How to Compare Data from Very Different Sources: A 4-level Longitudinal Model of Institutional Trust

By Claire Durand,
Luis Patricio Peña Ibarra, Nadia Rezgui & David Wutchiett

Presented at the 12th International Multilevel conference, April 9-10, Utrecht University, Utrecht, Netherlands.

Outline

- Institutional trust outside the western world - Data and challenges
  - Various data sets and measures
  - Various political/electoral/sociological & economic situations
- A multilevel approach and its problems
- How does it work?
- Results:
  - Description
  - Multilevel analysis
- Discussion and conclusion

The goal

- Similar concepts -- trust in institutions in our case -- are measured using different question wordings and scales.
- We have samples at the different levels, i.e., measures, respondents, years and countries.
- Our goal is to combine all the information on institutional trust from all the international survey projects and keep as much information as possible in order to be able to compare between countries, regions, etc.

Why use a multilevel approach?

- We can assess the different effects at the level where they operate, within individuals, at the individual level, over time and at the country level
- We do not have to deal with missing values and keep only the cases where the same information is available for all the cases, years, or countries.
- We can model cross-level effects,
  - Like the possible effect of time on trust in the army.
  - Or the impact of age-group or of time on trust in religion.
  - And a different impact of time in different regions.

Data sources

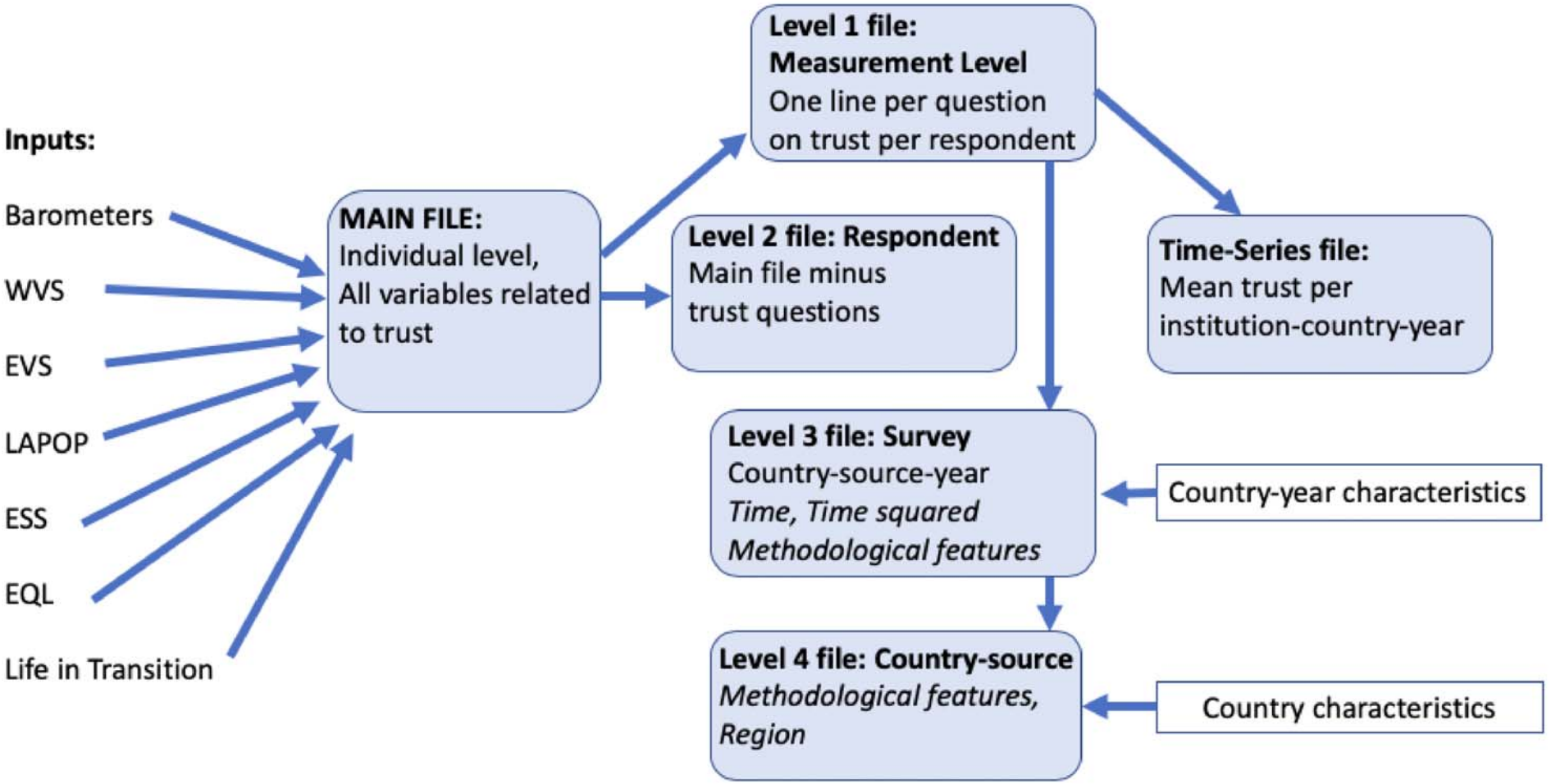
1 Latinobarometer	356
2 Asiabarometer	50
3 East_Asiabarometer	50
4 Arabarometer	37
5 Africabarometer	100
6 Lapop	169
7 WVS	178
8 Caucasus Barometer	17
9 Consolidation of democracy	13
10 European quality of life	57
11 European social survey	80
12 European values study	45
13 Life in transition	80
14 New baltic barometer	9
15 New Europe barometer	52
16 New Russia barometer	10
18 South Asia Barometer	10
Total	1313

