Quality & Quantity (2022) 56:1755–1797
https://doi.org/10.1007/s11135-020-01088-1

How to combine and analyze all the data from diverse sources: a multilevel analysis of institutional trust in the world

Claire Durand[1] · Luis Patricio Peña Ibarra[1] · Nadia Rezgui[1] · David Wutchiett[1]

Accepted: 17 December 2020 / Published online: 18 January 2021

Abstract
Scholars who want to perform cross-national comparative research rely on data provided by International survey projects, which study the same concepts in varying countries and periods using different question wordings and scales. In this article, we propose a process to combine and analyse the data pertaining to the same concept—institutional trust—when measures and sources differ. We show how we combined 1327 surveys conducted from 1995 to 2017 by 17 survey projects in 142 countries. The database comprises close to 2 M respondents and 21 M answers to trust questions. We use local regression to visualize the trends in trust for different institutions and sources of data in different parts of the world. We complete these analyses with a 4-level longitudinal analysis of repeated measures. These analyses lead to reliably conclude that institutional trust is a property of the institutions themselves and of the context in which they operate since there is much more variance within respondents than between respondents and more variance between countries than over time. This research contributes to the current debates in political trust research. Since the process presented here can be applied to other fields of research, the research also contributes to enhance the possibilities for comparative cross-national analysis.

Keywords Data combination · Institutional trust · 4-level multilevel analysis · Harmonization · Survey data recycling · Longitudinal analysis · Repeated measures · Comparative analysis · Cross-national analysis

Accumulation of knowledge is central to science in general and certainly to social science. However, scholars who want to perform cross-national comparative research face several issues. They rely on data provided by international survey projects like the Barometers, the Values Survey, the Social Surveys and other regional survey projects. The questions used to measure the same concepts vary in terms of question wording, answer scales used and specific object or focus. They also vary within survey projects as well as between projects, between countries and over time. In addition, each project does not cover all the countries

✉ Claire Durand
Claire.Durand@umontreal.ca

[1] Department of Sociology, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, QC H3C 3J7, Canada

over all the period of interest. Few projects aim at an international coverage; some are conducted yearly, others in different waves covering varying periods. Some regions—for example, the former soviet republics (Oleksiyenko 2017)—are under covered if we rely on only one international survey project. These observations are very similar to those put forward by Tomescu-Dubrow and Slomczynski (2016) to introduce the Survey Data Recycling (SDR) project.

It is difficult to list all the restrictions that researchers acknowledge when using data produced by the various international and regional survey projects. Authors tend to restrict themselves to one survey project and use only the question(s) asked in the same way in the countries where the data are available over all the period of interest (Catterberg and Moreno 2006; Schneider 2017, Tomescu-Dubrow and Slomczynski 2016). Researchers sometimes run analyses on several surveys but separately (Tomescu-Dubrow and Slomczynski 2016).

This restricts substantially the possibilities for cross-national and longitudinal comparison. The teams responsible for international survey projects devote much attention to harmonization ex ante. They try to make sure that data will be comparable between countries and over time for their specific projects. However, the context of the different countries is not always comparable. In the end, only few measures end up being similar for some of the countries of interest for a short period for a given project. In such a situation, how can we compare countries and regions of the world? How is it possible to trace trends reliably, assess whether different measures of the same construct are equivalent and whether different survey projects give a similar portrait of the situation in each country and period?

The main aim of the research presented here is to develop methods allowing for the combination and analysis of all the available data that include measures of similar concepts in order to expand the possibilities for comparative cross-national longitudinal research. We initiated this project independently from other data recycling projects, the SDR project for example (Tomescu-Dubrow and Slomczynski 2016), which focusses on protest behavior and political participation. We bring data recycling one step further in adding within-individual differences. To this aim, we combine a large number of dependent variables pertaining to the same concept asked from the same respondents.

We focus on institutional trust as the concept under study. Many researchers have shown that political and institutional factors have a substantial impact on trust in institutions (Hooghe and Zmerli 2011). Most studies limit themselves to an analysis of trust in new or established European democracies (Armingeon and Ceka 2014; Marien 2011; 2017; van der Meer and Dekker 2011; Van der Meer 2010a, b; van Erkel and van der Meer 2016; van der Meer and Ouattara 2019) and in the United States (Brewer et al. 2017; Cawvey et al. 2017; Dalton 2017; Hetherington and Rudolf 2017; Wilkes 2014). Most studies of trust restrict themselves to political trust, a central concept in the political science literature. If we are interested in how the political context influences institutional trust, variation in countries of Europe or of the Western World is restricted. In established democracies, economic conditions are rather good, and most countries use the same type of electoral system (Bormann and Golder 2013; Transparency International, 2017). In this research, we focus on the regions of the world that are less studied and where the context is more diverse, that is, the countries outside Western Europe and North America.

1 Why study institutional trust?

Trust in institutions is a prerequisite for democratic stability and for the proper functioning of representative democracies (see for example, Easton 1965; Marien 2011; Zavecz 2017). People must be able to trust that their government acts in their interest, that police will arrest and fine people when and only when they break the law, that the army will protect their country and will not overthrow an elected government, that banks will keep their money safe, that the Church and religious leaders will abide by the rules and principles that they promote.

The importance of trust as a measure of the smooth functioning of societies has fostered many international survey projects to ask questions about trust. The World Values Surveys and the Barometer Surveys together with, in the Americas, the Latin American Public Opinion Project (LAPOP), also known as the Americas Barometer, all ask numerous questions on trust. After the collapse of the Soviet Union, many survey projects were conducted in the former soviet republics and socialist countries of Eastern Europe and Central Asia, usually for a restricted period. Empirical research on trust is largely based on the data gathered by these projects. However, research using these databases suffer from some limitations.

First, *research tends to focus on only one region or country*. In the recently published books dedicated to political trust—Zmerli and Hooghe 2011; Zmerli and van der Meer 2017; Uslaner 2017—most analyses of institutional trust are restricted to countries of one region of the world, usually Europe or North America (Bargsted et al. 2017; Helliwell et al. 2017; Marien 2011, 2017; Mayne and Hakhverdian 2017; McLaren 2017; Mondak et al. 2017; Park 2017; Rose and Mishler 2011; Torcal 2017; Zavecz 2017; van der Meer and Dekker 2011) or to only one country (Bovens and Wille 2011; Brewer et al. 2017; Cawvey et al. 2017; Dalton 2017; Hetherington and Rudolf 2017; Thisted Dinesen and Mannemar Sonderskov 2017; Trudinger and Bollow 2011; Van Deth (2017); Wilson 2017), usually the United States. Few trust researchers compare different regions, or countries from more than one region (Catterberg and Moreno 2006; Chan et al. 2017; Hutchison and Johnson 2017; Letki 2017; Mattes and Moreno 2017; Muñoz 2017; van der Meer and Ouattara 2019; Uslaner 2017; Zmerli and Newton 2011, 2017).

Second, *research is limited in terms of the institutions that are studied*. A review of the literature available on Sociological Abstracts from 2015 to 2018 shows that more than 700 articles studied trust in political or social institutions. However, research focusses mostly on the institutions of the political system. Trust in institutions of the administration—police, courts, public administration—are often used as indicators of political trust (see Schneider 2017, van der Meer and Ouattara 2019). Trust in institutions of the civil society—the church, trade unions, the media, for example—and of the economic system—banks, enterprises—is rarely analyzed despite its substantial interest.

Third, *most comparative studies focus on only one institution or on a mean of trust in political institutions* (Schneider 2017; Van der Meer and Ouattara 2019). In doing so, they study trust as if it were an attribute of individuals who trust rather than of the objects that are trusted (van der Meer and Ouattara 2019; Mattes and Moreno 2017; Zmerli and Newton 2017). There is indeed a debate among trust researchers regarding whether political trust is unidimensional—an indicator of political culture according to Hooghe (2011) –, whether it may be measured in an equivalent way in different contexts (Schneider 2017; van der Meer and Ouattara 2019) and whether hierarchies of trust in different institutions vary between context (van der Meer and Ouattara 2019). Many researchers (see for example Bargsted

et al. 2017; Catterberg and Moreno 2006; Mattes and Moreno 2017; Zmerli and Newton 2017) conclude that trust is an evaluation of the performance of institutions since the level of trust in specific institutions differs within as well as between countries.

Fourth, *most researchers rely on only one database*. Consequently, research is dependent upon the availability of the data for a given country and year. This hampers the capacity to cover all the countries of interest (Oleksiyenko 2017; Tomescu-Dubrow and Slomczynski 2016). In addition, longitudinal analyses are not common (Van Erkel and van der Meer 2016) and researchers tend to restrict themselves to institutions measured in the same way over time (van Ham and Thomassen 2017).

Finally, *the statistical procedures used suffer from methodological limitations*. Wilkes (2014) underlines thoughtfully that only multilevel models are appropriate if we want to analyze concurrently individual and contextual determinants of institutional trust. Although resorting to this type of analysis is common in recent edited books on political trust—19 chapters on the 34 chapters using empirical data analysis in Zmerli and Hooghe (2011), Zmerli and van der Meer (2017) and Uslaner (2017)—it seems less common in articles. Very few articles published in the Journal of Trust Research, for example, use multilevel analyses.

Since multilevel analyses are often performed on only a few European countries, researchers must combine country and year in order to have enough units at the higher level. If it is possible to keep time on its own level in multilevel analysis, it becomes possible to test whether some events—a change in government, an uprise or an economic crisis—had an impact on trust in specific institutions and it is possible to partial out which part of the variation in trust belongs to the countries themselves and which part pertains to change over time (Van der Meer 2010a, b; van Erkel and van der Meer 2016). It also allows for an analysis of whether the impact of some variables—age for example—varies over time (Dalton 2005). The within-individual level must also be taken into account to estimate the variation in trust in different institutions that is due to varying evaluations by the same respondents.

Many authors have sought to understand the determinants of political trust. Economic determinants are probably the most studied (Mauk 2019; Bargsted et al. 2017). Authors conclude to "paradoxical" findings where strong effects are found at the individual level—personal situation and perceptions of the global economic situation—but inconsistent or inexistent effects at the macro level—macroeconomic indicators like unemployment rate, economic growth, inflation rate and budget deficit (Mishler and Rose 2001; Van der Meer 2010a, b; van Erkel and van der Meer 2016; van der Meer 2017; Zavecz 2017). Macro-level political determinants—the level of democracy, corruption, the government ideological orientation, the characteristics of representation—are also studied. It is interesting to note that the level of democracy is negatively associated with political trust (Letki 2017; Zavecz 2017; Mishler and Rose 2001).

Authors have also examined social—religious affiliation (Mauk, 2019; Bargsted et al. 2017; Van der Meer 2010a, b; van der Meer and Dekker 2011)—and political determinants at the individual level—political interest (Mauk 2019).

In summary, institutional trust has been studied extensively. It is beyond the scope of this paper to perform a complete review of the literature. However, this short review convince that institutional trust is a very appropriate and interesting field of research to apply the proposed methods. This field of research benefits from much interest in its measurement – and therefore measures are available – but methodological issues prevent researchers from using the full potential of —the available data.

Table 1 Basic Information on Data Sources

Sources	Number of countries	Number of surveys	Number of questions on trust		Data collection
			Minimum	Maximum	
Barometers					
Africa Barometer	34	100	3	19	2001–2015
Arab Barometer	14	37	4	11	2008–2016
Asia Barometer	29	50	10	19	2003–2007
Caucasus Barometer	3	17	13	15	2009–2015
East-Asia Barometer	14	49	8	15	2001–2016
Latino Barometro	19	356	6	17	1995–2017
New Baltic Barometer	3	9	8	13	1996–2004
New Europe Barometer	14	52	6	13	1995–2004
New Russia Barometer	1	10	3	9	1996–2009
South Asia Barometer	5	10	6	12	2005–2013
Value Surveys					
European Value Survey	27	45	13	18	1999–2009
World Value Survey	85	181	4	23	1995–2014
Other sources					
Consolidation of Democracy	13	13	17	17	1999
European Quality of Life	21	57	5	8	2007–2016
European Social Survey	16	80	6	7	2002–2014
Latin American Public Opinion Project	34	169	3	22	2004–2016
Life in transition	32	92	10	13	2006–2016
Total		1327			
Average	21.4		7.4	14.8	

2 Methodology

In this section, we first present the data and the process used to combine them. We then present the dependent and independent measures used and the harmonization process performed. Finally, the analytical procedures used to analyse these data are introduced.

2.1 Data

We identified 17 different survey projects that include questions pertaining to trust in institutions. For reasons explained above, we focussed on regions of the world outside Western Europe and North America. The selected survey projects were conducted in 142 countries overall, in South and Central America, North and Sub-Saharan Africa, Asia and Eastern Europe and Post-communist countries. Table 1 presents the information on these projects.

We selected 1995 as the starting date for this research, that is, the earliest date when the first Barometers started outside of Europe, in order to insure the best possible availability of data each year. Therefore, we dropped the few surveys conducted before 1995 by the selected projects. The latest available data sets when the current step in the research project was completed is 2017.

