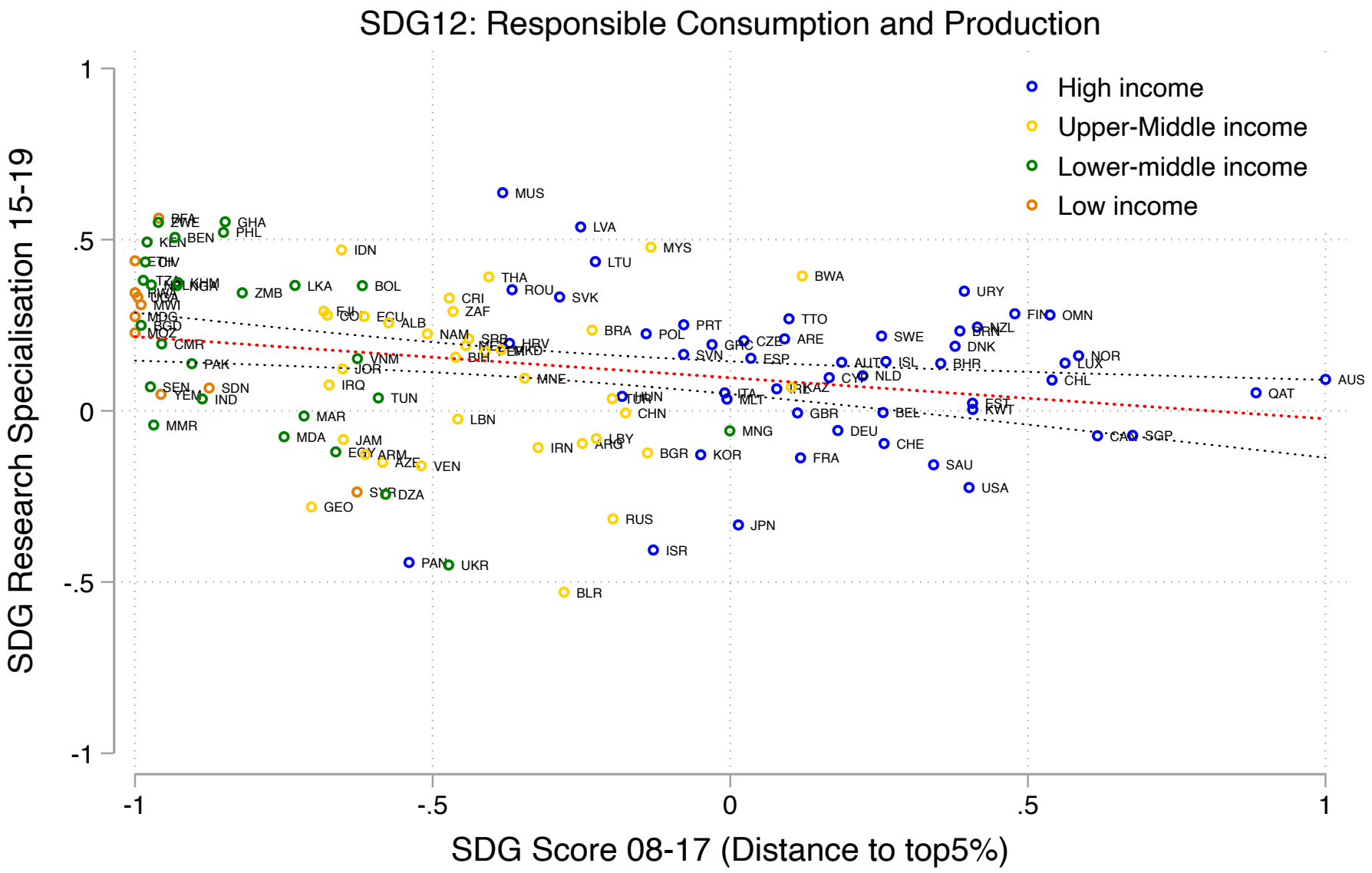
# Positive relation (alignment) in SDG 2




Targets related to:

- levels of hunger
- malnutrition
- agricultural productivity
- sustainable food production systems
- genetic diversity of seeds, etc.

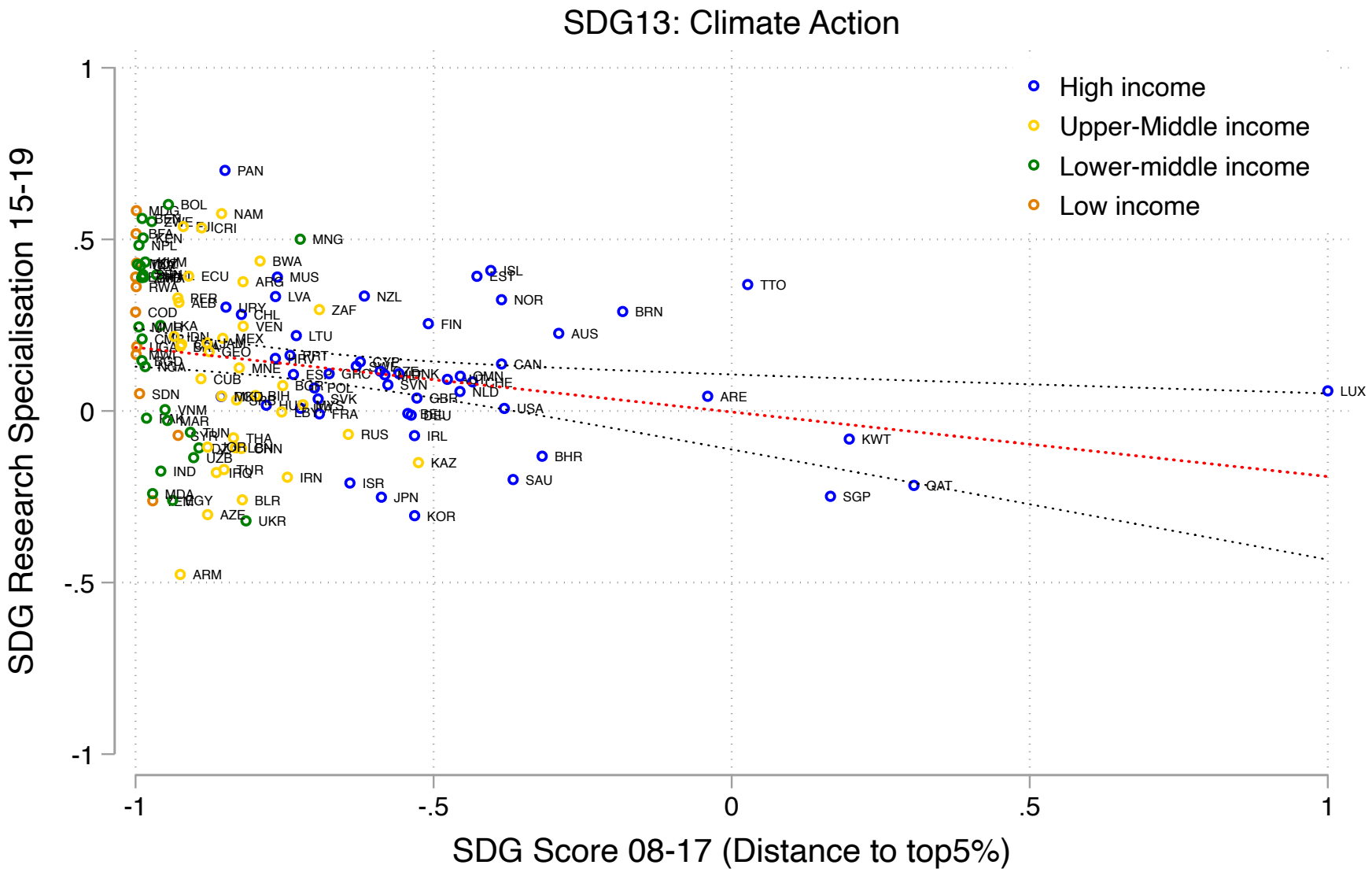
# Negative relation (misalignment) in SDG 12