- 1323 surveys collected from 1995 to 2017
- Restricted mostly to non-WEIRD (western, educated, industrialized, rich & democratic) countries.
- 17 different sources:
  - Barometers
  - World & European Value Surveys
  - LAPOP (Latin America)
  - Many european sources for Eastern Europe

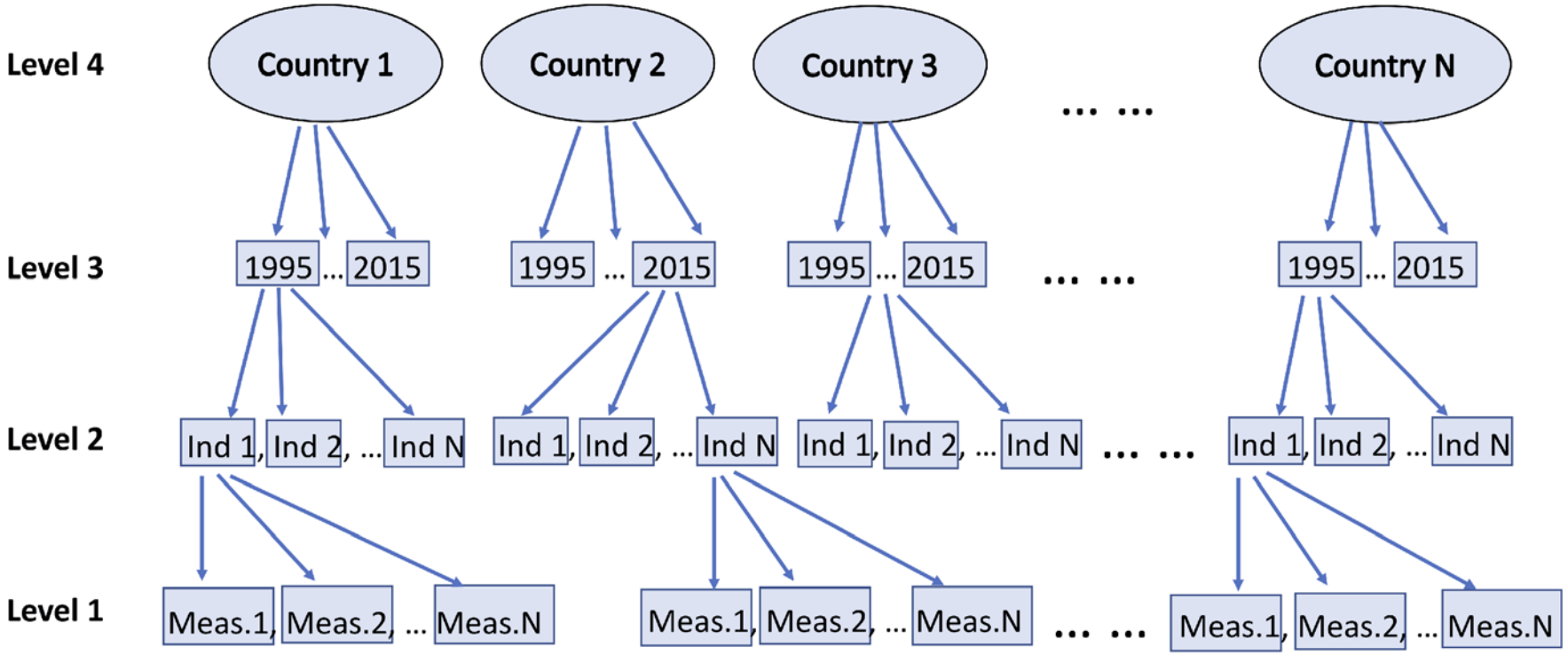
Synthesis of the data (excluding missings)

- Level 1: Measures: 21,124,464 questions on trust
- Level 2: 1,820,693 respondents
- Level 3: Countries-source-years: 1323 surveys
  - Country-years: 1082 (241 double/triple)
- Level 4: Countries-source: 364 units
  - Countries: 143 (219 double/triple)
- Time-series: 11,268 country-year-institutions.

Concretely, synthesis of the process



A 4-level longitudinal model with repeated measures



A multilevel approach and its problems

At level 1, the measurement level

- There are 133 different institutions for which trust is asked in the surveys to date.
- Questions asked in a survey are samples of all the questions that can be asked to measure a concept.
- The method could be used for other concepts like attitude towards the protection of environment, towards democracy, etc.

A multilevel approach and its problems

At level 1, the measurement level

- The 133 institutions have to be grouped.
- Conceptual criteria :
  - Political (6): President, government, parliament, elections, political parties, international & regional organizations;
  - Administrative (4): Army, police, public administration, Legal system/fight corruption;
  - Civil Society (4): Media, Religious organizations, Trade Unions, NGO;
  - Economic (2): Financial institutions, enterprises.
- Empirical criteria:
  - The institutions have similar averages & Std within category and project.
  - The number of answers is sufficient in each grouping.

A multilevel approach and its problems

At level 1, the measurement level

- The measures have to be on the same **scale**, which means
  - Recode all scales so that the highest number corresponds to higher trust.
  - Expand or shrink the scales to a 1 to 7 scale.
- Different question wordings have to be controlled for.
  - Trust vs Confidence: In our case, not a problem because most surveys are not conducted in English. In most other languages, there is only one word for trust.
- Those are characteristics of survey projects and therefore are controlled at level 4.

A multilevel approach and its problems

At level 2, the respondents' level: ex-post harmonization issues

- The respondent level is where most harmonization problems occur. The only non problematic variable – for now – is sex.
- Some projects ask age in years, others in categories. Categories may not be the same in different projects.
- Level of education: the educational systems vary. May be difficult to place technical training. Not asked in 4 surveys.
  - Harmonize in 5 categories:
    - No formal education, primary, secondary, technical, university.