Table 2 Synthesis of the data

Surveys (source-country-year)	1327
Country-years	1082(245 double/triple)
Country-source combinations	364
Countries	142
Number of projects per country	Number of countries
1	45
2	36
3	37
4	8
5	3
6	4
7	7
8	2

Each project conducted between nine and 356 surveys for a total of 1327 surveys. The period covered by each project varies. The minimum number of questions on institutional trust in a given survey is three, the maximum, 23.

Some regions—and some countries within regions—are better covered than others. Table 2 presents detailed information on country coverage. Between one and 35 surveys were conducted in each country over the 22-year period; 245 surveys were conducted in the same country and year by at least two different projects.

Since the data come from different projects—or sources of data—we must be able to identify the source to assess its possible impact. We identified 364 country-source combinations in the 142 countries. Two-thirds of the countries benefitted from more than one survey project over the period. The number of projects conducted in each country ranges from one (45 countries) to four or more (24 countries). Eastern Europe and the former soviet republics benefitted from the most substantial coverage with nine different projects conducted in different countries over the period.

We combined the 1327 surveys in one database. The resulting data set includes 1,906,406 respondents for whom we kept the information on the demographics and the answers to the questions pertaining to institutional trust. This main file includes 133 trust variables, that is, one variable per institution for which trust was asked in any of the survey projects.

The advantage of combining multiple projects is two-fold. Different projects are not necessarily conducted in the same year in each country. Therefore, different projects may fill the holes in the time series and lead to data that are more complete. Second, two different survey projects conducted in the same country and period allow for cross validating the results from different projects.

In order to be able to analyse within respondent variance, we restructured the main data file to produce the level-1 file, that is, a file comprising one line per answer to a trust question per respondent. At this stage, we dropped the respondents who did not answer any of the trust questions (n = 77,188, 4%); this leaves 1,829,218 respondents. The restructuration procedure leads to a file with 21,209,889 lines, each comprising a respondent's answer to a trust question pertaining to one institution. In addition to the answer to the unique trust variable and an index that identifies the institution on which

trust was assessed, each line includes identifiers for each respondent together with identifiers for the country, year and source of data (see Durand et al. 2020 for replication data).

2.2 Measures

This section introduces the dependent and independent variables at each level of analysis. There are four levels: Measures, that is, answers to trust questions, are at level 1. This level includes one trust variable for each question answered by a respondent together with the variables identifying the institution for which the trust question was asked. Measures are nested within respondents (level 2). This level includes information on individual characteristics such as age, age category, sex, education level and occupation. Respondents are nested within surveys (level 3). Surveys are conducted by a survey project in a given country and year. Therefore, the survey level is a country-source-year level. It allows for studying variation over time. The variables related to time are introduced at this level. It is also at this level that it is possible to match time-varying information on each country, such as democracy indices and socio-economic and political characteristics. The surveys conducted by different projects are nested within countries. Therefore, level 4 is a country-source level. It allows for controlling unique effects related to each country and source of data that do not vary over time. It is at this level that it is possible to introduce fix characteristics of each country—like region and sub-region—and of the data sources—like the type of scale used to measure trust. We first present the dependent variable, institutional trust, which is necessarily at level 1 in multilevel analysis. We follow with a presentation of the independent variables at each level of analysis. At the measurement level, we need to harmonize the response scales used to measure trust and the institutions for which trust is asked. At the individual level, we need to select and harmonize respondents' demographics. At the survey level and at the country-source level, no harmonization is needed.

2.2.1 The dependent variable: Institutional trust

Trust is the dependent variable in the analysis. Since respondents answer multiple questions pertaining to trust in institutions, this level is akin to a repeated-measures level. Korsgaard et al. (2018) and Matsueda and Drakulich (2016) use a similar design and analysis to estimate within-respondent variations over time.

In order to harmonize the trust variables, we first identify the characteristics of the question wordings and answer scales used. Second, we deal with the question wording issue. Third, we select the appropriate procedure to transform the original scales into a common target scale. Finally, we justify which properties of the original scales must be controlled for and inform on how we will proceed.

The question wordings used by the different survey projects vary. Table 3 lists the English version of these questions and the characteristics of the response scales that are used. Surveys in English use two different words for trust: most surveys use the word *trust*; the European and World Values Surveys use the word *confidence*. Kolczynska and Slomczynski (2019) compared the European Social Surveys and the European Values Surveys for the same countries to assess whether there were differences in estimates with different question wordings. They identified four countries out of 17 where different terms were used for trust in the two projects, including Ireland, an English-speaking country. They found no difference in the distributions of answers for trust in parliament whatever the term used.

Table 3 Question wordings and scales used

Source	Question wording	Word used for trust	Scale length	Polarity	Direction	Low anchor	High anchor
Latinobarometer	…tell me how much trust you have in each of the following institutions. Would you say you have	confianza	**4**	**unipolar**	**descending**	no trust	a lot
Arabarometer	I will name a number of institutions, and I would like you to tell me to what extent you trust each of them:	trust	**4**	**unipolar**	**descending**	do not trust/none at all	a great deal/extent
Asiabarometer	Please indicate to what extent you trust the following institutions to operate in the best interests of society.	trust	**4**	**unipolar**	**descending**	don't trust at all	trust a lot
South Asia Barometer	I am going to name a number of institutions. For each one, could you tell me how much trust you have in them.	trust	**4**	**unipolar**	**descending**	none at all	a great deal
Consolidation of Democracy	I am interested in which of the following personalities you trust and to what degree you accept their statements.	trust	**4**	**unipolar**	**descending**	not at all	totally

Table 3 (continued)

Source	Question wording	Word used for trust	Scale length	Polarity	Direction	Low anchor	High anchor
World Values Survey	I am going to name a number of institutions. For each one, could you tell me how much confidence you have in them	confidence	**4**	**unipolar**	**descending**	none at all	a great deal
European Values Survey	…, how much confidence you have in them	confidence	**4**	**unipolar**	**descending**	none at all	a great deal
East Asia barometer	I'm going to name a number of institutions. For each one, please tell me how much trust do you have in them?	trust	**4**	**unipolar**	ascending (2001-2008) **descending (2010-2015)**	none at all	a great deal
Afrobarometer	How much do you trust each of the following,	trust	4	unipolar	ascending	not at all	a lot
New Baltic Barometer	To what extent do you trust the following public institutions?	trust	4	bipolar	descending	complete distrust	complete trust

Table 3 (continued)

Source	Question wording	Word used for trust	Scale length	Polarity	Direction	Low anchor	High anchor	
Caucasus Barometer	Please assess your level of trust toward each of social institutions and political unions	trust	**5**	**bipolar**	**ascending**	fully distrust	fully trust	
Life in Transition	To what extent do you trust the following institutions?	trust	**5**	**bipolar**	**ascending**	complete distrust	complete trust	
New Europe Barometer	To what extent do you trust each of the following institutions to look after your interests?	trust	**7**	**unipolar**	**ascending**	no trust at all	great trust	
New Russia Barometer	To what extent do you trust each of these political institutions to look after your interests?	trust	**7**	**unipolar**	**ascending**	no trust at all	great trust	
Latin American Public Opinion Project (LAPOP)	To what extent do you trust…?	trust	**7**	**unipolar**	**ascending**	nada (none)	mucha (a lot)	
European Quality of Life	Please tell me how much you personally trust each of the following institutions	trust	10	unipolar	ascending	do not trust at all	trust completely	2: some in 2005; quite a lot in 2013

Table 3 (continued)

Source	Question wording	Word used for trust	Scale length	Polarity	Direction	Low anchor	High anchor
European Social Survey	Please tell me on a score of 0-10 how much you personally trust each of the institutions	trust	11	unipolar	ascending	do not trust at all	complete trust

As Table 3 shows, there are many other differences between question wordings in English. However, it is not clear whether these differences are present in the many different languages used in the various projects. Since most languages have only one term for the concept of trust and most of the surveys used in this research are not conducted in English, we did not consider feasible nor necessary to control for the use of different terms in the question wordings.

The question of the response scales appears more complex. Simms et al. (2019) as well as Lee and Paek (2014) both show that the ideal number of options in response scales is between four and six and that under four options, reliability is highly problematic. We therefore decided to leave out the European Barometer, which uses only two response options.

As Table 3 shows, the survey projects analysed use five types of scale lengths, from Likert scales of four (ten projects) or five categories (two projects) to scales of seven (three projects), ten and eleven (one each) numbered categories. All the scales are unipolar, except the 5-point scales and the 4-point scale used by the New Baltic Barometer. Most of the 4-point scales are in the descending direction and all the scales with more than four options are in the ascending direction.

The response scales must be transformed into an identical target scale. Different types of transformation may be used (see De Jonge et al. 2017 for a review on this question). One proposed method is equating (Singh 2020), which is akin to reference distribution transformation (De Jonge et al. 2017; De Jonge et al. 2014). It applies best with few surveys and one obvious reference survey. It can hardly be used when there is much diversity in sources and contexts (De Jonge et al. 2017). Stretching—used by the SDR project (Slomczynski, Tomescu-Dubrow, Jenkins et al. 2016; Kolczynska and Slomczynski 2019; Kwak 2020) appears to be the most appropriate method.

The SDR project uses a stretching transformation where "each value in the source variable takes the mean value of the corresponding range of the 0–10 scale" (Kolczynska and Slomczynski 2019: 1023). This leads to a target scale where original scales do not have the same start and end values. For example, 4-point scales have target values of 1.25, 3.75, 6.25 and 8.75. The SDR project also proposes two other types of transformation, that is, a 0 to 1 transformation where the lowest and highest value are the same whatever the source scale and a more complex one related to the respondent's relative position in the distribution of trust in a given survey. The correlation between the original and target scales is very high: 0.88 at the survey level and more than 0.90 at the individual within survey level (Tomescu-Dubrow and Slomczynski, 2016).

Research on the impact of scale transformation has been performed in the context of the Subjective Well Being (SWB) literature. Both De Jonge et al. (2014) and Batz and Tay (2016) question whether the assumptions behind stretching are tenable, that is, that there are equal distances between the response options and that the labelling of options is not consequential. Using a transformation to a 1 to 10 scale in measures of happiness and life satisfaction and controlling for country and year using a multilevel model, Batz and Tay (2016) show that the impact of the source scale differs for the two measures. Besides, the impact is not linear, where longer scales would result in higher or lower levels of SWB. Finally, the regression coefficient between GDP and happiness is cut by half after controlling for the length of the original scale. In short, control for the original scale length is essential.

While Batz and Tay (2016) control only for the original scale length, Kolczynska and Slomczynski (2019) and Kwak (2020) use three indicators of the source scale, that is, length, polarity and direction, in their analysis of political trust variables. Kolczynska and

Slomczynski (2019) examine 1314 surveys of the SDR project while Kwak (2020) examines only the surveys conducted in the same country and year (137 country-years covering 53 countries between 1995 and 2012). They both show that the length and direction of the source scales have an impact. However, collinearity—long scales are usually unipolar and ascending—leads to substantial variation in estimates of coefficients and standard deviations when the three indicators are introduced together in the analysis (Kwak 2020).

In conclusion, the scale properties are only one aspect of the "cluster of methods" used by the different projects. These clusters of methods may also explain differences between projects, including between scales.

In order to harmonize the scales used to measure trust, we reversed the scales that were in the descending direction and we recoded the source scales to a 7-point scale. The length of the target scale has no impact on the substantive results since it is a linear transformation. However, the 7-point scale appeared preferable to the scales found in the literature for several reasons. It has an obvious middle point (4), it requires both stretching the shorter scales and squeezing the longer ones towards a middle ground, therefore "balancing" the transformations. It starts at 1, like all the original scales except the 11-point scale used by the ESS. Finally, the process is very easy to figure out conceptually and mathematically. The 4-point scales are recoded to 1, 3, 5, and 7, the 5-point scales to 1, 2.5, 4, 5.5 and 7, the 11 point scale is rescaled so that it starts at 1 and both the 10 and 11-point scales are shrunk proportionally to the 7-point scale. The resulting scale has a mean of 3.88 and a standard deviation of 1.94 at the level of measures (n = 21,209,889). Control for this operation is performed at the country-source level and will therefore be presented below.

2.2.2 Independent variables

This section presents the independent variables at each level of analysis and the harmonization process used, when relevant.

2.2.2.1 At level 1: institutions The institution for which trust is assessed is an independent variable. We aim at estimating whether individuals evaluate differently the various institutions for which they are asked their level of trust.

The large number of institutions for which trust is measured—133—is a challenge. We could have tried to figure out how to categorize and group the institutions before combining the data. Four issues prevented us from proceeding that way. First, we would have lost the original information on the specific institutions. Second, with each added survey, we could have run into a new institution, which may have forced us to modify the predetermined categories. Third, since two institutions may fit in the same broad category in a given survey, we would have had to introduce multiple variables for these categories or average the answers for some variables. Finally, since we were in the process of combining the data, we had no empirical criteria to help decide and validate which institutions could be grouped together.

We therefore introduced one variable per institution during the combination process. We define an institution as a stable social structure. We kept the questions about the representatives of these institutions, that is politicians or religious leaders, for example, but we excluded those that pertained to foreign governments (Chinese or U.S. government) or to specific institutions restricted to one country and period (like the Peace Commission in Columbia). Since we did not perform any a priori grouping, the original data are available (Durand et al. 2020).