Targets related to:

- Domestic material consumption per capita, (total tonnes)
- Electronic waste generated, per capita (Kg)

# Negative relation (misalignment) in SDG 13




Targets related to:

- $CO_2$ emissions embodied in imports
- Energy-related $CO_2$ emissions
- Integrate climate change measures into national policies, strategies and planning (13.1, 13.2, 13.3 no data)

# Different patterns depending on the SDG

| | Pwcorr t ↔ t | RQ1 t → t | RQ2 t → Δ | RQ3 Δ → Δ |
|---|---|---|---|---|
| 1 No poverty | + | | | |
| 2 Zero hunger | + | | + | |
| 3 Health well-being | + | | | + |
| 4 Quality education | - | - | | |
| 5 Gender equality | | | | + |
| 6 Clean water sanitation | + | + | + | |
| 7 Affordable clean energy | - | | | |
| 8 Decent work growth | | | | |
| 9 Industry infrastructure innovation | | | + | - |
| 10 Reduced inequalities | | | | |
| 11 Sustainable cities | | | | |
| 12 Responsible consumption | - | | | |
| 13 Climate action | - | | + | + |
| 14 Life below water | | | | |
| 15 Life on land | | - | - | |
| 16 Peace justice institutions | | - | | |

**Note:** Multivariate regression analysis (OLS), we use scientific specialisation (2015-2019) by SDG/country as our dependent variable, and SDG Score (2013-2017) as our main independent variable. In model (1) we control for previous scientific specialisation due to the path-dependant nature of scientific production. In models (2 and 3) we work with growth rates and control for country's scientific productivity.

- Controlling for other factors, we found **negative or inconclusive relationship between research prioritisation and SDG challenges for the majority of SDGs.**
- **Historical patterns of research specialisation** matter more than SDG challenges (SDG1, SDG2, SDG3)
- Countries that have major challenges in **SDG6 (Clean Water and Sanitation)** are relatively (or becoming) specialised (or becoming) in research related to this SDG.
- Countries that have major challenges in **SDG15 (Life on Land) are NOT** relatively (or becoming) specialised in research related it.

# In a nutshell

- We have developed **new methods** to analyse the alignment between countries' research priorities and their SDG-related challenges. Results might help policy makers to understand these relations and bring **directionality** to the debate.

- While most SDG challenges are worst in lower income countries, only a small fraction of SDG related research is performed in those regions
  - Nevertheless, lower income countries are on average specialised in SDG related research.
  - Global effort to allocate research funds to lower income contexts?

- At the SDG level, we find different patterns depending on the SDG:
  - Countries with higher challenges relating to **SDG1** (No poverty), **SDG2** (Zero hunger), **SDG3** (Good health and well-being) and **SDG6** (Clean water and sanitation) are relatively specialised in research in these areas -> **Alignment strongly related to historical patterns of research specialisation**
  - **For all other SDGs, we found a misalignment or inconclusive relationship**. For example, the countries with the most unsustainable consumption/production patterns, generate more CO2 emissions, and contribute more to biodiversity loss, are usually higher income countries that are not specialized in research related to **SDG12** (Responsible consumption and production), **SDG13** (Climate action) or **SDG15** (Life on land).

# Limitations

- Uncertainty and ambiguity of our estimates of SDG research priorities (Armitage et al., 2020) and SDG challenges (Miola and Schiltz, 2019)

- Synergies & trade-offs (e.g. Sharlemann et al. 2020)

- The marginal impact of increasing research investments in areas related to a certain SDG on the improvement of that SDG might be different across SDGs (e.g., local research in health (SDG3) may lead to significant improvements in the health outcomes of a country. In contrast, more local research on Poverty (SDG1) may not lead to similar marginal improvements.