A multilevel approach and its problems

At level 2, the respondents' level: missing, methods

- Missing values:
  - Occupation: 44% not asked
    - Common categories: employed, out of work, homemaker, retired, student.
  - Subjective Income: (62% not asked)
    - Four categories from "sufficient, can save" to "not sufficient, have big problems"
  - Attitudes as independent variables.
    - Satisfaction with democracy: 60% not asked
- Methods:
  - item non-response: 74% answered all questions

A multilevel approach and its problems

At level 3: country-source-year = survey level

- Multiple surveys are conducted over time in each country. The time level is intermediary, i.e., both
  - nested within country-source and
  - having respondents nested within each survey.
- Add variables identifying time and time squared
- Introduce methodological characteristics of surveys: See Survey Data Recycling (SDR) Project
- Introduce country characteristics that vary over time.

A multilevel approach and its problems

At level 4: country or country-source?

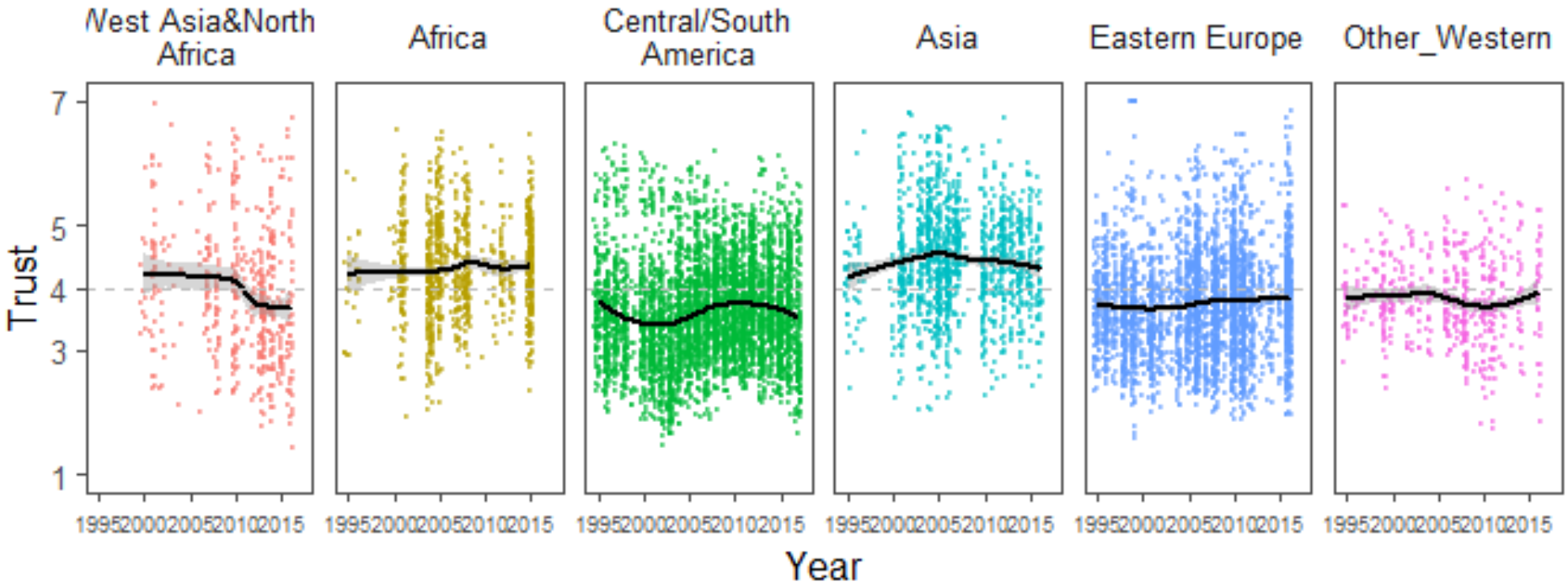
- In some countries, more than one survey project conducts surveys,
  - We need to be able to test whether, on average, there are differences according to the source of data.
  - Solution: The highest level is a “country-source” level.
- Which allows for:
  - Adding variables identifying the source of data and the methodological features -- answer scale, question wording, etc. -- of the different projects.
- Fixed effects: Region, fixed country characteristics.

An extra step: match external data

- Different sources of data can be matched at level 3 -- country-year -- or at level 4 -- country:
  - Political: Polity index, World Governance indicators, Participation in elections, etc.
  - Economic: Gini, GDP per capita,
  - Sociological: Proportion of urban population, ethnic and religious diversity, proportion of young & old people, etc.
  - Quality of governement data
  - V-DEM project data
- The main problems:
  - Some indices do not vary enough over time: preferable to introduce them at the country level.
  - Lack of data outside the western world.