Since we have 133 institutions but an average of only 12.5 institutions per survey, we need to group the institutions into larger categories. The proportion of missing data for each variable is way too high to perform an exploratory or confirmatory factor analysis (Tabachnick and Fidell 2019) or use an item response theory model like van der Meer and Ouattara (2019). Therefore, we must rely on "classical" methods, that is, perform a conceptual grouping and use empirical criteria to validate the proposed categories.

Table 6 in appendix 1 lists the institutions that were grouped into 16 broad categories pertaining to four domains. In order to validate the groupings, we checked that the means and standard deviations of trust in the institutions that are grouped together were similar in a given region and source and that there were enough respondents in a category to keep it separate. The four domains are *political trust* (six categories), *trust in the institutions related to public administration* (four categories), *trust in institutions of the civil society* (four categories) and *trust in financial institutions* (two categories). Appendix 1 informs on the proportion of each institution and the proportion of measures that fit in each category together with the proportion of respondents who were asked about at least one institution grouped in a broad category.

These groupings led to compute 15 dummy variables for the institutional categories minus one. *The media were selected as the reference category* to which the other institutional categories are compared because its level of trust was generally in the middle range and because the media are a rather "neutral" institution, that is, their main role is to report what happens, not to influence it or provide services.

2.2.2.2 At level 2: demographics and controls at the individual level The only main indicator of socio-demographics that does not need harmonization is sex. It was transformed into a dummy variable with Woman taking the value of one. Age is usually measured as a continuous variable (year of birth) but categories are used in some projects, and these may vary across countries or years. We managed to harmonize age in an ordinal variable comprising 14 categories—from 15 to 17 years old to 75 years and older—with only 0.4 percent of missing values.

Two indicators of the respondents' interaction with the survey instrument were introduced. The number of trust questions asked to the respondent is an indicator of burden. A high number may induce fatigue and satisficing from respondents. We also computed an indicator of *item non-response,* that is, the proportion of the questions on trust asked to the respondent that remained unanswered. A negative relationship between non-response and trust would mean that less trusting respondents are more reluctant to answer some questions.

Other demographic indicators—education, main occupation—would be relevant. They were harmonized but were not introduced in the analyses due to the high proportion of missing values. Level of education is missing in four percent of the surveys and main occupation, in 43 percent.

2.2.2.3 At level 3: trends at the survey—country-source-year—level The 22-year period over which the surveys are conducted is long enough to estimate the trends in institutional trust. To model the trends, year is centered at mid-point to avoid statistical collinearity, and subsequently squared and cubed (Singer and Willet 2003). These variables allow for estimating the global trends, but they may also be used in cross-level interactions to estimate the trends specific to some parts of the world or to specific institutions.

The quality of the available methodological information and of the methodology itself varies between surveys, projects, and over time (Kolczynska and Schoene 2019; Oleksiyenko et al. 2019). According to Kolczynska and Schoene (2019), the quality of methodological reporting has improved over time. On the contrary, Oleksiyenko et al. (2019) estimate that the presence of processing errors has increased. Kwak (2020) has shown that the SDR indices of Data Documentation and Processing Error are significantly related to trust in parliament and in the legal system.

Indicators of the quality of reporting and of the level of processing errors are not available for all the surveys used in this research, which prevents us from using this information for the time being. However, we combined the surveys one at a time and tried to identify and repair all the processing errors that we could identify. We also introduced dummies (see level 4 below) for some survey projects in order to control for their specific methodology.

At this level, it is also possible to match any time-varying characteristic of the countries. Numerous data bases compiled by different organizations—the World Bank, the United Nations, the Quality of Government Institute, Polity IV, etc.—make these indicators available. However, only a small proportion of these are available for most of the countries "outside the Western World" for all the years. If we introduced the Gini index and the GDP per capita for example, we would lose 174 surveys and 64 of the 364 country-source units. Therefore, to use these data, we would need to impute some values, which is beyond the scope of this paper.

2.2.2.4 At level 4: characteristics of countries and data sources The projects differ in terms of the question wordings and the scales used, their specific methodological features and the context in which the surveys are conducted. However, these characteristics are stable, that is, survey projects tend to harmonize ex ante and use the same methodological features for all their surveys. It is not possible to disentangle the possible impact of each of these specificities but we must control as much as possible for the different characteristics of the projects. It is the only way to make sure that observed differences between regions are not due to methodological artefacts. We keep three indicators that may explain some of the variance between country-source units: Scale length, source of data and country groupings.

Scale length is an indicator of the source scale. Since the relationship between scale length and the dependent variable may not be linear (Batz and Tay 2016; Kolczynska and Slomczynski 2019), we use dummies for the different scale lengths and compare them to the 7-point scale, the reference category, which is the only untransformed scale. Therefore, we control for scale length and for the transformation process at the same time.

In order to control for other possible methodological features of the survey projects themselves, we compute dummy variables for the World and European Values Surveys (WVS and EVS) and for the Latin American Public Opinion Project (LAPOP) that are compared with the other projects. This aims at controlling for "house effects", that is, unique features of the survey projects that cannot be controlled for due to lack of information. We selected these projects because the WVS and EVS are the only international projects and the only ones that use a different wording (confidence instead of trust). In addition, we observed that LAPOP estimates differ from those of other projects conducted in the same countries and periods.

A third variable aims at taking into account the regional historical, political, social or cultural context in which the different survey projects take place. Since most survey projects are conducted in specific parts of the world, differences between survey projects may

be due to characteristics of the projects themselves and not to contextual factors acting at the regional level.

How should we group countries together in order to take the contextual factors into account? While there are significant differences between geographical regions in trust and in various socio-economic indicators, there is heterogeneity within regions (Durand, et al. 2018). This led us to look for refined criteria to achieve more homogenous country groupings. We use two criteria, sequentially. The first one is historical and political. It pertains to the level of democracy and to the democratization process. The second one is geographical. It is applied after the first criterion has been applied. The process by which we decided on the country groupings is described in appendix 2. We end up with five regional groupings—Post-communist countries (28 countries), South and Central America (30), West Asia and North Africa (16), Sub-Saharan Africa (30) and Asia (22)—that are compared with Consolidated democracies (17 countries that were withdrawn from their respective geographical region).

Table 8 in appendix 2 shows that the "Consolidated Democracies" group stands out as substantially more democratic and more homogenous than the regional groupings according to all the indices. Therefore, by isolating this group, we removed some of the heterogeneity present in the regional groupings.

2.3 Analysis

Our goal is to assess the variation that can be attributed to each level of analysis and explain that variation. We fulfill this goal using two types of analysis, local regression, also called Loess regression, and 4-level longitudinal multilevel analysis of repeated measures.

When analysing longitudinal data, visualizing is essential to understand the trends. To do so, we produced time-series data by aggregating the level-1 data by institutional category, source of data, country and year (Durand et al. 2020). We analyze these data using local regression (Cleveland and Devlin 1988; Fox 2000a, b; Jacoby 2000; Loader 1999), a smoothing procedure that allows for a flexible estimation of trends without a priori assuming the form that they may take. Smoothing uses polynomials of degree 2. The bandwidth, estimated by trial and error as suggested by Fox (2000a), is fixed at 0.65 as suggested by Jacoby (2000). The Epanechnikov function, used to weight data according to the distance to the estimation points, is considered among the best performing functions (see for example Zucchini, 2003; Loader 1999).

While local regression illustrates trends, it does so without controlling for the possible differences due to the presence of different institutions, countries, and survey projects at each time point. To analyze concurrently trust in various institutions over time in different countries while taking these differences into account, multilevel analysis is the way to go (Hox 2010; Luke 2004; Raudenbush et al. 2016; Snijders and Boskers 2012; Tabachnik and Fidell 2019). It allows for taking into account the nesting of the data without restricting ourselves to the measures that are similar for all the countries, periods or constructs. The way multilevel analysis deals with missing data is a clear advantage, particularly in the current context (Hox 2010; Snijders and Boskers 2012).

Multilevel analysis offers a global method that keeps variance and its determinants at their conceptual level, therefore preventing ecological fallacy—also called the Robinson effect. Cross-level inference, that is, imputing at the lower level relations that occur at a higher level or vice versa, could generate misleading and erroneous conclusions (Snijders and Boskers 2012). Multilevel analysis allows for dealing with problems of

heteroskedasticity (Hox 2010; Snijders and Bosker 2012). It also allows for using to our advantage the fact that the same respondents answered multiple questions on trust. Finally, it facilitates the testing of cross-level interactions.

This use of multilevel analysis to perform meta-analysis of individual data, that is adding a measurement level, conforms to the theory. Hox (2010) underlines how interesting it would be to perform that type of analysis *if such data were available.* We did not find any publication where the authors took advantage of the nesting of measures within individuals except for Matsueda and Drakulich (2016) and Korsgaard et al. (2018). The software used is HLM version 7.03 (Raudenbush et al. 2016).

3 How has trust in institutions varied across the world since 1995?

3.1 Visualizing data

This section presents the trends in trust estimated by local regression for each country grouping. These trends are first presented by institutional category. They trace a portrait of the hierarchy of trust in various institutions and allow for assessing whether this hierarchy varies by country grouping. Trends are also presented by survey projects. In order to identify differences in average trust and in trends between projects conducted in the same context.

3.1.1 Trends by institutional category

Figure 1a–e show the trends in average trust by institutional category, by dimension, for each country grouping. The time series data used to compute these graphs are available in Dataverse (Durand et al. 2020). Each dot on the graphs corresponds to an average of trust in one institutional category and country at a given point in time for a specific source of data. Lines correspond to local regression estimates of trends.

Figure 1a illustrates the data pertaining to the three first institutional categories of the political trust dimension. The graph shows that the hierarchy of trust in the different institutions is similar across country groupings. However, trust varies in level and in trends by institution and country grouping. Trust in the President is generally higher than trust in the government or parliament, except in the rest of Asia. It has increased or stayed stable from 1995 to 2005 and is decreasing since then in most parts of the world. Trust in the government tends to follow similar trends, except in Post-communist countries where the trend goes in the opposite direction than trust in the president. Finally, trust in the parliament follows similar trends than trust in the government. Some trends are difficult to estimate due to sparse data (in the WANA region shortly after the Arab spring, for example). We notice a substantial downward trend in the WANA region and a cubic—fish-like—trend in South and Central America.

Figure 1b shows varying trends by country grouping for trust in elections, political parties, and supranational organizations. One clear trend is for political parties to be among the low trusted organizations, although they benefit from a small increase in Sub Saharan Africa and in the rest of Asia. Trust in elections is high and stable in consolidated democracies, in the rest of Asia and in Sub-Saharan Africa; it is lower and decreasing in South and Central America. Trust in elections is not asked in the Post-Communist countries and

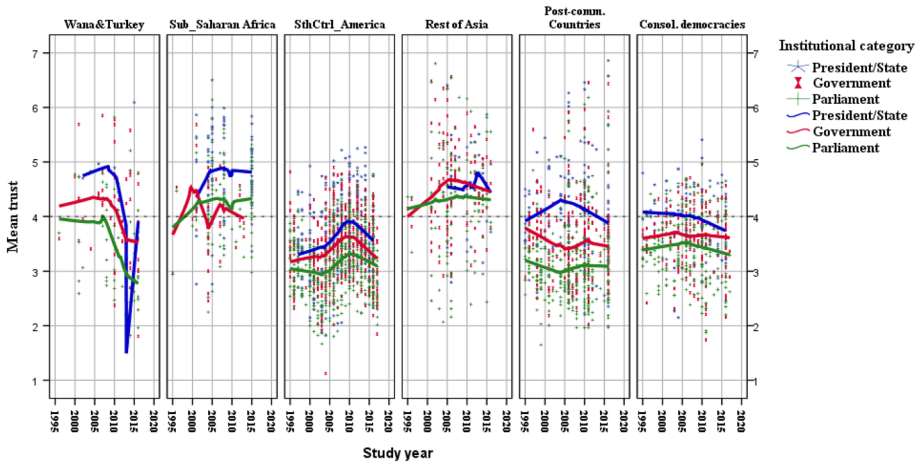

Fig. 1 **a** Trust in institutions of the political system—three first institutions. **b** Trust in institutions of the political system—three last institutions. **c** Trust in institutions of the administration. **d** Trust in institutions of the civil society. **e** Trust in institutions of the economic system

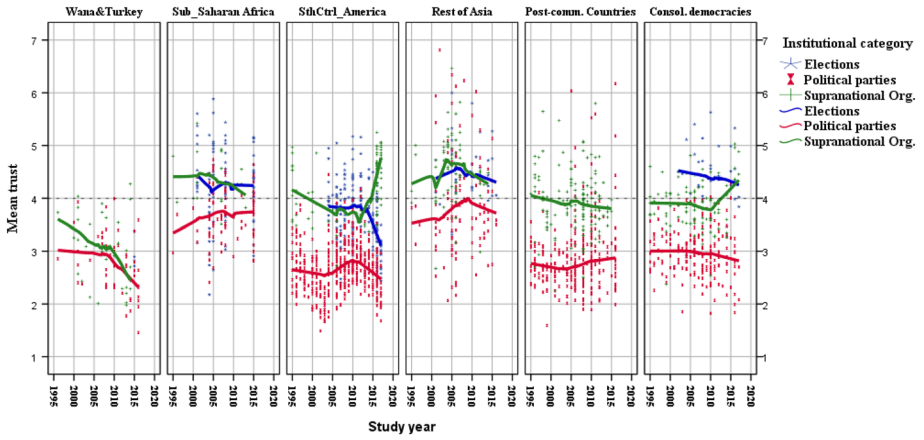

Fig. 1 (continued)

in the WANA region. Again, the hierarchy of institutions is similar in the different country groupings and trust is decreasing sharply after 2011 in the WANA region.