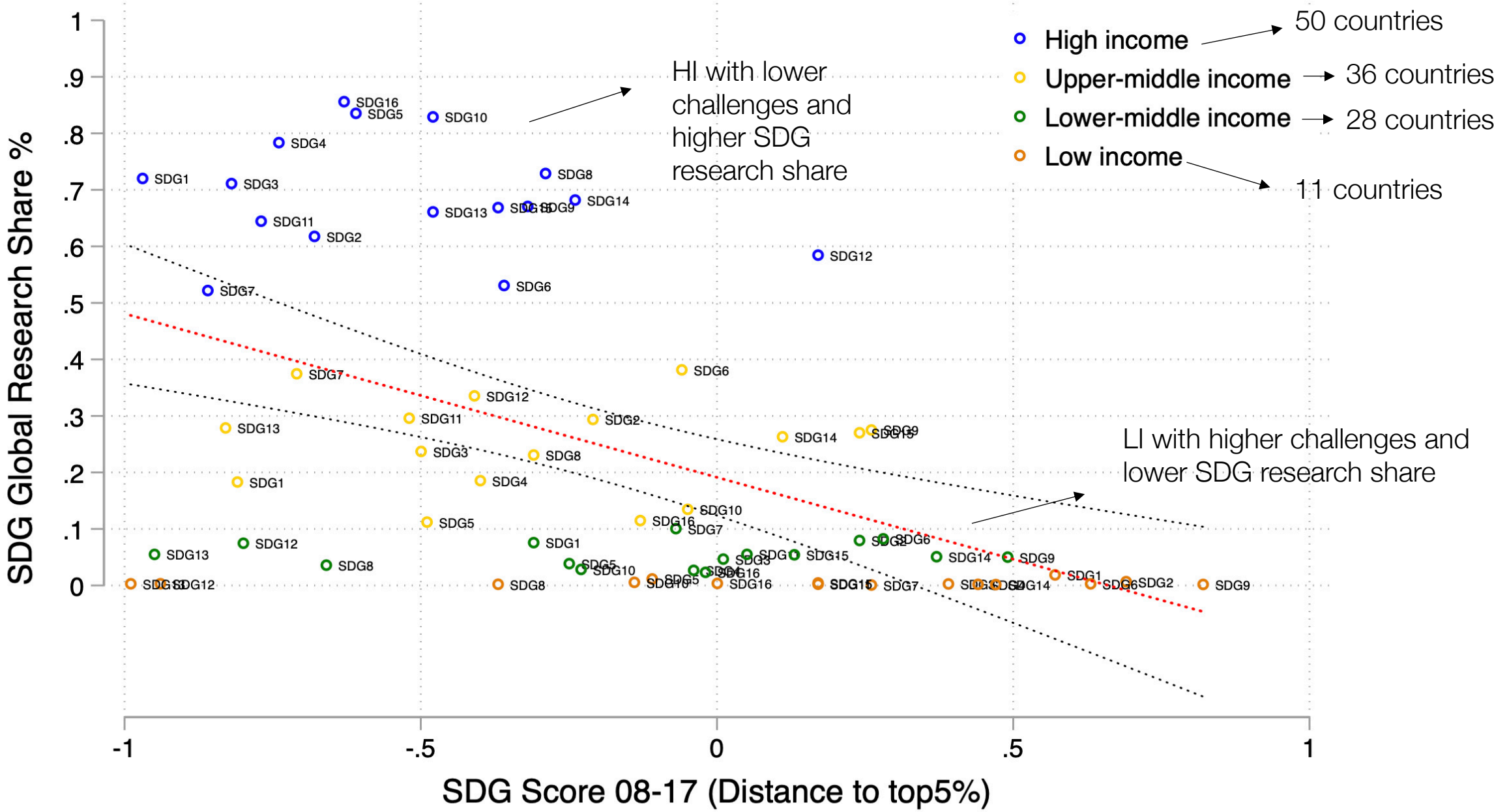
# Thank you!

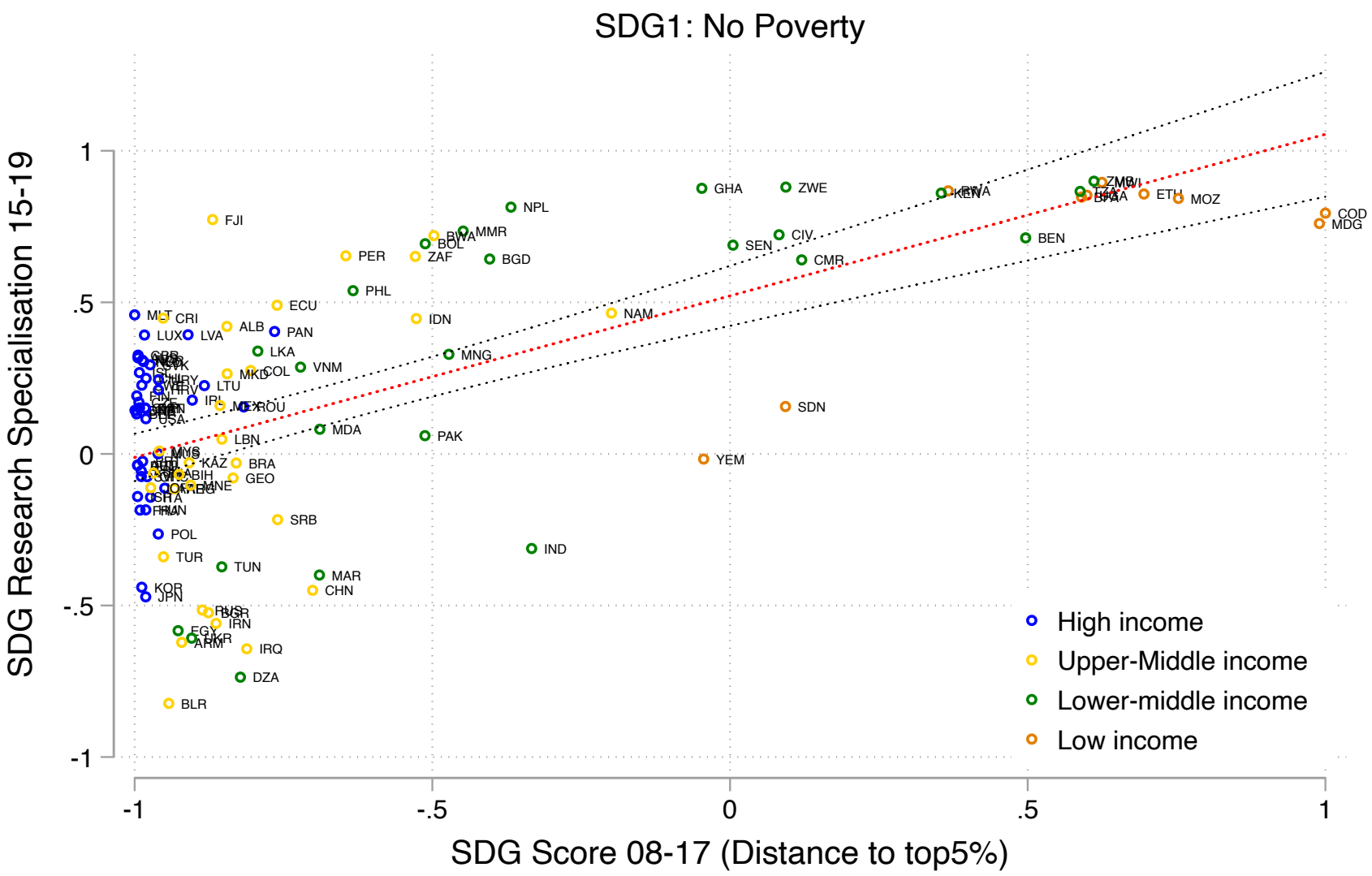
If you want to know more about this research and STRINGS check tomorrow's special session (SS -2C1) on:

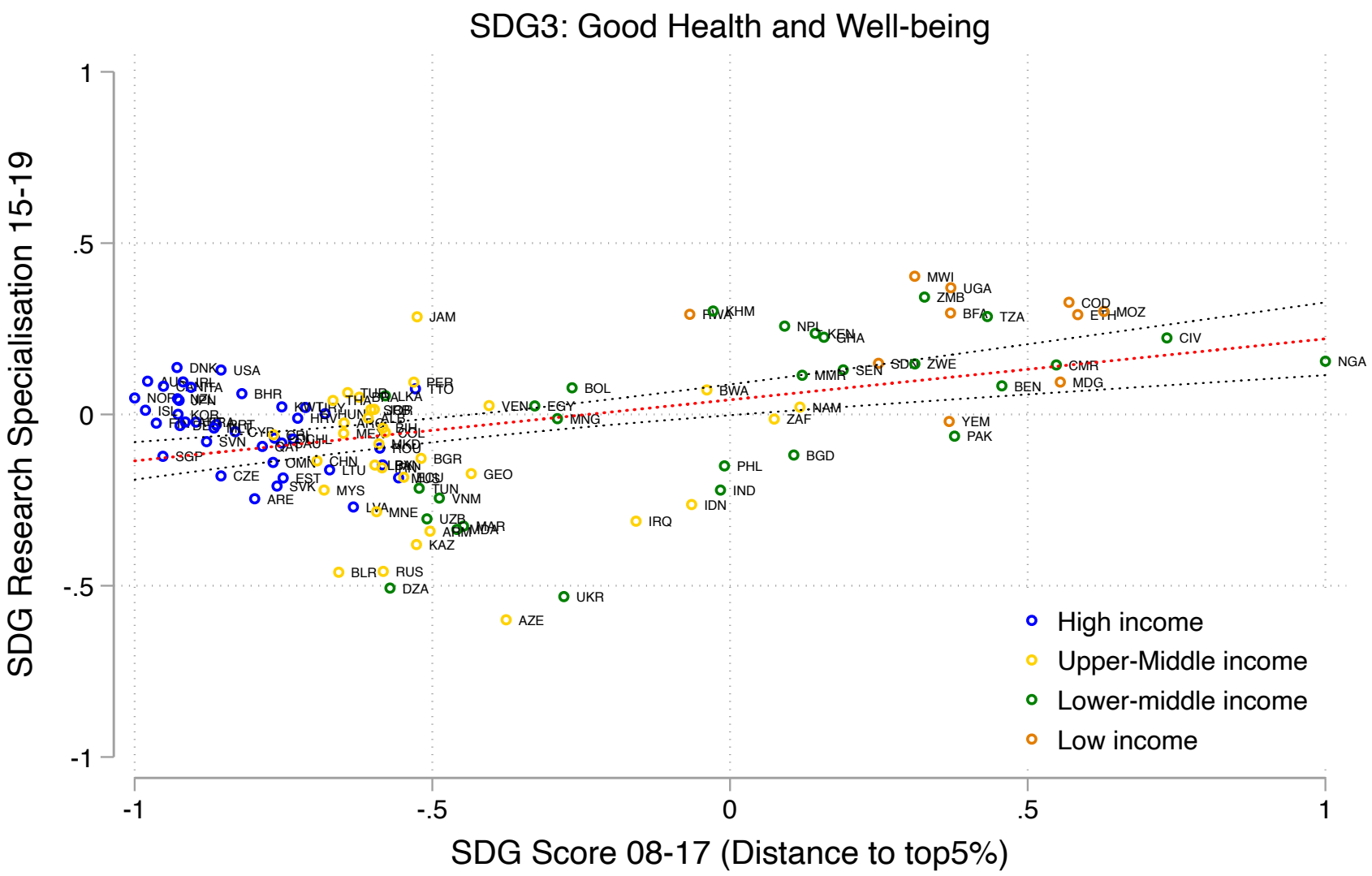
**”Mapping and interpreting the relation between STI and the SDGs"**