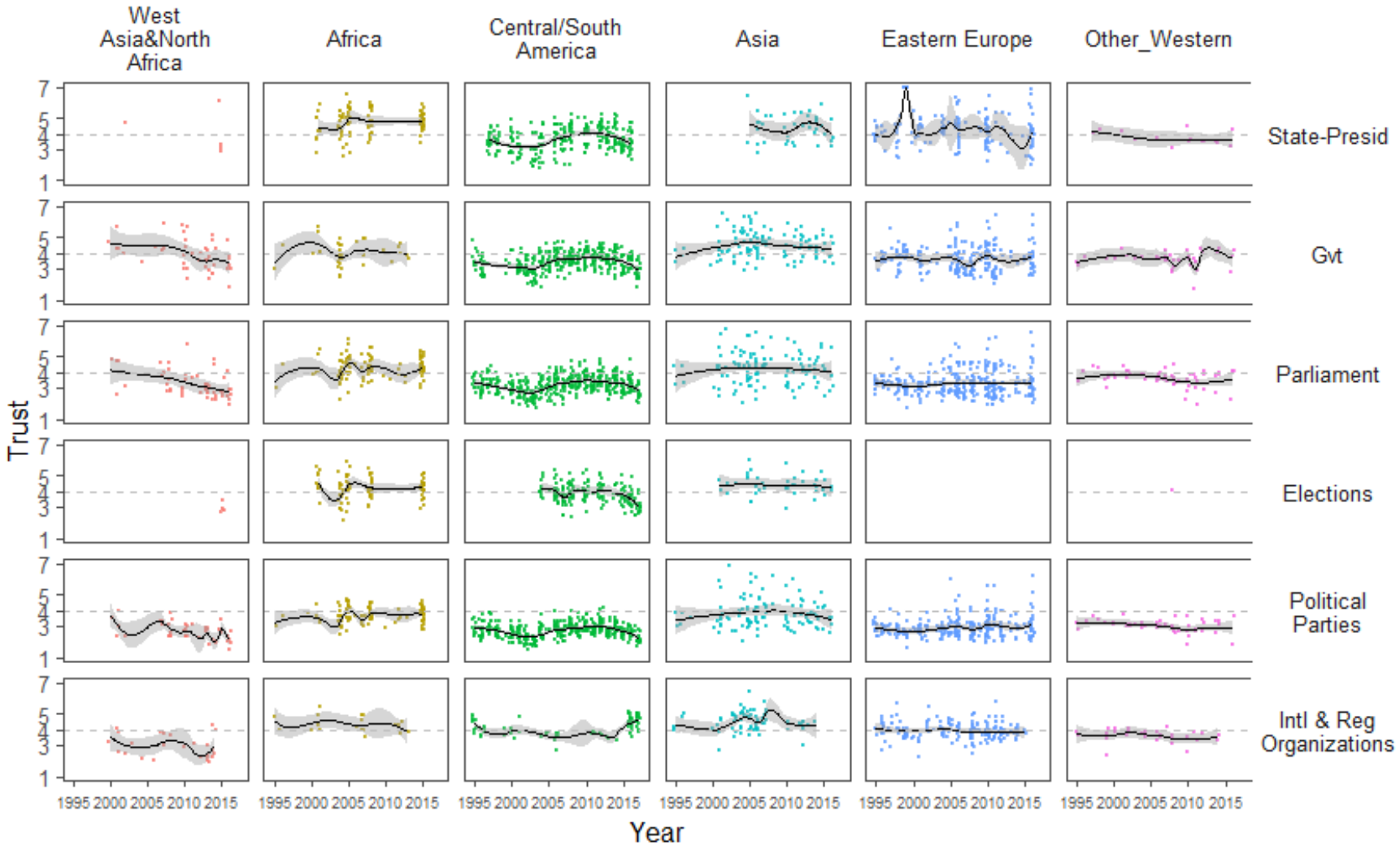
Results: First, describe

Trust over time by region

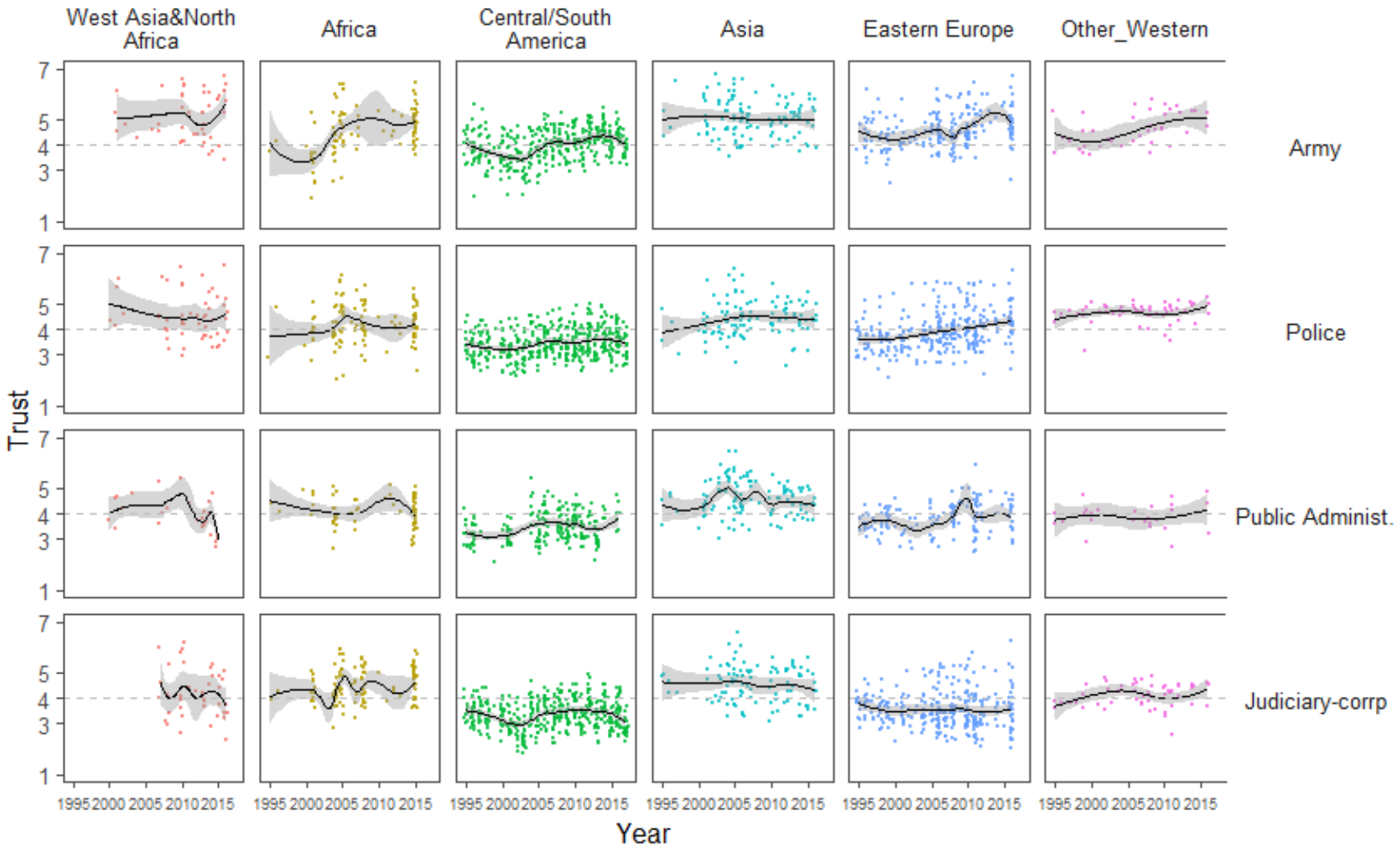


- Each point represents an estimate of trust in a given institution in a given country and year.
- Differences between regions & much variation within region by institution & country.