Figure 1c illustrates that, except in consolidated democracies, trust in the police is way lower than trust in the army. This finding illustrates why it is important to compare different regions. It would seem justified to group together trust in the army and in the police in consolidated democracies, but nowhere else. Trust in the justice system, including trust in the institutions in charge of fighting corruption, vary much between parts of the world. It is low in South and Central America and in the Post-communist countries, average in Consolidated democracies and high elsewhere (though decreasing in the WANA region). The

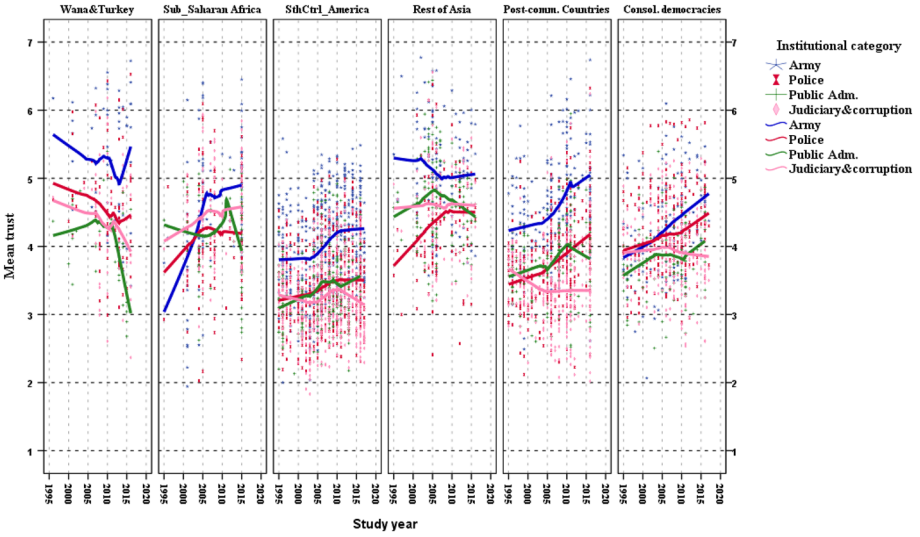

Fig. 1 (continued)

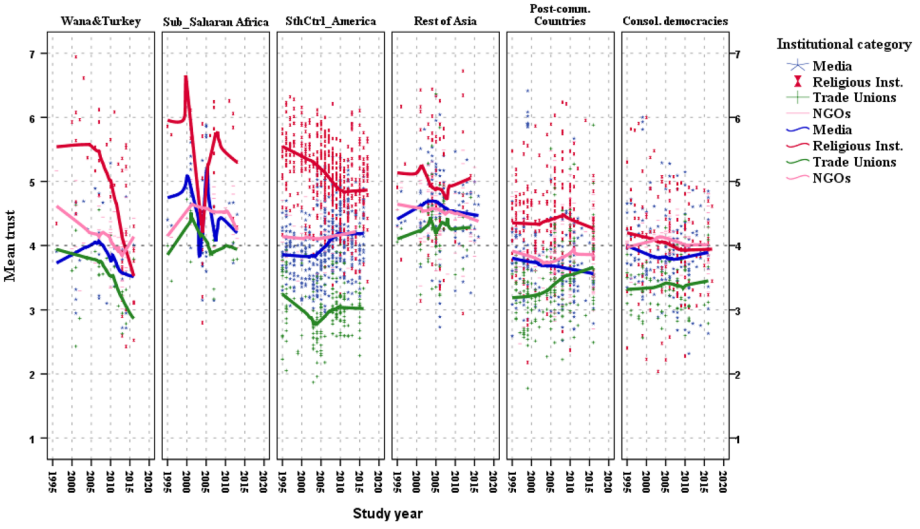

Fig. 1 (continued)

level of trust in the public administration varies much between parts of the world but it is increasing everywhere except in the WANA region.

If we summarize the trends in the institutions that are often grouped together in measures of political trust, we notice that trends in trust in these different institutions seem closer and more similar in consolidated democracies than in other countries. However, trust in political parties stands apart. Even in consolidated democracies, it does not seem

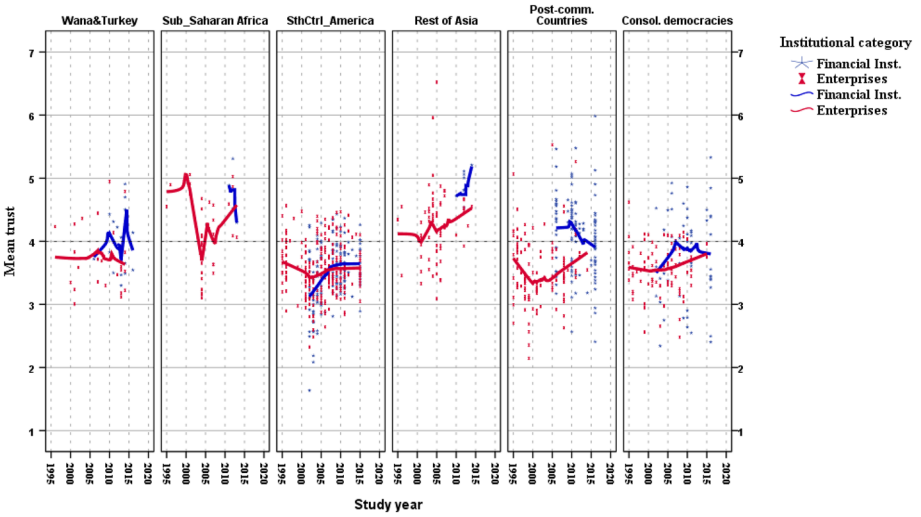

Fig. 1 (continued)

justified to group it with other political institutions. Trends in trust differ in level and in form between parts of the world and similarities observed in consolidated democracies are not always present in other parts of the world.

Figure 1d illustrates trends that are not often presented in the literature, that is, trends in trust in civic institutions. In Sub Saharan Africa and in the rest of Asia, trends are difficult to estimate due to lack of data. Levels and trends in trust are more similar to each other in consolidated democracies than in most other parts of the world. In all the regions, trust in religion is highest and in most, it is declining. It is particularly obvious in the WANA region. Trust in the media is rather stable and average in most regions. Trust in Non-Governmental Organizations (NGOs) is rather stable. Finally, trust in the trade unions is the lowest everywhere. It is however increasing in Post-communist countries and declining in the WANA region.

It is difficult to draw reliable conclusions on trust in economic institutions due to lack of data. They are however the only institutions for which trust is not declining in the WANA region. Besides, trends in trust in these institutions are rather stable and similar in consolidated democracies and in South and Central America.

In terms of country groupings, it is important to notice that trust in almost all institutions has been decreasing sharply since 2010 in the WANA region—and only in that region. Besides, trends in South and Central America often follow a fish-like pattern. These graphs also illustrate that the variance of estimates differ by institution and between some country groupings. For example, the dots representing average trust in religious institutions show high homogeneity, that is, low dispersion, in South and Central America. On the opposite, there is much more variation in the Post-communist countries and Consolidated democracies. Trust in political parties however is rather homogenous in all parts of the world. These graphs are representation of the data. They give cues regarding the validity and the reliability of the estimates. Statistical analyses will confirm or invalidate the "impressions" gathered from the graphs.

Fig. 2 **a** Trends in trust by source—WANA and Turkey. **b** Trends in trust by source—Sub-Saharan Africa. **c** Trends in trust by source—South and Central America. **d** Trends in trust by source—Rest of Asia. **e** Trends in trust by source—Post-Communist countries

3.1.2 Trends by source of data

The next graphs show the trends in average trust by source of data. They allow for visualizing whether different survey projects conducted in the same group of countries give a similar picture of trends in average institutional trust.

Figure 2a shows the trends for the WANA and Turkey region. Most surveys are conducted by two sources, the Arab Barometer and the World Values Survey (WVS). The trends for the other survey projects are not presented since these data points pertain to only one country (Turkey). The figure illustrates that there is convergence in the portrait traced by the Arab Barometer and WVS, particularly in recent years. The two sources show a downward trend after 2011.

Figure 2b, c, and d show that the trends traced by the WVS and the specific regional Barometers are similar. They also illustrate that the two projects complete each other, since they are often conducted in alternating years (as illustrated by the presence of different types of dots in different years). However, Fig. 2c shows that LAPOP, while tracing a trend that is similar to the two other sources, estimates trust at a level that is systematically higher. Since LAPOP also uses a different scale—a 7-point scale compared to the 4-point scale used by the other projects—multilevel analysis will tell us whether the observed

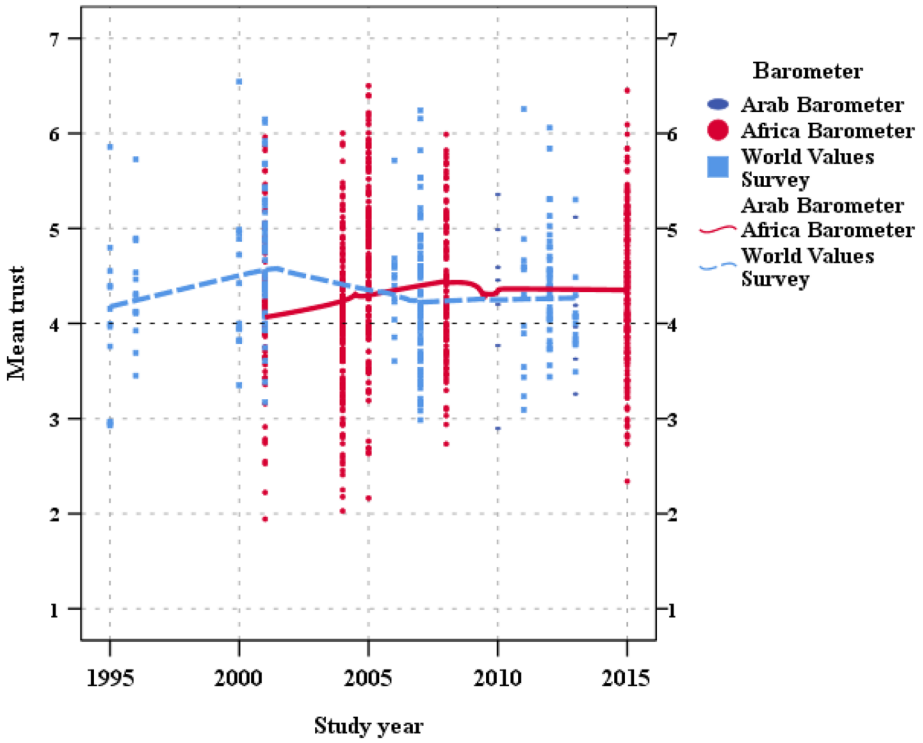

Fig. 2 (continued)

difference is uniquely due to the scale length used. The fish-like average trend in trust for this region is confirmed.

Figure 2e illustrates the trends traced by 11 different projects that conducted surveys in Post-communist countries. They cannot be easily compared with each other since some are conducted in specific countries—Russia—or regions—the Baltic or the Caucasus countries. In addition, the periods of fieldwork also differ. However, although there is much variation, the different sources trace similar trends, generally stable or increasing, when conducted in the same countries and period. Two trends stand as different. The Caucasus Barometer—conducted in Azerbaijan, Armenia and Georgia—shows a clear decreasing trend in trust between 2009 and 2015. On the opposite, the New Russia Barometer, shows a clear increasing trend in trust from 1996 to 2008.

Since consolidated democracies are spread all over the world, there are too many sources of data for these countries for a reliable comparison between sources, "everything else equal". Therefore, we do not present a graph for these countries. We can safely assume that the general absence of differences between sources—and the presence of differences for LAPOP—in the specific geographical regions apply as well to consolidated democracies.

In summary, the graphs comfort us in the assumption that the different sources complete each other and give a similar portrait of trust in institutions, except for LAPOP. This informs the following multilevel analyses.