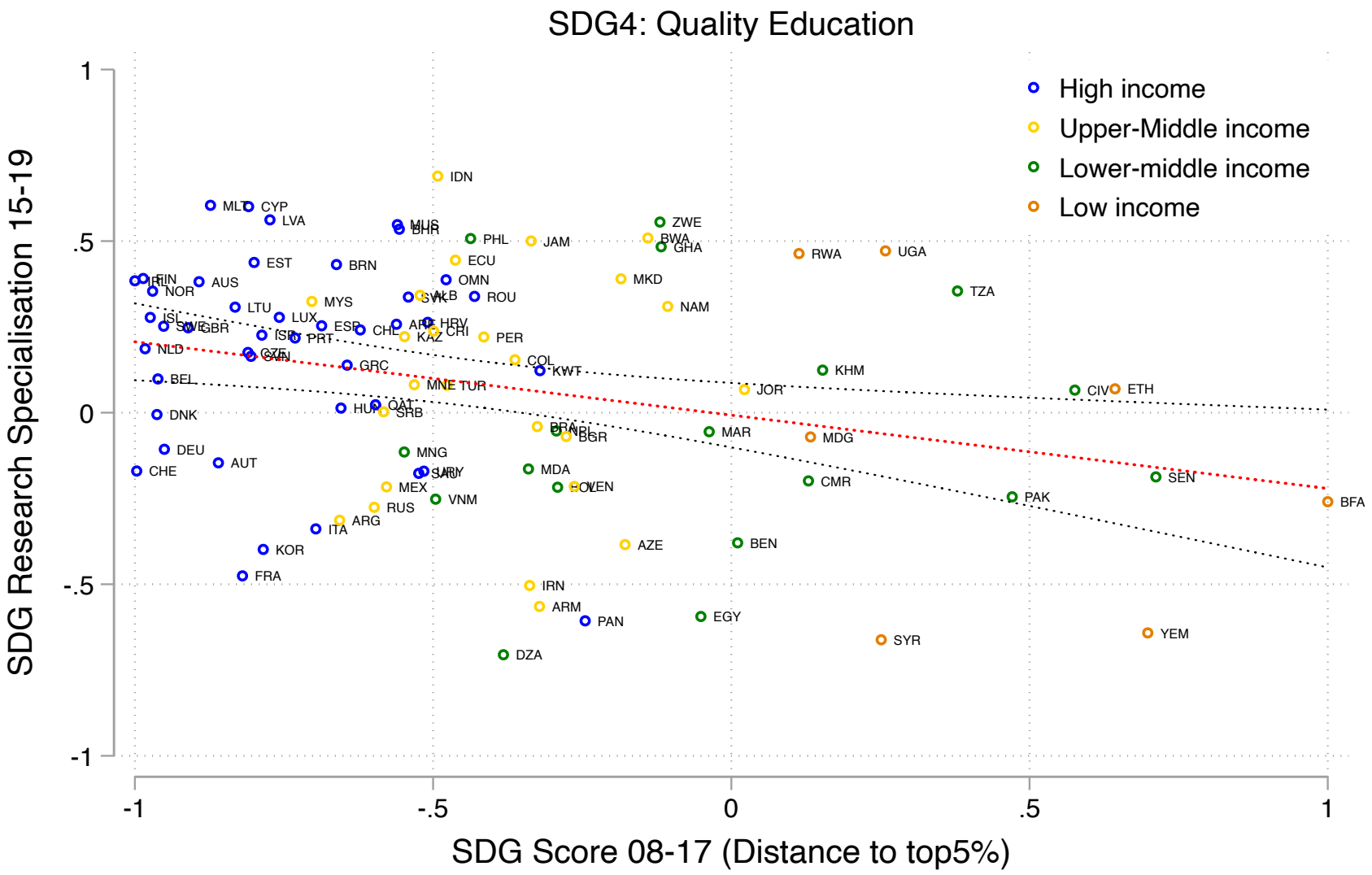
Thursday 4 November at 10:30-12:00 local time (GMT+6).

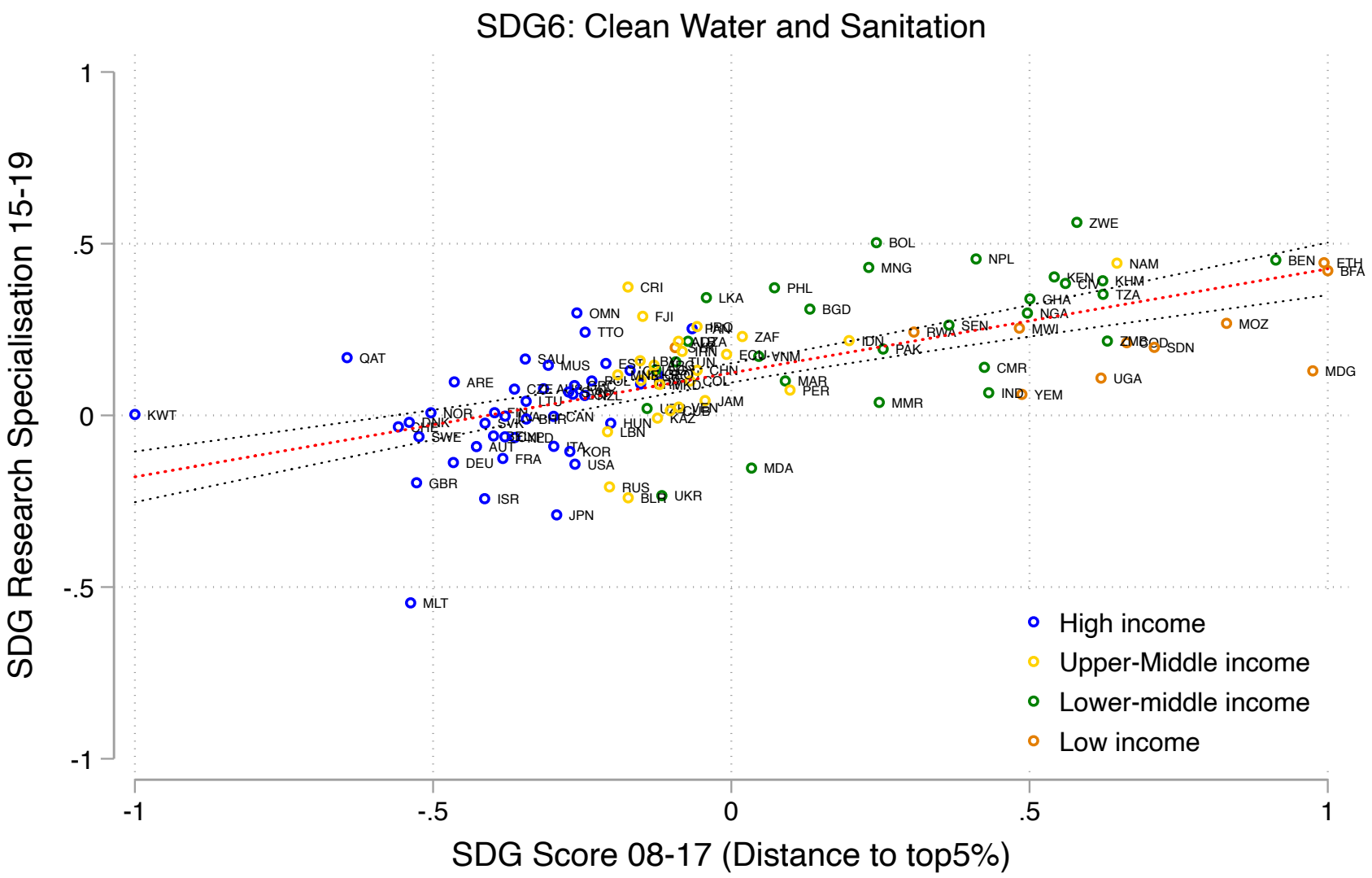
# Misalignment between SDG challenges and SDG research by income group