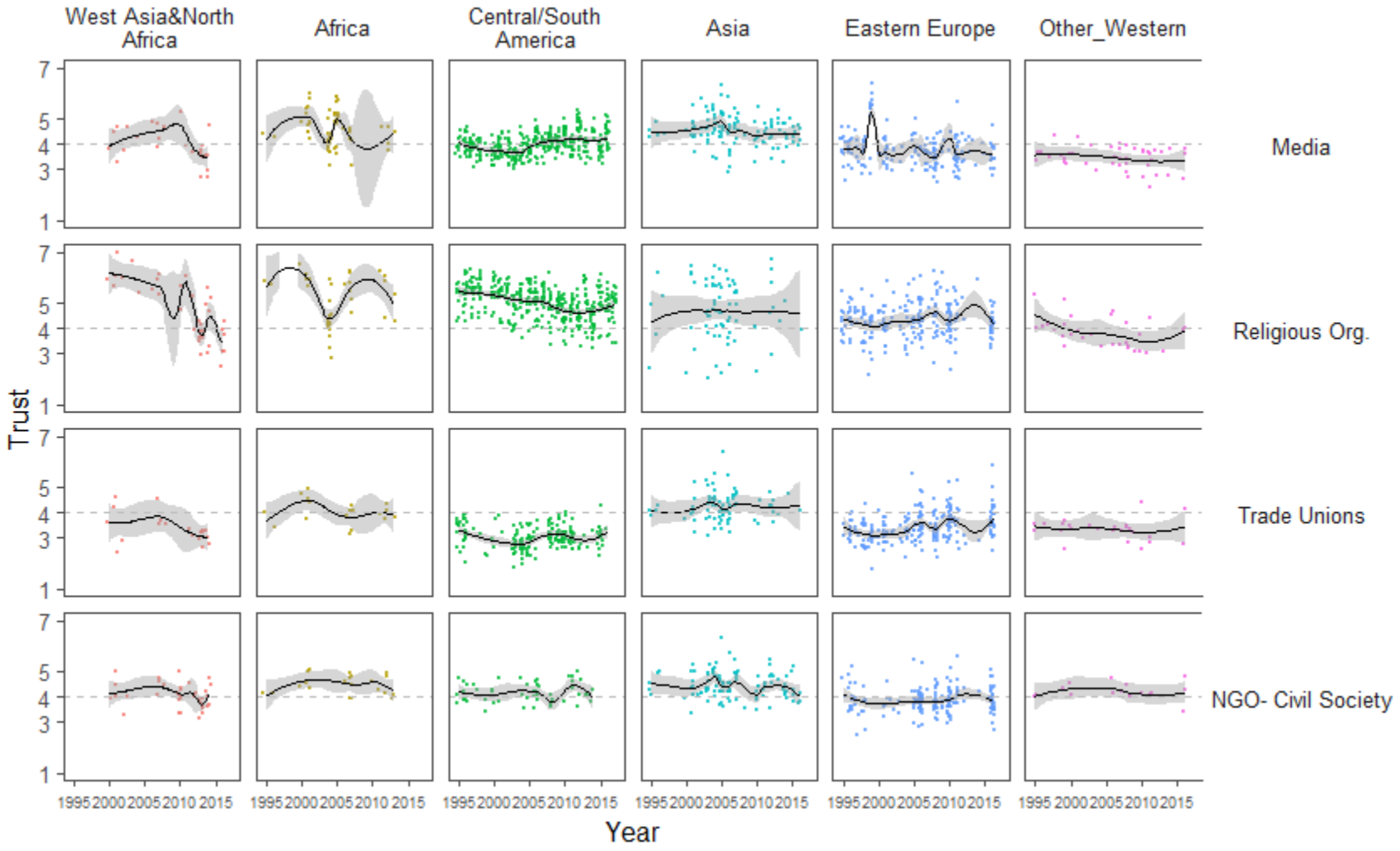
Trust in political institutions



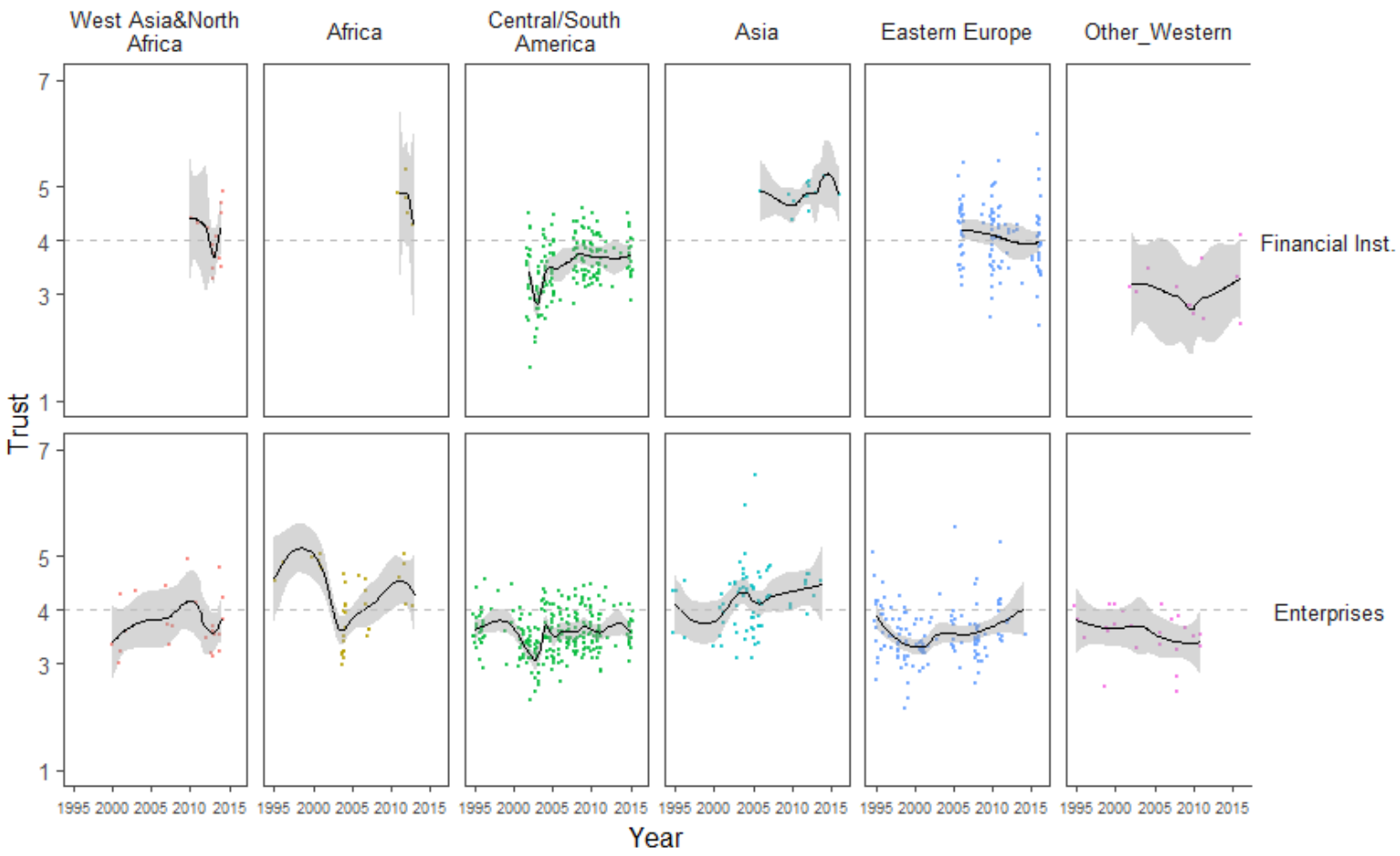
Trust in Public Administration



Trust in institutions of the civil society