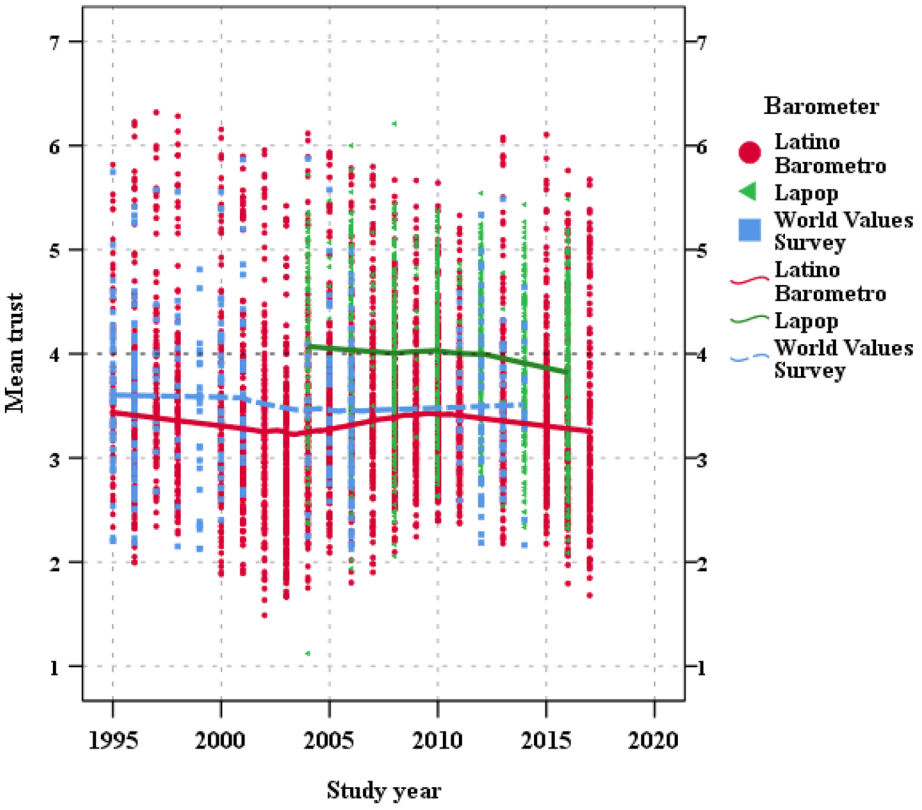

Fig. 2 (continued)

3.2 Multilevel analyses

Tables 4 and 5 present the results of the 4-level multilevel analysis in a classical way. We present the maximum-likelihood estimates for each coefficient and for the variance components together with the deviance. The analyses converged rapidly, and the reliability estimates are all higher than 0.8.

The first three models are presented in a separate table to ensure readability. Model 0 in Table 4 informs on the overall mean trust (3.96 on the 7-point scale) and on the distribution of the variance between levels of analysis. Close to two-thirds (63%) of the variance occurs between measures within respondents and 27.3% between respondents, which means that more than 90 percent of the variance is at the individual level. The rest of the variance is split between surveys, that is, at the country-source-time level—2.3 percent—and between countries and sources of data—7.4 percent.

The substantial variance between measures within individuals, reinforces the assertion that there is no unidimensional scale of trust in institutions (Rose and Mishler 2011; Schneider 2017; van der Meer and Ouattara 2019; Zmerli and Hooghe 2011). Trust does not appear to be a property of individuals who would not make much difference between institutions (Hooghe 2011) but on the contrary a property of institutions (Bargsted et al. 2017; Catterberg and Moreno 2006; Mattes and Moreno 2017; Zmerli and Newton 2017). In the

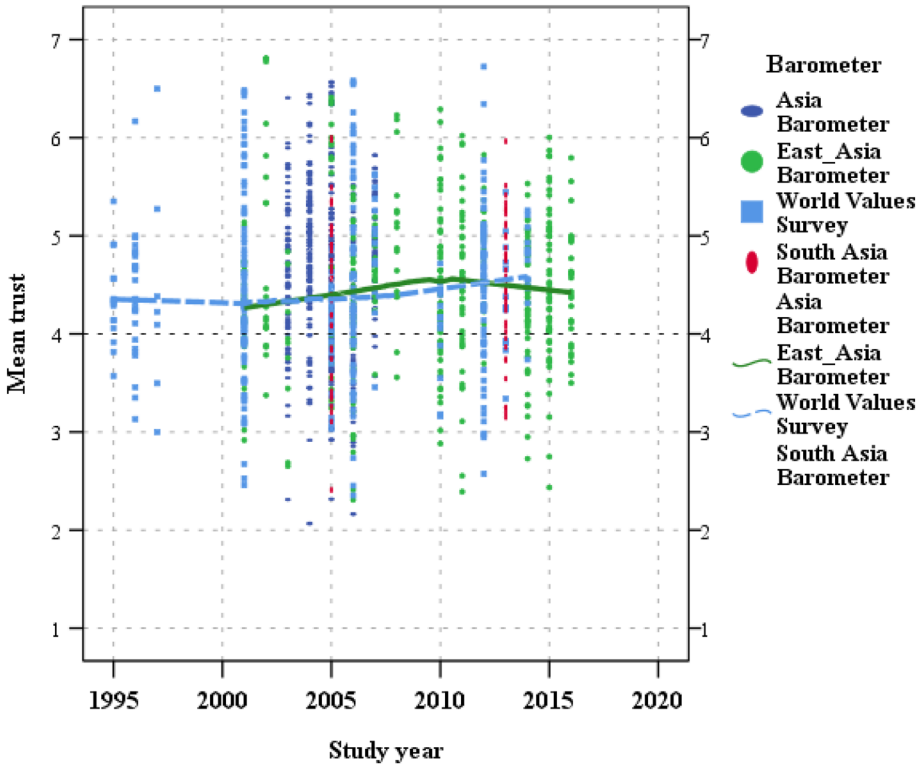

Fig. 2 (continued)

same way, at the higher levels, variation over time accounts for much less variance than variation between countries and sources, which validates the perception that institutional trust is contextual and rather stable.

This distribution of variance is consistent with other researchers. For example, van der Meer and Dekker (2011) estimate that 91 percent of the variance in trust in parliament in Europe in 2002–2003 is between respondents and 9 percent between countries. Using different measures of trust, Van der Meer (2010a, b)—15 percent between countries—as well as Hutchison and Johnson (2017)—13% at the country-year level in the Arab world (1999–2009) and 17% in Sub-Saharan Africa (2006–2011)—get similar results. In short, variance between individuals generally accounts for more than 80 percent of the variance in answers to trust questions. In this study, we show that respondent-level variance is split in two parts. The major part is due to the variance in the evaluations that individuals make when asked about different institutions. The rest is constituted by differences in trust between individuals.

Model 1 introduces the dummy variables for the different categories of institutions—compared with the media, the reference category—as predictors of variance in trust within respondents. After control for the institutional categories, the mean increases to 4.22, which is the mean trust in the media. The level of trust in the political parties is the lowest, at 1.12 point lower than trust in the media, followed by trust in the Parliament (−0.65) and the Trade Unions (−0.62). On the opposite, the Church (+0.62) and the Army (+0.31)

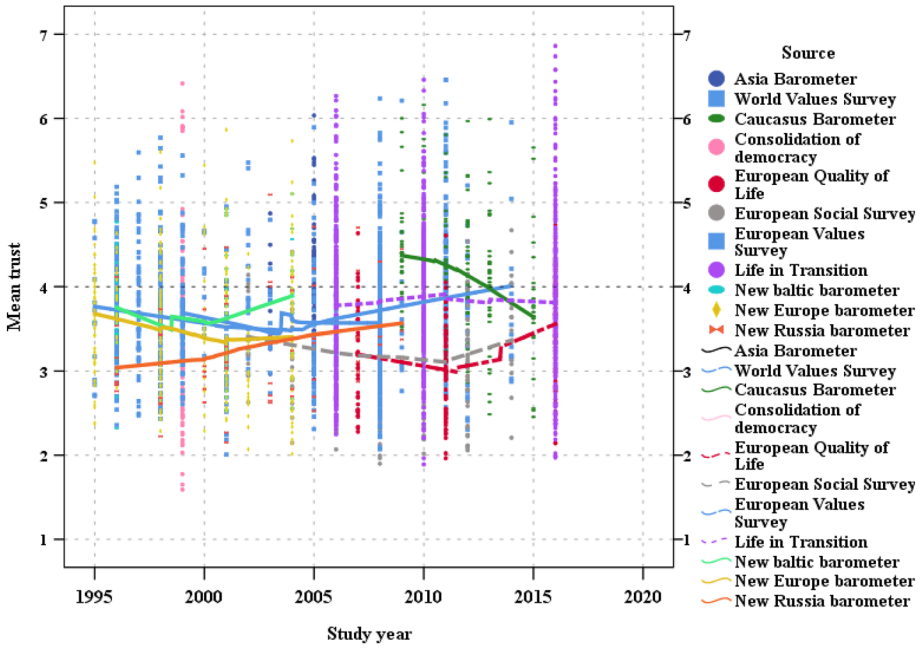

Fig. 2 (continued)

are the only institutions that are, on average, more trusted than the media. The institutional categories explain seven per cent of the variance at that level.[1] They also explain close to four percent of the variance at the country-source level, which means that part of the difference between country-source units is due to the fact that different institutions are assessed in different countries or by different sources. The difference in deviance, distributed as a Chi square distribution, is 105,716 with 15 degrees of freedom, a highly significant value.

Model 2 adds determinants at the individual level. On average, women tend to be slightly more trusting (0.028) than men. Trust increases by 0.011 by age category (k = 14), which means that there is a 0.14-point difference in average trust between the younger and the older respondents. Two indicators measure the respondent's burden and reluctance to answer some questions. The number of questions asked is not related to trust. However, the higher the item non-response, the more trusting the respondents are (+0.27). Our fear that less trusting respondents would tend to be reluctant to answer some questions is not confirmed. A possible interpretation is that respondents who are less trusting refrain from attributing low trust evaluations to some institutions. Hence, their level of trust for the questions that they answer is higher.

Given the sample size, small effects tend to be significant. Examination of the variance shows that virtually no respondent-level variance is explained by these variables, which tend to further confirm that institutional trust is more a matter of the institutions themselves, than of individuals, at least as age and sex are concerned. The difference in deviance is nonetheless highly significant at 3607 with four degrees of freedom. These results

[1] The proportion of explained variance is computed as (variance in model 1 minus variance in model 0) divided by the variance in model 0. In this case: (2.427-2.257)/2.427.

Table 4 Trust in Institutions - Measurement and respondent levels

	Model 0	Model 1	Model 2
Intercept	3.956***	4.221***	4.185***
Level Measurement			
Institutional Groupings (reference: Media)			
President/State		−0.048***	−0.048***
Government		−0.388***	−0.388***
Parliament		**−0.655*** **	**−0.655*** **
Elections		−0.323***	−0.323***
Political Parties		**−1.119*** **	**−1.119*** **
Supranational Organizations		−0.173***	−0.173***
Army		0.314***	0.314***
Police		−0.214***	−0.214***
Public Administration		−0.202***	−0.202***
Judiciary - Legal System		−0.380***	−0.380***
Church		0.619***	0.619***
Trade Unions		**−0.619*** **	**−0.619*** **
Non Governmental Organizations (NGO)		−0.069***	−0.069***
Financial Organizations		−0.173***	−0.173***
Enterprises		−0.354***	−0.354***
Level Respondent			
Woman			0.028***
Age (14 categories)			0.011***

Table 4 (continued)

	Model 0		Model 1		Model 2	
Number of questions asked					− 0.005	
Proportion of Non Response (Trust questions)					0.274***	
Variance	Model 0		Model 1		Model 2	
Measures	2.427	63.0%	2.257	61.3%	2.257	61.2%
Respondents	1.051	27.3%	1.064	28.9%	1.061	28.8%
Country-Year-Source	0.088	2.3%	0.090	2.4%	0.089	2.4%
Country-Source	0.284	7.4%	0.273	7.4%	0.278	7.6%
Total	***3.849***		***3.682***		***3.685***	
Deviance (diff. with preceding model)	80,555,452		80,449,736	*(105,716)*	80,446,129	*(3607)*
Degrees of freedom (diff. with preceding model)	5		20	*(15)*	24	*(4)*

The models include only the level 1 and level 2 predictors

blank: $p > 0.05$; * < 0.05; ** $p < 0.01$; *** $p < 0.001$

Table 5 Trust in Institutions - Country-source-year and Country-source levels

	Model 2	Model 3	Model 4a	Model 4b
Intercept	4.185***	4.180***	3.908***	3.913***
Level Country-Year-Source				
Time		0.010**	0.010*	0.000
$Time^2$		0.000	0.000	0.000
$Time^3$		0.000**	0.000**	0.000
Level Country-Source				
Source (reference: Other sources)				
LAPOP			0.587***	0.582***
WVS-EVS			-0.072	-0.088
Scale length (reference: 7-point scale)				
4-point scale			0.259*	0.265*
5-point scale			0.308*	0.288*
10-point scale			-0.150	-0.163
11-point scale			0.024	0.009
Country groupings (reference: Consolidated Dem.)				
West Asia - North Africa			0.036	0.298**
Time				*-0.021*
$Time^2$				*-0.004***
Sub-Saharan Africa			0.473**	0.453**
Central/South America			-0.431**	-0.362**

Table 5 (continued)

	Model 2		Model 3		Model 4a		Model 4b	
Time							*0.026***	
Time^2							*0.000*	
Time^3							*0.000****	
Rest of Asia					0.683**		0.665**	
Post-Communist Countries					-0.089		-0.086	
Variance	Model 2		Model 3		Model 4a		Model 4b	
Measures	2.257	61.2%	2.257	61.2%	2.257	63.3%	2.257	63.4%
Respondents	1.061	28.8%	1.061	28.8%	1.061	29.8%	1.061	29.8%
Country-Year-Source	0.089	2.4%	0.088	2.4%	0.088	2.5%	0.085	2.4%
Country-Source	0.278	7.6%	0.281	7.6%	0.159	4.5%	0.157	4.4%
Total	***3.685***		***3.687***		***3.565***		***3.560***	
Deviance (diff. with preceding model)	80,446,129	*(3607)*	80,446,119	*(10)*	80,445,951	*(168)*	80,445,909	*(210)*
Degrees of freedom (diff. with preceding model)	24	*(4)*	27	*(3)*	38	*(11)*	43	*(5)*

These models introduce the level 3 and 4 predictors (in addition to the level 1 and 2 predictors)

blank: $p > 0.05$; * < 0.05; ** $p < 0.01$; *** $p < 0.001$

are also in line with previous studies. For example, Van der Meer (2010a, b) explains four percent of the variance in trust in parliament at the individual level and van der Meer and Dekker (2011), five percent, although they introduced much more indicators of individual level characteristics—education, urbanization, household size, religion—in their model than we could do in this research.