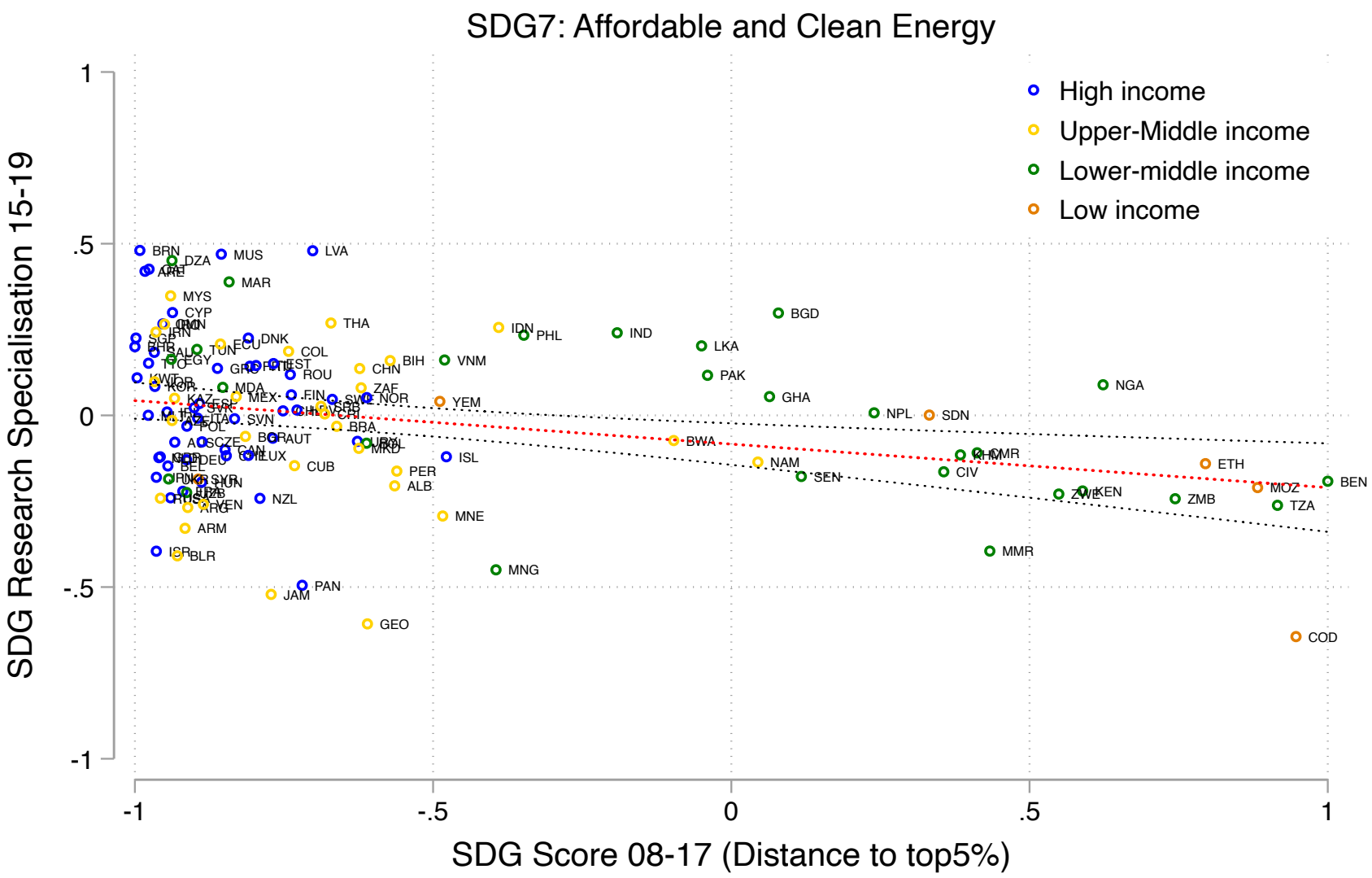


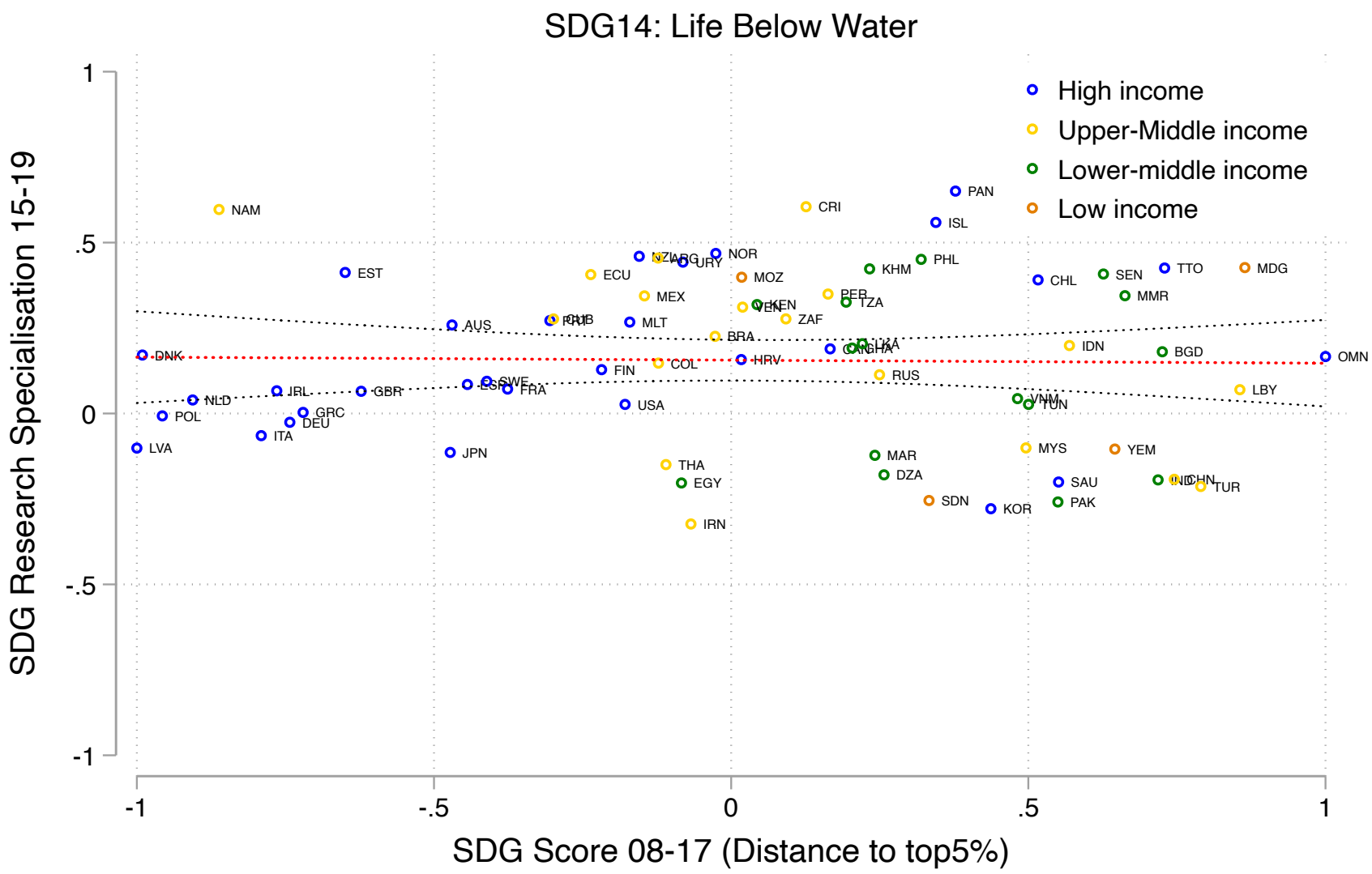

SDG1: No Poverty
SDG Research Specialisation 15-19
SDG Score 08-17 (Distance to top5%)
1
.5
0
-.5
-1
High income
Upper-Middle income
Lower-middle income
Low income
GHA
ZWE
NPL
FJI
MMR
BOL
ZAF
PER
BGD
PHL
ECU
CRI
ALB
PAN
LUX
LVA
LKA
VNM
COL
MKD
LTU
IDN
MNG
NAM
SDN
MDA
PAK
LBN
BRA
GEO
YEM
SRB
POL
TUR
TUN
IND
MAR
CHN
KOR
JPN
IRN
IRQ
DZA
BLR
SEN
CIV
CMR
BEN
KEN
ETH
MOZ
COD
MDG