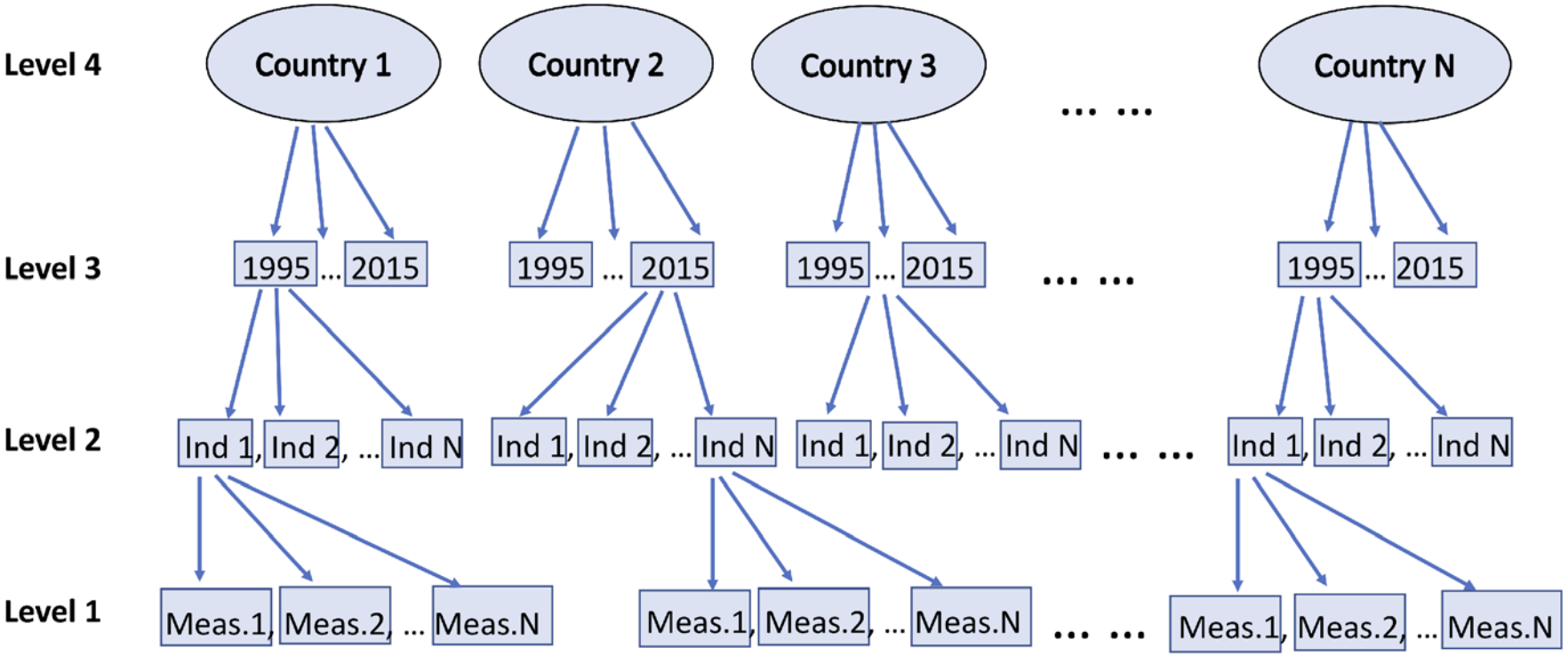
Trust in economic institutions



There is much data missing!
Trust in financial institutions low in Other Western countries (2008 crisis?).