Table 5 focusses on the level 3 and level 4 models. Model 3 adds time, time squared, and time cubed at level 3. Time and time cubed are significant. The linear effect of time is positive, which means that on average, trust has slightly increased during the period. The cubic effect of time is negative, which models a fish-like trend, a form that appears to be present at least in South and Central America. These variables explain only 1.31 per cent more variance compared with the preceding model. The difference in deviance (10 with 3 degrees of freedom) shows that the global contribution of time is negligible. However, as Fig. 1a–e illustrate, average trends may hide heterogeneity between country groupings and between institutions.

At the country-source level, we enter scale length using four dummy variables for each of the transformed scales (four, five, ten and 11 points), which are compared with the untransformed 7-point scales. We also enter two dummies for projects, one for LAPOP and one for WVS-EVS, for reasons stated above. Finally, we enter the country groupings using consolidated democracies as the reference category. We first entered these variables separately to understand their independent contribution. We present the final model with all the variables.

Although the different sources of data use different scales and different methodologies and are present in different regions of the world, methodological factors by themselves do not explain much variance. This confirms the portrait traced by Fig. 2a–e, that is the absence of substantial differences between sources. The variance at level four decreases from 0.281 to 0.254, that is, 9.6 percent, when we enter scale length. It further decreases to 0.239 when we enter the dummy variables for the survey projects (5.3% more variance explained). Therefore, we can be reassured that methodological differences between survey projects do not have a major impact on the estimates. However, this impact is not negligible and it can and must be controlled for.

When we enter the country groupings, the variance decreases to 0.159. Country groupings explain by themselves 28.5 percent more variance. The three factors together explain 43.4 percent of the variance at the country-source level. The difference in deviance is highly significant at 168 with 11 degrees of freedom. These results may be compared to those presented by Van der Meer (2010a, b) and van der Meer and Dekker (2011) for trust in parliament in Europe; they explain respectively 65 percent and 86 percent of the variance at the country-time level using indicators of the political context (corruption, proportional electoral system, former communist country).

The model shows that, compared with 7-point scales, 4-point and 5-point scales lead to slightly higher estimates of trust while 10-point and 11-point scales show no difference. LAPOP's estimates are more than half a point (0.59) higher, on average, than those of other projects, once controlled for scale length. Finally, compared with consolidated democracies, trust in the Rest of Asia (+0.68) and in Sub-Saharan Africa (+0.47) is significantly higher, trust in South and Central America is significantly lower (−0.31) and trust in the Post-communist countries and the countries of the WANA region is not different.

Guided by the portrait traced by local regressions, we hypothesized that the trends in average trust were different in the WANA region—a substantial drop after 2011—and in South and Central America—a fish-like cubic trend—than in consolidated democracies. To estimate these trends, we add time and time^2 for the WANA region and time, time^2

and time[3] for South and Central America in model 4b. These cross-level interactions show a more refined portrait. WANA appears to have benefitted from more trust than the consolidated democracies at the beginning of the period (+0.298) but experienced a quadratic negative trend in trust afterwards. Average trust in South and Central America still appears lower than in consolidated democracies. However, both the linear and cubic trends are significant; they portray a fish-like change over time (down-up-down). These additional trends explain 1.3 percent of the variance at level 4 and an additional 3.4 percent of the variance at level 3. The change in deviance is highly significant (210 with four degrees of freedom). In addition, the time variables for the overall model become non-significant, which means that the trends for these country groupings take care of most of the global differences over time.

4 Discussion

The process presented here informs on major issues raised by researchers who study institutional trust. In order to analyse trends in trust comparatively and longitudinally, we combined all the data pertaining to trust in institutions without performing any a priori grouping or selection of institutions. At this stage, we only harmonized the scales used and the demographics. Therefore, all the data are available (Durand et al. 2020), and the decisions taken relative to groupings of institutions and countries may be modified by other researchers. Analyses may also be performed on a subset of institutions or countries. The same process followed for institutional trust may be applied to other concepts of interest, which would contribute to extend the possibilities of cross-national comparative analysis.

Many analyses may be performed to check for the robustness of our results. We did perform quite a few—with different country groupings and different ways of controlling for methodological factors. We concluded that our results are highly consistent and robust. The fact that we have a high number of cases, including at the highest level, also helps avoid an impact of influential cases on our results (van der Meer, Te Grotenhuis and Pelzer 2006).

One issue is the way we grouped the countries together using both historical-political criteria and geographical ones. There are interpretable differences between the country groupings that we used. However, we continue looking for factors that may explain these differences and eventually validate our groupings. Different researchers have explored economic, social and political determinants of political trust—inequality, GDP, perception of corruption, level of democracy, proportion of urban population, ethnic diversity, etc.—in specific regions, with mixed or paradoxical results (Letki 2017; Mishler and Rose 2001; Zavecz 2017). Van Erkel and van der Meer (2016) and Durand et al. (2018), also obtained paradoxical results when introducing time-varying indicators in an attempt at explaining variation between and within countries.

The fact that we could combine data over a long enough period to have a separate level for time open opportunities. Cross-level interactions would allow for testing whether the trends in trust differ by institutional category. They would also allow for assessing whether the strength of some relationships—between gender, age or education and trust for example—decreases or increases over time.

The data file will be regularly updated to include the most recent data and the data for the countries of the "Western World". The version used in this paper, is available via the Dataverse link "Institutional Trust in the World" (Durand et al. 2020). New analyses planned with these data include analyses conducted in specific regions, analyses exploring

the influence of education on trust, analyses exploring the characteristics of very trusting or very untrusting respondents, and analyses of the social, political and economic determinants of trust at the country level.

5 Conclusion

Since there is much more variance between institutions than between individuals and more variance between countries than over time*, we may conclude that institutional trust informs on the evaluation of specific institutions in a given context.*

Since we compare countries that experienced different political regimes, including authoritarian ones, it is interesting to observe that the consolidated democracies included in this research do not show a higher level of institutional trust, on average, than Post-communist societies or Sub-Saharan and Asian countries. It leads to question the hypothesized strong relationship between democracy and overall trust in institutions.

Our results are in line with previous research on institutional trust. However, the process presented allows to go a step further in the research. We analyse trust in all institutions and in all countries at once taking into account within-individual variability, which is a first.

The database allows for much more research on this topic, including on the impact of the social, political, and economic context in which institutions play their role. The method illustrated here can be applied to other fields of research, like attitudes related to topics where many surveys are conducted, with many questions asked of the same respondents. The process is rather easy to carry on once the main questions related to harmonization are dealt with. The main point to keep in mind is that the combination process should be performed in such a way that the original data remain available. The decisions related to categorization can be taken after the combination process is completed, and modified afterwards, if needed.

Authors' contributions The order of authors reflects their contribution.

Funding This research is funded by the Social Sciences and Humanities Research Council of Canada (SSHRC) Grants No. 435-2019-0899 and 430-2015-01208.

Availability of data and material Replication Data: *Institutional Trust in the World*, Université de Montréal Dataverse, V1, https://doi.org/10.5683/SP2/TGJV6G.

Code availability The software used in HLM 7. It is menu driven.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interests.

Table 6 Institutional groupings

Dimension	Number of answers	Proportion of measures	Proportion of respondents	Original institutions
Political Institutions				
State-President	899,573	4.2%	46.4%	President (66.9%); Executive (26.5%); State (6.6%)
Government	1,949,447	9.2%	72.8%	Government (52%); Local government (34.5%); Regional government body (6.4%); *Less than 5%*: Prime minister; Cabinet of ministers; National Government; Local government body
Parliament	1,700,163	8.0%	92.9%	Congress_Parliament (83.9%); National Legislature (16.1%)
Elections	723,795	3.4%	27.9%	National electoral commission (65%); Elections (21.1%); Local government elections (13.2%); *Less than 5%*: Official Results of the Next Election; Secret vote
Political Parties	1,829,600	8.6%	84.2%	Political parties (76.5%); Politicians (8.6%); Political parties ruling (7.3%); Opposition political parties (6.5%); *Less than 5%*: Largest opposition party; Nationalist parties; Islamist parties; Arab nationalist parties; Socialist leftist parties

Table 6 (continued)

Dimension	Number of answers	Proportion of measures	Proportion of respondents	Original institutions
Supra national Organizations including regional organizations (0.8%) and international economic organizations (0.9%)	1,279,942	6.0%	30.8%	United Nations (42.2%); European Union (23.2%); NATO (5.8); *Less than 5%*: World Bank; International Monetary Fund; Organization of American States; World Trade Organization; Organization for African Unity (OAU); Arab League; Commonwealth of Independent States (CIS); Inter-American Development Bank; NAFTA; United Nations Development Programme; Mercosur; Development Bank of Latin America; Central American Bank for Economic Integration; APEC; Association of South East Asian Nations (ASEAN); South Asian Association for Regional Cooperation (SAARC); Free Commerce Treaty; United American States Organization; Southern African Development Community (SADC); Andean Pact; Caribbean community (CARICOM); Arab Maghreb Union; Commission for Human Rights; Cooperation Council for the Arab states of Gulf (GCC); Economic Cooperation Organization (ECO); East African Cooperation; UNASUR
Administrative institutions				
Army	1,372,308	6.5%	75.0%	Armed forces (100%)
Police	1,783,116	8.4%	96.9%	Police (99.3%); *Less than 5%*: Federal Service of the Russian Federation; Police station; Policía Militar de Orden Público; Border service; Police to Catch Robbery or Assault Criminal

Table 6 (continued)

Dimension	Number of answers	Proportion of measures	Proportion of respondents	Original institutions
Public Administration	1,292,717	6.1%	48.5%	Civil service (30.2%); Public administration (13.3%); Public education system (10.6%); Ombudsman office (8.5%); Public health system (8.1%); Tax department (6.2%); Universities (6.0%); Public ministry (5.8%); *Less than 5%*: Social security system; Local authorities; State enterprise; State student's admissions committee; Civil servants; Instituto de Acceso a la Información Pública; Family office; Trust in Office of Utilities Regulations (OUR)
Judiciary and corruption control	2,054,842	9.7%	90.3%	Judiciary (72.1%); Supreme Court of Justice (9%); Justice System (6.6%); *Less than 5%*: Attorney office; Superior Court of Accounts; Corruption Commission; Constitution Court; Judiciary will Punish the Guilty; International Commission against Impunity in Guatemala (CICIG); Prosecutor to Investigate Corruption; Judicial System to Punish Corruption; Misión de Apoyo contra la Corrupción y la Impunidad en Honduras; Tribunal de Ética Gubernamental; Independent Commission of Investigations (INDECOM); Director of Public Prosecution; Major organised crime and anticorruption agency (MOCA); Office of the Contractor General (OCG)
Institutions of the civil society				
Media	2,164,447	10.2%	68.2%	Television (32%); Medias (18.3%); Press (15.9%); Newspapers (14.8%); Radio (9.5%); *Less than 5%*: Government broadcasting service; Independent broadcasting services; Independent Newspapers; TV&Radio; Government newspapers; National broadcasting; TV News

Table 6 (continued)

Dimension	Number of answers	Proportion of measures	Proportion of respondents	Original institutions
Religion	1,342,865	6.3%	66.4%	Church (70.8%); Catholic Church (16.2%); Evangelical/Protestant Church (8%); *Less than 5%*: Religious Leaders; Traditional leaders; Muslim Brotherhood; Organization of the Islamic World (OOTIW)
Trade Unions	684,730	3.2%	37.4%	Trade unions (100%)
NGO	913,678	4.3%	29.9%	Environmental Protection movement (29.9%); Women´s movement (25.4%); ONG (21.9%); Charitable or humanitarian organizations (16.9%); *Less than 5%:* Civil society institutions (associations, clubs, volunteer youth groups, etc.); Indigenous Movements; Farmers organizations; Movimiento en pro de Vieques (Puerto Rico)
Economic Institutions				
Financial Institutions	456,628	2.2%	23.2%	Banks (92.7%); Stock exchange (7.3%)
Enterprise	762,038	3.6%	37.3%	Private enterprises (44.3%); Major companies (38%); Multinational companies operating in your country (5.4%); Large domestic companies (5.3%); *Less than 5%*: Local market; Small businesses/shopkeepers; Employers
Total	**21,209,889**	**21,209,889**	**1,829,218**	

Appendix 1: The grouping of institutions

Table 6 presents the groupings of the 133 different institutions in 16 broad categories referring to four dimensions. *Political trust* comprises six categories, that is, president/state, government, parliament, political parties, elections, and supra-national organizations. Between 28% (elections) and 93% (parliament) of the respondents were asked about their trust in an institution grouped in one of these institutional categories. There are one or two major institutions that account for the larger part of the data in each category. For example, 66% of the respondents in the President/State category were asked specifically about the president and 84% were asked about the Government or Congress in the Government category. Therefore, it is unlikely that the inclusion of any institution that has been less surveyed had an impact on the results. The most heterogenous category groups all the international and regional institutions into a "supra-national institutions" category.