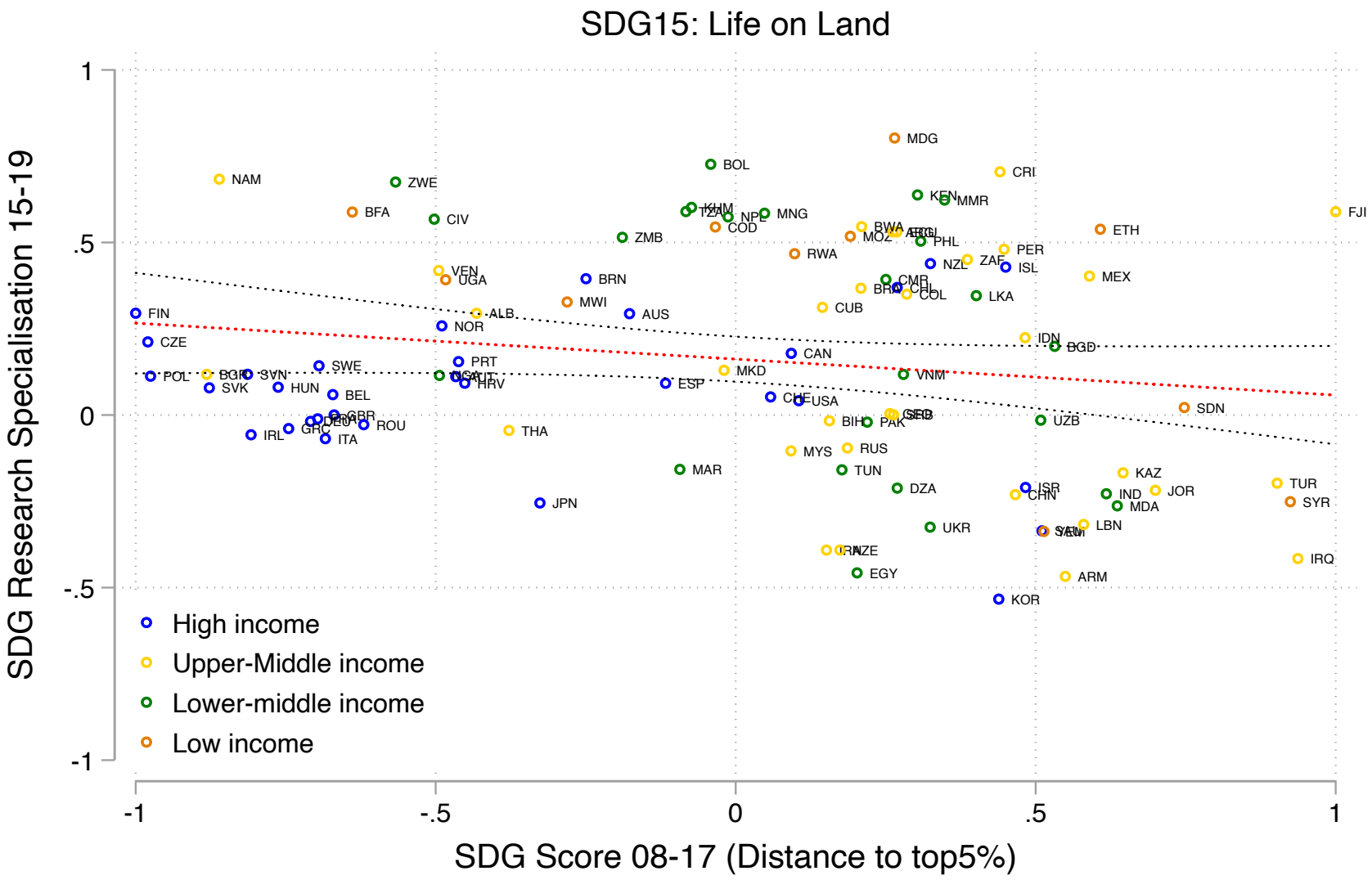

SDG3: Good Health and Well-being
SDG Research Specialisation 15-19
SDG Score 08-17 (Distance to top5%)
1
.5
0
-.5
-1
High income
Upper-Middle income
Lower-middle income
Low income
JAM
DNK
USA
BHR
PER
TTO
BOL
VEN
EGY
MNG
BWA
KHM
RWA
NPL
KEN
GHA
MWI
UGA
ZMB
BFA
TZA
COD
ETH
MOZ
CIV
NGA
CMR
MDG
BEN
YEM
PAK
NAM
ZAF
BGD
PHL
IND
IDN
IRQ
GEO
BGR
CHN
LTU
CZE
EST
SVK
MYS
ARE
LVA
MNE
UZB
KAZ
BLR
RUS
DZA
UKR
AZE
SGP
SVN
OMN
TUN
VNM
MAR
MDA
ZWE
SEN
MMR


SDG4: Quality Education
High income
Upper-Middle income
Lower-middle income
Low income
SDG Research Specialisation 15-19
SDG Score 08-17 (Distance to top5%)
1
.5
0
-.5
-1
IDN
MLT
CYP
LVA
MUS
BHR
PHL
JAM
ZWE
BWA
GHA
RWA
UGA
EST
BRN
ECU
OMN
MKD
TZA
FIN
AUS
NOR
MYS
SVK
ROU
NAM
LTU
LUX
ESP
CHL
ARE
HRV
CRI
KAZ
PER
ISL
SWE
GBR
ISR
PRT
NLD
CZE
SVN
GRC
COL
KWT
KHM
BEL
MNE
TUR
JOR
CIV
ETH
DNK
HUN
QAT
SRB
BRA
BGR
MAR
MDG
DEU
MNG
AUT
CHE
URY
SAU
MDA
MEX
COL
VEN
CMR
SEN
PAK
BFA
RUS
VNM
ARG
ITA
AZE
BEN
KOR
FRA
IRN
ARM
PAN
EGY
SYR
YEM
DZA


SDG6: Clean Water and Sanitation
SDG Research Specialisation 15-19
SDG Score 08-17 (Distance to top5%)
1
.5
0
-.5
-1
-1
-.5
0
.5
1
High income
Upper-Middle income
Lower-middle income
Low income
ZWE
BOL
MNG
NPL
NAM
BEN
ETH
BFA
KEN
CMR
KHM
TZA
CRI
LKA
PHL
BGD
GHA
NGA
OMN
FJI
TTO
SEN
MWI
MOZ
ZAF
IDN
PAK
SDN
QAT
SAU
MUS
VNM
CHN
MDG
ARE
MAR
PER
UGA
COL
JAM
MMR
IND
YEM
KWT
NOR
CAN
HUN