Results: Multilevel analysis

A 4-level longitudinal model with repeated measures



Concretely: a glimpse at results

Trust in institutions - basic models

	Model 0	Model 1	Model 2	Model 3	M4b with scale	M4c with region
Intercept	3.956 ***	4.2209 ***	4.185 ***	4.180 ***	4.45221 ***	4.221827 ***
Level Measure						
Media (REF)						
State/President		-0.048 ***	-0.048 ***	-0.048 ***	-0.048 ***	-0.048 ***
Governments		-0.388 ***	-0.388 ***	-0.388 ***	-0.388 ***	-0.388 ***
Parliament		-0.655 ***	-0.655 ***	-0.655 ***	-0.655 ***	-0.655 ***
Elections- Elect. Commis.		-0.323 ***	-0.323 ***	-0.323 ***	-0.323	-0.323
Political Parties		-1.119 ***	-1.119 ***	-1.119 ***	-1.119 ***	-1.119 ***
International Org.		-0.173 ***	-0.173 ***	-0.173 ***	-0.173 ***	-0.173 ***
Army		0.314 ***	0.314 ***	0.314 ***	0.314 ***	0.314 ***
Police		-0.214 ***	-0.214 ***	-0.214 ***	-0.214 ***	-0.214 ***
Public Admin.		-0.202 ***	-0.202 ***	-0.202 ***	-0.202 ***	-0.202 ***
Judiciary		-0.380 ***	-0.380 ***	-0.380 ***	-0.380 ***	-0.380 ***
Church		0.619 ***	0.619 ***	0.619 ***	0.619 ***	0.619 ***
Trade Unions		-0.619 ***	-0.619 ***	-0.619 ***	-0.619 ***	-0.619 ***
ONG- Civil Society		-0.069 ***	-0.069 ***	-0.069 ***	-0.069 ***	-0.069 ***
Finance		-0.173 ***	-0.173 ***	-0.173 ***	-0.173 ***	-0.173 ***
Enterprises		-0.354 ***	-0.354 ***	-0.354 ***	-0.354 ***	-0.354 ***
Level Respondent						
woman			0.028 ***	0.028 ***	0.028 ***	0.028 ***
age_cat			0.011 ***	0.011 ***	0.011 ***	0.011 ***
Nb questions asked			-0.005	-0.005	-0.009 *	-0.008 *
Prop_Non-resp.			0.274 ***	0.274 ***	0.274 ***	0.274 ***
Level Country-Year						
Time				0.000 ns	-0.002 ns	-0.002 ns
Time2				0.000 ns	0.001 ns	0.001 ns
Level Country-Source						
Other Western (REF)						
West Asia N. Africa *Time*						-0.069
Africa						0.426 **
Central/South America						-0.366 **
Asia						0.557 ***
Eastern Europe						-0.058
LAPOP					0.192	0.484 ***
WVS-EVS					-0.267 ***	-0.075
medium scale (5-7)					-0.393 ***	-0.129
Long scale (10-11 pts)					-0.650 ***	-0.374 ***

- Distribution of variance - model 0:
- Institutions' level: 63%
- Respondents: 27.3%
- Time: 2.3%
- Country: 7.4%

At level 1: Trust in institutions

	Model 0		Model 1	
Intercept	3.956	***	4.221	***
Level Measure				
Media (REF)				
State/President			-0.048	***
Governments			-0.388	***
Parliament			-0.655	***
Elections- Elect. Commis.			-0.323	***
Political Parties			-1.119	***
International Org.			-0.173	***
Army			0.314	***
Police			-0.214	***
Public Admin.			-0.202	***
Judiciary			-0.380	***
Church			0.619	***
Trade Unions			-0.619	***
ONG- Civil Society			-0.069	***
Finance			-0.173	***
Enterprises			-0.354	***

- Highest trust: religious organizations, army.
- Lowest trust: Political parties, parliament & trade unions.
- Institutions explain
  - 7% of the level 1 variance.
  - And 4% of level 4 variance.

At level 2: respondents

Level Respondent				
woman			0.028	***
age_cat			0.011	***
Nb questions asked			-0.005	
Prop_Non-resp.			0.274	***

- Significant but very small effects.
  - Women and older people →higher trust
- Proportion of item non-response associated with higher trust: more non-response →higher trust.
- Practically no variance explained.