The second group comprises the *institutions related to the public administration.* Some of the institutions in this grouping are often grouped with the political institutions in a unidimensional scale. When considered separately, they receive various designations: institutions of implementation (Marien, 2011, 2017), of the state (Mattes and Moreno 2017), regulatory (Catterberg and Moreno 2006; Zavecz 2017) or impartial (van der Meer and Ouattara 2019). This dimension includes four categories. Trust in the police (97% of the respondents), in the judicial system, including institutions that fight corruption (90%) and in the army (75% of respondents) are frequently measured. Public administration itself (48% of the respondents) is a quite heterogenous category. It groups all the institutions of the public service including those of the education and health systems.

The third group comprises the institutions of the civil society (Catterberg and Moreno 2006), also called civic institutions (Zavecz 2017). It includes trust in the media (asked of 68% of the respondents), in the Church or religious leaders (66%), in the trade unions (37%) and in the non-governmental organizations (NGO), 30%. Trust in the Media accounts for 10.2% of the measures and therefore is the most frequent institutional category.

Finally, the fourth group comprises the *economic institutions*, that is, the financial institutions and the enterprises. At most 37% of the respondents were asked about these institutions.

Some institutional categories comprise few different institutions—the army, for example—while others group many different institutions—supra national organizations, for example. However, our empirical and statistical criteria hold and, although the institutions themselves may be different, they may not be assessed differently by respondents (Bovens and Wille, 2011). For some of these institutions—politicians and political parties for example—research shows high correlations when trust about them are asked from the same respondents (Hooghe, 2011; Marien, 2011; Torcal, 2017). Finally, given the way we set up the data base, it remains possible to drop some of these institutions to check whether the choices that we made biased the results.

Other groupings may be criticized. Why group trust in the Prime Minister with trust in the government and not trust in the President? Different political systems have different roles for the President and the Prime Minister. In some systems, the President represents the state and has a protocolary role while the Prime minister leads the government. Empirically, the average trust in the President is usually higher than trust in the government or the Prime Minister, particularly in countries where the two roles coexist. This led us to group trust in the Prime Minister with trust in the government.

Appendix 2: The grouping of countries

We must first deal with the fact that some regional survey projects include countries of the "Western World". The Latino Barometro includes Spain; the Americas Barometer (LAPOP) includes Canada and the United States. The European Quality of Life project, the European Social Survey and the European Values Surveys are conducted in Austria, Finland, Germany and Greece, countries that are geographically in the eastern part of Europe. Given our stated goal to concentrate on the countries outside the Western World, we could have dropped these countries. However, if we keep them, we benefit from more variance in the contexts and therefore more possibilities for relevant comparisons with the other countries.

What do these "Western World" countries have in common? All of them received the highest grade (10) on the Polity IV index (Marshall, Gurr and Jaggers, 2019; Center for Systemic Peace 2019) over all the period studied. The Polity IV index has two scales, one of democracy and one of autocracy and an index computed as the difference between the two. Therefore, it varies from minus 10—fully autocratic regime—to plus 10—fully democratic regime. We selected this index because it is more fact-based (Marshall, Gurr and Jaggers, 2019) than other indices that are commonly used. Though other indices produced by Freedom House (Freedom House, 2019), V_DEM (Coppedge et al. 2018) and the Economist Intelligence Unit (Economist Intelligence Unit 2019) are commonly used and have their own specificities, all the indices are highly correlated (Elff and Ziaja, 2018).

Since we have a criterion that groups these "Western World" consolidated countries together, we must check whether other countries in our database meet the same criterion, that is, a Polity IV score of 10 over all the period. Ten countries in four regions meet the criterion: Japan and Mongolia in Asia, Costa Rica, Trinidad and Tobago and Uruguay in South and Central America, Cabo Verde and Mauritius in Sub-Saharan Africa and Hungary, Lithuania and Slovenia in Eastern Europe. We group these countries with the "Western World" countries in a "Consolidated Democracies" category. These countries are not a random selection of the consolidated democracies. The next iteration of data combination will seek to include all the countries of the World.

It is also on historical and political criteria that we group together the "Post-communist countries" that did not qualify as consolidated democracies. Indeed, authors who study trust in European countries generally group together the former socialist countries of Eastern Europe and the former soviet republics (Quaranta and Martini, 2016; Zavecz 2017). Some authors divide these countries in two or more sub-regions (Catterberg and Moreno 2006; Zavecz 2017). While Catterberg and Moreno (2006) found differences between the Eastern European countries and the former soviet republics, Zavecz (2017) did not find any. We decided to group all these countries together (except the three countries that are considered consolidated democracies) to have enough units in the category.

The countries that did not meet the former two criteria are grouped according to a socio-geographical criterion. We use the same divisions as Zmerli and van der Meer (2017). We group the other countries of South and Central America in one group, of West Asia and North Africa (WANA)[2] including Turkey in a second group, of Sub-Saharan Africa in a third group and of Asia in a fourth group. Consolidated democracies include 17 countries, Post-communist countries, 28, South and Central America, 30, Rest of Asia, 22, Sub-Saharan Africa, 30 and the WANA and Turkey grouping, 16. Table 7 in lists all the countries by grouping.

[2] This region in often called the Middle East and North Africa (MENA), a designation that is not geographically based and that is criticized for its occidental bias..

Table 7 Country Groupings

Region	Countries
Consolidated democracies	Austria; Canada; Cabo Verde; Costa Rica; Finland; Germany; Greece; Hungary; Japan; Lithuania; Mongolia; Slovenia; Spain; Trinidad and Tobago; United States; Uruguay
Post-communist Countries	Albania; Armenia; Azerbaijan; Belarus; Bosnia; Bosnia&Herzegovinia;; Bulgaria; Croatia; Czech Republic; Estonia; Georgia; Kazakhstan; Kosovo; Kyrgyzstan; Latvia; Macedonia; Moldova; Montenegro; Poland; Romania; Russia; Serbia; Serbia and Montenegro; Slovakia; Tajikistan; Turkmenistan; Ukraine; Uzbekistan
South & Central America	Antigua & Barbuda; Argentina; Bahamas; Barbados; Belize; Bolivia; Brazil; Chile; Colombia; Dominica; Dominican Republic; Ecuador; El Salvador; Grenada; Guatemala; Haiti; Honduras; Jamaica; Mexico; Nicaragua; Panama; Paraguay; Peru; Puerto Rico; Saint-Kitts&Nevis; Santa Lucia; Saint-Vincent &Grenadines; Suriname;Venezuela
Wana + Turkey	Algeria; Bahrain; Egypt; Iran; Iraq; Jordan; Kuwait; Lebanon; Libya; Morocco; Palestine; Qatar;Turkey; Saudi Arabia; Tunisia; Yemen
Sub_Saharan Africa	Benin, Botswana; Burundi; Burkina Faso; Cameroon; Côte d'Ivoire; Ethiopia; Ghana; Guinea; Kenya; Lesotho; Liberia; Madagascar; Malawi; Mali; Mauritius; Mozambique; Namibia; Niger; Nigeria; Rwanda; Senegal; Sierra Leone; South Africa; Sudan; Swaziland; Tanzania; Togo; Uganda; Zambia; Zimbabwe
Rest of Asia	Afganistan; Bangladesh; Bhutan; Brunei; Cambodia; China; Hong Kong; India; Indonesia; Laos; Malaysia; Maldives; Myanmar; Nepal; Pakistan; Philippines; Singapore; South Korea; Sri Lanka; Taiwan; Thailand; Viet Nam

We are aware that these groupings and the criteria we used are not common and may be criticized. Do they represent cultural, political and economic contexts that are homogenous within regions, or even, in our case, characteristics of survey projects? In terms of our first criterion regarding democracy, Table 8 presents the mean and standard deviation of common indices of Democracy—V_DEM Polyarchy Index of Electoral Democracy (Coppedge et al. 2018), Freedom House Political Rights and Civil Liberties indices (Freedom House, 2019) and the Economist Intelligence Unit Democracy Score (Economist Intelligence Unit, 2019)—used in the literature in order to validate our recourse to the Polity IV index as a first criterion. All the indices have been rescaled to a 0–1 scale for comparison purposes. Table 8 validates our decision to group together the consolidated democracies. Whatever the measure of Democracy used, this group of country stands as rather homogenous and different from the other groups of countries.

References

Armingeon, K., Ceka, B.: The loss of trust in the european union during the great recession since 2007: the role of heuristics from the national political system. European Union Politics **15**(1), 82–107 (2014)

Bargsted, M., Somma, N., Castillo, J.C.: Political Trust in Latin America. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on Political Trust, pp. 395–417. Northampton, Elgar (2017)

Batz, Pa, Tay, : The impact of scale transformations on national subjective well-being scores. Soc. Indic. Res. **129**, 13–27 (2016)

Bormann, N-Ch., Golder, M.: Democratic electoral systems around the world, 1946–2011. Electoral. Stud. **32**(2), 360–369 (2013)

Table 8 Mean Indices of Democracy According to Country Groupings

Country Groupings		PolityIV Democracy	PolityIV Autocracy	V_DEM Polyarchy-Electoral Democracy index	Freedom House Political Rights Index	Freedom House Civil Liberties Index	Economist Intelligence Unit Democracy Index
Consolidated democracies	**Mean**	1.000	0.000	0.859	0.986	0.942	0.795
	Standard Deviation	0.000	0.000	0.052	0.049	0.081	0.066
Post-communist Countries	**Mean**	0.694	0.108	0.615	0.664	0.680	0.588
	Standard Deviation	0.316	0.236	0.234	0.332	0.267	0.167
South and Central America	**Mean**	0.762	0.014	0.666	0.730	0.677	0.635
	Standard Deviation	0.150	0.055	0.136	0.168	0.149	0.078
West Asia, North Africa and Turkey	**Mean**	0.311	0.329	0.372	0.344	0.393	0.423
	Standard Deviation	0.310	0.281	0.181	0.221	0.164	0.107
Sub-Saharan Africa	**Mean**	0.539	0.107	0.535	0.593	0.589	0.517
	Standard Deviation	0.284	0.177	0.166	0.278	0.209	0.147
Rest of Asia	**Mean**	0.479	0.238	0.443	0.476	0.511	0.583
	Standard Deviation	0.368	0.298	0.239	0.351	0.245	0.166
Total	**Mean**	0.722	0.077	0.639	0.703	0.685	0.621
	Standard Deviation	0.294	0.192	0.216	0.290	0.237	0.155
	N	1267	1267	1304	1308	1308	726

The means are computed for the countries and years analysed. The Economist index is available from 2006 on