KAZ
LBN
KOR
USA
DEU
FRA
MDA
GBR
RUS
UKR
ISR
JPN
MLT
High income
Upper-Middle income
Lower-middle income
Low income


SDG7: Affordable and Clean Energy
High income
Upper-Middle income
Lower-middle income
Low income
SDG Research Specialisation 15-19
SDG Score 08-17 (Distance to top5%)
1
.5
0
-.5
-1
BRN
DZA
MUS
LVA
MAR
MYS
CYP
THA
IDN
PHL
IND
BGD
LKA
PAK
GHA
NGA
NPL
SDN
DNK
COL
BIH
VNM
CHN
ZAF
NOR
YEM
MDA
ROU
BWA
NAM
SEN
CIV
CMR
ETH
BEN
MOZ
TZA
ZMB
KEN
ZWE
ISL
PER
ALB
CUB
NZL
MNE
ARM
BLR
MMR
MNG
PAN
JAM
GEO
COD


SDG14: Life Below Water
High income
Upper-Middle income
Lower-middle income
Low income
SDG Research Specialisation 15-19
SDG Score 08-17 (Distance to top5%)
1
.5
0
-.5
-1
NAM
EST
ECU
MEX
AUS
MLT
DNK
NLD
POL
LVA
JRL
GBR
GRC
DEU
ITA
JPN
FIN
COL
USA
THA
EGY
IRN
NOR
URY
MOZ
BRA
HRV
ZAF
CRI
PAN
ISL
PHL
KHM
PER
TZA
RUS
MAR
DZA
SDN
KOR
SAU
PAK
MYS
YEM
CHL
SEN
TTO
MDG
MMR
IDN
BGD
OMN
LBY
TUN
TUR


SDG15: Life on Land
SDG Research Specialisation 15-19
SDG Score 08-17 (Distance to top5%)
1
.5
0
-.5
-1
High income
Upper-Middle income
Lower-middle income
Low income
MDG
BOL
CRI
NAM
ZWE
KEN
MMR
BFA
CIV
MNG
NPL
COD
ZMB
BWA
MOZ
PHL
ETH
FJI
RWA
PER
NZL
ZAF
ISL
VEN
UGA
BRN
CMR
MEX
MWI
COL
LKA
CUB
FIN
ALB
AUS
NOR
CZE
IDN
BGD
CAN
PRT
SWE
POL
SVK
HUN
HRV
MKD
VNM
ESP
BEL
CHE
USA
SDN
GBR
ROU
BIH
PAK
UZB
IRL
GRC
ITA
THA
MYS
RUS
MAR
TUN
KAZ
DZA
ISR
CHN
IND
JOR
TUR
SYR
JPN
MDA
UKR
LBN
NZE
IRQ
EGY
ARM
KOR