At level 3 & 4 Country-year-source & country-source

Level Country-Year						
Time	0.000	ns	-0.002	ns	-0.002	ns
Time2	0.000	ns	0.001	ns	0.001	ns
Level Country-Source						
Other Western (REF)						
West Asia N. Africa					-0.069	
Africa					0.426	**
Central/South America					-0.366	**
Asia					0.557	***
Eastern Europe					-0.058	
LAPOP			0.192		0.484	***
WVS-EVS			-0.267	***	-0.075	
medium scale (5-7)			-0.393	***	-0.129	
Long scale (10-11 pts)			-0.650	***	-0.374	***
Variance	**Model 3**		**Model 4b**		**Model 4c**	
Measures	2.257	61.3%	2.257	61.9%	2.257	63.1%
Respondents	1.056	28.7%	1.061	29.1%	1.061	29.7%
Country-Year	0.089	2.4%	0.088	2.4%	0.089	2.5%
Country-Source	0.279	7.6%	0.242	6.6%	0.172	4.8%
Total	***3.680***		***3.648***		***3.578***	

- Mean trust stable over time.
- Methods -- source, scale -- explain 15% of the variance at level 4.
- Region explains 25% more.
  - Lower trust: Latin America
  - Higher trust: Africa & Asia

Introducing cross-level interactions with time

Level Measure				
Army	0.314	***	0.271	***
			0.026	***
Church	0.619	***	0.643	***
- Time			*-0.030*	***
Level Country-Source				
Other Western (REF)				
West Asia N. Africa	-0.069		0.204	
Time			*-0.048*	***
Africa	0.426	**	0.425	**
Central/South America	-0.366	**	-0.363	**
Asia	0.557	***	0.533	***
Eastern Europe	-0.058		-0.066	
LAPOP	0.484	***	0.499	***
WVS-EVS	-0.075		-0.126	
medium scale (5-7)	-0.129		-0.126	
Long scale (10-11 pts)	-0.374	***	-0.332	***
Variance	**Model 4c**		**Model 4e**	
Measures	2.257	63.1%	2.253	63.1%
Respondents	1.061	29.7%	1.061	29.7%
Country-Year	0.089	2.5%	0.088	2.5%
Country-Source	0.172	4.8%	0.168	4.7%
Total	***3.578***		***3.570***	

- Trust in Army lower, sign., increasing over time.
- Trust in religious inst. higher, sign., decreasing over time.
- Trust in the WANA region higher, decreasing over time.
- 3% more variance explained at the levels of measurement, year & country.

Introducing complementary data: Polity2

	M4e ss polity		M4e + polity2	
Intercept	4.342634	***	4.357351	***
Polity2		***	*-0.014*	***
Level Measure				
Media (REF)				
State/President	-0.051	***	0.278	***
Polity2			*-0.051*	***
Governments	-0.401	***	-0.155	***
Polity2			-0.039	***
Parliament	-0.668	***	-0.356	***
Polity2			-0.048	***
Elections- Elect. Commis	-0.339	***	-0.394	***
Polity2			0.009	***
Political Parties	-1.133	***	-0.839	***
Polity2			-0.045	***

- Polity2, an index of democracy
  - Is **related negatively** to overall trust, and to trust in most political institutions, **except elections**.
- Introducing polity2 explains 14% of the variance in trust between countries.

What next? Literature review in order to...

- Figure out relevant between level interactions.
- Figure out which slopes should be random.
- Find which characteristics of institutions, individuals and countries may be relevant.
  - Impute missing values for country characteristics.
- What should we do about weighting?
  - At the individual level: not all files have equivalent weights, or even weights. Weighting not available in HLM 7.03 for 4-level models.
  - At the country-level? It would give a weight that is way too large to countries like Brasil in Latin America or China in Asia.

What next?

- The method is now well developped, systematized and described.
  - Document the process and the files
  - Make them available to the research community via Dataverse
  - Integrate with the Survey Data Recycling project.
- Integrate the rest of the World (essentially the WEIRD countries)?

Conclusion

- This method allows for concluding that trust is a property of institutions in a given environment more than an individual characteristic, in part because it allows for comparison between much varied countries.
- The distribution of variance between levels show how important it is to take into account the within individual-between measures variance.
- Another advantage is the possibility of cross-level interactions and random slopes.

At level 3: Country-year-source

Level Country-Source-Year						
Time			0.023	***	0.021	**
Time2			0.001	n.s.	0.001	n.s.
Prop_urban population			-0.015	***	-0.010	**
LnGDP			0.128	*	0.122	*
Polity2			-0.002	n.s.	0.004	n.s.
Gini evolution			0.001	n.s.	-0.002	n.s.
Level Country-Source						
Central/South America (REF)						
Asia					0.339	*
Africa					0.574	***
West Asia N. Africa					0.444	*
LAPOP					0.267	*
WVS					0.417	**
Variance						
Measures	2.537	63.6%	2.294	61.5%	2.294	62.7%
Respondents	1.046	26.2%	1.064	28.6%	1.064	29.1%
Country-Source-Year	0.099	2.5%	0.168	4.5%	0.167	4.6%
Country-Source	0.304	7.6%	0.202	5.4%	0.132	3.6%
Total	***3.986***		***3.728***		***3.657***	
Deviance	38255971		32450232	*5805739*	32450199	*33*
dl	5		34	*29*	*37*	*3*

- Trust increase with time.
- Hi Prop urban population= lower trust.
- Hi GDP= higher trust.

At level 4: Country-source

Level Country-Source-Year						
Time			0.023	***	0.021	**
Time2			0.001	n.s.	0.001	n.s.
Prop_urban population			-0.015	***	-0.010	**
LnGDP			0.128	*	0.122	*
Polity2			-0.002	n.s.	0.004	n.s.
Gini evolution			0.001	n.s.	-0.002	n.s.
Level Country-Source						
Central/South America (REF)						
Asia					0.339	*
Africa					0.574	***
West Asia N. Africa					0.444	*
LAPOP					0.267	*
WVS					0.417	**
Variance						
Measures	2.537	63.6%	2.294	61.5%	2.294	62.7%
Respondents	1.046	26.2%	1.064	28.6%	1.064	29.1%
Country-Source-Year	0.099	2.5%	0.168	4.5%	0.167	4.6%
Country-Source	0.304	7.6%	0.202	5.4%	0.132	3.6%
Total	***3.986***		***3.728***		***3.657***	
Deviance	38255971		32450232	*5805739*	32450199	*33*
dl	5		34	*29*	*37*	*3*

- Higher trust when source is LAPOP or WVS.
- Higher trust outside Latin America, even more in Sub Saharan Africa.