Bovens, M., Wille, A.: Falling or Fluctuating Trust Levels? The Case of the Netherlands. In: Zmerli, S., Hooghe, M. (eds.) Political Trust: Why Context Matters, pp. 47–66. ECPR Press, Colchester (2011)
Brewer, P., Gross, K., Vercellotti T.: Trust in international actors. In: Uslaner, EM. (ed) *The Oxford Handbook of Social and Political Trust*. (2017) https://doi.org/10.1093/oxfordhb/9780190274801.013.32
Catterberg, G., Moreno, A.: The individual bases of political trust: trends in new and established democracies. Int. J. Publ. Opin. Res. **18**(1), 31–48 (2006)
Cawvey, M., Hayes, M., Canache D., Mondak, J.: Biological and psychological influences on interpersonal and political trust. In: Uslaner, E.M. (ed). The Oxford Handbook of Social and Political Trust. (2017) https://doi.org/10.1093/oxfordhb/9780190274801.013.11
Center for Systemic Peace (2019). The Polity Project. https://www.systemicpeace.org/polityproject.html. Accessed August 16, 2019
Chan, H., Wangsit, M., Torgler, B.: Trust and Tax Morale. In: Uslaner, E.M. (ed) The Oxford Handbook of Social and Political Trust (2017). https://doi.org/10.1093/oxfordhb/9780190274801.013.23
Cleveland, W.S., Devlin, S.J.: Locally weighted regression: an approach to regression analysis by local fitting. J. Am. Statist. Assoc. **83**(403), 596–610 (1988)
Coppedge, M., Gerring J., Knutsen, C.H. et al.: V-Dem [Country-Year/Country-Date] Dataset v8, Varieties of Democracy (V-Dem) Project (2018)
Dalton, R.: Political Trust in North America. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on Political Trust, pp. 375–394. Elgar, Northampton (2017)
Dalton, R.: The social transformation of trust in Government. Int. Rev. Sociol. **15**(1), 133–154 (2005)
De Jonge, T., Veenhoven, R., Kalmijn, W.: Diversity in Survey Questions on the Same Topic: Techniques for Improving Comparability. Spinger, Berlin (2017)
De Jonge, T., Veenhoven, R., Arends, L.: Homogenizing responses to different survey questions on the same topic: proposal of a scale homogenization methods using a reference distribution. Soc. Indic. Res. **117**(1), 275–300 (2014)
Durand, C., Peña Ibarra, L.P., Rezgui, N., Wutchiett, D.: Replication Data for: Institutional Trust in the World, Université de Montréal Dataverse, V1 (2020). https://doi.org/10.5683/SP2/TGJV6G
Durand, C., Pelletier, P., Wutchiett, D.: Looking for ways to characterize countries politically and economically using longitudinal data. Harmonization Newsletter, Summer (2018), V4 (1), 3-16. https://www.asc.ohio-state.edu/dataharmonization/harmonization-newsletter-summer-2018/
Easton, D.: A Framework for Political Analysis. Prentice Hall, Englewood Cliffs (1965)
Economist Intelligence Unit (2019). Democracy Index. https://www.eiu.com/topic/democracy-index
Elff, M., Ziaja, S.: Method factors in democracy indicators. Polit. Governance **6**(1), 1 (2018). https://doi.org/10.17645/pag.v6i1.1235
Fox, J.: Multiple and Generalized Nonparametric Regression. Sage, Thousand Oaks (2000a)
Fox, J.: Non parametric Simple Regression. Sage, Thousand Oaks (2000b)
Freedom House (2019). Freedom in the world 2019. https://freedomhouse.org/report/freedom-world/freedom-world-2019
Helliwell, J., Huang, H., Wang, S.: New Evidence on Trust and Well-Being. In: Uslaner, E.M. (ed). The Oxford Handbook of Social and Political Trust (2017). https://doi.org/10.1093/oxfordhb/9780190274801.013.9
Hetherington, M., Rudolf, T.: Political Trust and Polarization. In: Uslaner, E.M. (ed). The Oxford Handbook of Social and Political Trust (2017). https://doi.org/10.1093/oxfordhb/9780190274801.013.15
Hooghe, M., Zmerli, S.: Political Trust: Why Context Matters. ECPR Press, Colchester (2011)
Hooghe, M.: Why there is basically only one form of political trust. British J. Polit. Int. Relat. **13**, 269–275 (2011)
Hox, J.: Multilevel Analysis. Techniques and Applications, 2nd edn. Routledge, New York (2010)
Hutchison, M., Johnson, K.: Political trust in Sub-Saharan Africa and the Arab Region. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on Political Trust, pp. 461–487. Northampton, Elgar (2017)
Jacoby, W.G.: Loess: a nonparametric, graphical tool for depicting relationships between variables. Electoral. Stud. **19**, 577–613 (2000)
Kolczynska, M., Slomczynski, K.: Item metadata as controls for ex post harmonization of International Survey projects. In: Johnson, T.P., Pennell, B.-E., Stoop, I.A.L., Dorer, B. (ed), Advances in Comparative Survey Methodology. Wiley, New York (2019), pp. 1011–1034
Kołczyńska, M., Schoene, M.: Survey data harmonization and the quality of data documentation in cross-national surveys. In: Johnson, T.P., Pennell, B.-E., Stoop, I.A.L., Dorer, B. (eds.) Advances in Comparative Survey Methodology, pp. 963–984. Wiley, New York (2019)

Korsgaard, M.A., Kautz, J., Bliese, P., Samson, K., Kostyszyn, P.: Conceptualising time as a level of analysis: new directions in the analysis of trust dynamics. J. Trust Res. (2018). https://doi.org/10.1080/21515581.2018.1516557
Kwak, J.: Inter-survey methodological variability in institutional trust from the survey data recycling project. Harmonization **6**(1), 18–27 (2020)
Lee, J., Paek, I.: In search of the optimal number of response categories in a rating scale. J. Psychoeducational Assess. **32**(7), 663–673 (2014)
Letki, N.: Trust in Newly Democratic Regimes. In: Uslaner, E.M. (ed). The Oxford Handbook of Social and Political Trust (2017). https://doi.org/10.1093/oxfordhb/9780190274801.013.28
Loader, C.: Local Regression and Likelihood, p. 290. Springer, New York (1999)
Luke, D.A.: Multilevel Modelling. Sage Publications, Thousand Oaks (2004)
Marshall, M.G., Gurr, T.R., Jaggers, K.: Polity IV project: Dataset users' manual. Vienna, VA: Center for Systemic Peace (2019. Accessed from http://www.systemicpeace.org/inscr/p4manualv2018.pdf
Marien, S.: Measuring Political Trust across Time and Space. In: Zmerli, S., Hooghe, M. (eds.) Political Trust: Why Context Matters, pp. 13–46. ECPR Press, Colchester (2011)
Marien, S.: The Measurement Equivalence of Political Trust. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on Political Trust, pp. 89–103. Northampton, Elgar (2017)
Matsueda, R.L., Drakulich, K.M.: Measuring collective efficacy: a multilevel measurement model for nested data. Sociol. Methods Res. **45**(2), 191–230 (2016)
Mattes, R., Moreno, A.: Social and political trust in developing countries: sub-saharan Africa and Latin America. In: Uslaner, E.M. (ed). The Oxford Handbook of Social and Political Trust (2017). https://doi.org/10.1093/oxfordhb/9780190274801.013.10
Mauk, M.: Disentangling an elusive relationship: how democratic value orientations affect political trust in different regimes. Political Res. Q., 1–15 (2019)
Mayne, Q., Hakhverdian, A.: Education, Socialization and Political Trust. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on Political Trust, pp. 176–196. Northampton, Elgar (2017)
McLaren, L.: Immigration. Ethnic Diversity and Political Trust. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on Political Trust, pp. 316–337. Northampton, Elgar (2017)
Mishler, William, Rose, Richard: What are the Origins of Political Trust? Testing institutional and cultural theories in post-communist societies. Comparat. Political Stud. **34**(1), 30–62 (2001)
Mondak, J., Hayes, M., Canache, D.: Biological and Psychological Influences on Political Trust. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on Political Trust, pp. 143–159. Northampton, Elgar (2017)
Muñoz, J.: Political Trust and Multilevel Government. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on Political Trust, pp. 69–88. Elgar, Northampton (2017)
Oleksiyenko, O.: Harmonization of ethnic minority status in international survey projects: the case of russian-speaking minority in former-soviet states. Harmonization Newsletter **2**(2), 16–19 (2017)
Oleksiyenko, O., Wysmułek, I., Vangeli, A.: Identification of Processing Errors in Cross-national Surveys. In Johnson, T.P., Pennell, B-E., Stoop, I. A. L., Dorer, B. (ed) Advances in Comparative Survey Methodology, Wiley. 985–1010 (2019)
Park, C.-M.: Political Trust in the Asia-Pacific Region. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on Political Trust, pp. 488–508. Northampton, Elgar (2017)
Raudenbush, S.W., Bryk, A.S., Cheong, Y.F., Congdon, R.T., Du Toit, M.: *HLM7 Hierarchical Linear and Nonlinear Modeling User Manual: User Guide* for Scientific Software International's (S.S.I.) Program, SSI, U.S.A., Skokie, Il (2016)
Rose, R., Mishler, W.: Political Trust and Distrust in Post-Authoritarian Contexts. In: Zmerli, S., Hooghe, M. (eds.) Political Trust: Why Context Matters, pp. 117–140. ECPR Press, Colchester (2011)
Schneider, I.: Can we trust measures of political trust? assessing measurement equivalence in diverse regime types. Soc. Indic. Res. **133**, 963–984 (2017)
Simms, L.J., Zelazny, K., Williams, T.F., Bersntein, L.: Does the number of response options matter? psychometric perspectives using personality questionnaire data. Psychological Assess. **31**(4), 557–566 (2019)
Singer, J.S., Willet, J.B.: Applied longitudinal data analysis: Modeling change and event occurrence (2003)
Singh, R.K.: Harmonizing Instruments with Equating. Harmonization **6**(1), 11–18 (2020)
Snijders, T., Bosker, R.: Multilevel Analysis An introduction to Basic and Advanced Multilevel Modeling, 2nd edn. Sage Publications, London (2012)
Tabachnik, Fidell: *Using Multivariate Statistics*, 7th Edition, Pearson (2019)
Thisted Dinesen, P., Mannemar Sonderskov, K.: Cultural Persistence or Experiential Adaptation? A Review of Studies using Immigrants to Examine the Roots of Trust. In: Uslaner, E.M. (ed). The

Oxford Handbook of Social and Political Trust (2017). https://doi.org/10.1093/oxfordhb/9780190274801.013.27
Tomescu-Dubrow, I., Slomczynski, K.M.: Harmonization of cross-national survey projects on political behavior: developing the analytical framework of survey data recycling. Int. J. Sociol. **46**(1), 58–72 (2016)
Torcal, M.: Political Trust in Western and Southern Europe. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on Political Trust, pp. 417–439. Northampton, Elgar (2017)
Trudinger, E.-M., Bollow, U.: Evaluations of Welfare State Reforms in Germany: Political Trust Makes a (Big) Difference. In: Zmerli, S., Hooghe, M. (ed.), Political Trust: Why Context Matters, Colchester, ECPR Press, p. 187–212 (2011)
Uslaner, E.: Political Trust, Corruption and Inequality. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on Political Trust, pp. 302–315. Northampton, Elgar (2017)
Van der Meer, T.W.G., Ouattara, E.: Putting 'political' back in political trust: an IRT test of the unidimensionality and cross-national equivalence of political trust measures. Quality Quantity **53**, 2983–3002 (2019). https://doi.org/10.1007/s11135-019-00913-6
Van der Meer, T.: Democratic input, macroeconomic output and political trust in Zmerli, S, and T. Van der Meer (2017). *Handbook on Political Trust*, Northampton, U.S.A. 270–284
Van der Meer, T., Dekker, P.: Trustworthy States, Trusting Citizens? A Multilevel Study. In: Zmerli, S., Hooghe, M. (eds.) Political Trust: Why Context Matters, pp. 95–116. ECPR Press, Colchester (2011)
Van der Meer, T.: In What We Trust? A Multi-Level Study into Trust in Parliament as an Evaluation of State Characteristics. Int. Rev. Administrat. Sci. **76**(3), 517–536 (2010a)
Van der Meer, T.: En quoi fait-on confiance? Une étudie multi-niveaux de la confiance dans le parlement en guise d'évaluation des caractéristiques de l'État. Revue des Sciences Administratives **76**, 543–561 (2010b)
Van der Meer, T., Te Grotenhuis, M., Pelzer, B.: Influential Cases in Multilevel Modeling: a Methodological Comment. Am. Sociol. Rev. **75**(1), 173–178 (2006)
Van Deth, J.: Compliance, Trust and Norms of Citizenship. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on Political Trust, pp. 212–227. Northampton, Elgar (2017)
Van Erkel, P.F.A., van der Meer, T.W.G.: Macroeconomic Performance, Political Trust and the Great Recession: a Multilevel Analysis of the Effects of Within-Country Fluctuations in Macroeconomic Performance on Political Trust in 15 EU Countries, 1999–2011. Eur. J. Polit. Res. **55**(1), 177–197 (2016)
Van Ham, C., Thomassen, J.: The myth of legitimacy decline: an empirical evaluation of trends in political support in established democracies. In: van Ham, C., Thomassen, J., Aarts, K., Andeweg, R. (eds.) Myth and Reality of the Legitimacy Crisis. Explaining Trends and Cross-National Differences in Established Democracies, pp. 17–36. Oxford University Press, Oxford (2017)
Wilson, R.: Trust Experiments, Trust Games, and Surveys. In: Uslaner, E.M. (ed). The Oxford Handbook of Social and Political Trust (2017). https://doi.org/10.1093/oxfordhb/97801902748O1.013.2
Wilkes, R.: Trust in Government: a Micro-macro Approach. J. Trust Res. **4**(2), 113–131 (2014). https://doi.org/10.1080/21515581.2014.889835
Zavecz, G.: Post-communist Societies of Central and Eastern Europe. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on Political Trust, pp. 440–4460. Northampton, Elgar (2017)
Zmerli, S., Newton, K.: Winners, Losers and Three Types of Trust. In: Zmerli, S., Hooghe, M. (eds.) Political Trust: Why Context Matters, pp. 67–94. ECPR Press, Colchester (2011)
Zmerli, S., Hooghe, M.: Political Trust: Why Context Matters. ECPR Press, Colchester (2011)
Zmerli, S., Newton, K.: Objects of Political and Social Trust: Scales and Hierarchies. In: Zmerli, S., van der Meer, T.W.G. (eds.) Handbook on political Trust, pp. 104–124. Northampton, Elgar (2017)
Zmerli, S., Van der Meer, T.: Handbook on Political Trust, Northampton, U.S.A. (2017)