Cross-national Studies: Interdisciplinary Research and Training Program (CONSIRT)
The Ohio State University and the Polish Academy of Sciences

Volume 4, Issue 1
Summer 2018

Editors
Irina Tomescu-Dubrow
and
Joshua K. Dubrow
CONSIRT

consirt.osu.edu/newsletter
ISSN 2392-0858

Harmonization: Newsletter on Survey Data Harmonization in the Social Sciences

Harmony

Welcome to the new issue of *Harmonization: Newsletter on Survey Data Harmonization in the Social Sciences*. Harmony, as defined by the Oxford English Dictionary (online), is "the combination of simultaneously sounded musical notes to produce a pleasing effect." We are not musicians, so our task is to amplify *scientific* notes sounded by the ever growing community of scholars, institutions, and government agencies who work on the harmonization of data. The international social science community is composed of many disciplines and each scholar and research team strikes chords on data, methods, and theory simultaneously; their diverse sounds produce a pleasing effect.

This issue features notes on a variety of topics. **Tom W. Smith** writes about how cross-national research programs can improve comparativeness at the project and program levels. **Claire Durand, Paul Pelletier** and **David Wutchiett** build on a previous *Harmonization* item (vol. 2, no. 2) to address why institutional trust varies across world regions. **Ewa Jarosz** discusses the harmonization of time-use surveys. **The V-Dem Team** presents the latest version of their renowned democracy project. The **Survey Data Recycling** team reflects on their participation in the latest Comparative Survey Design and Implementation (CSDI) workshop.

As with every issue of *Harmonization*, we welcome your articles and news. Please send them to the newsletter co-editor Josh Dubrow at dubrow.2@osu.edu.

In This Issue…

Articles

Making Comparative Survey Research More Comparative: Collaboration and Coordination across Cross-National Survey Research Programs

by Tom W. Smith

Both the comparison error perspective under the total survey error paradigm and the traditional functional equivalence approach need to be applied to maximize the comparability in cross-national and cross-cultural research. But beyond these essential steps, comparative survey research can and should do much more to enhance comparativeness at the project and program levels.

First, cross-national programs should pay close attention to what other, similar projects are doing. They can improve their surveys both methodologically and substantively by studying the methods and content of the other comparative surveys. John Donne observed, "No man is an island entire of itself; every man is a piece of the continent, a part of the main…" This is also true regarding comparative surveys. Each should be seen as a complement to the others with all together serving the general goal of understanding global human society.

Second, cross-national programs should, when possible, standardize study and sample designs, demographics, and other components to facilitate cross-project analysis. For example, standards and formats of documentation should be pre-harmonized, demographics should either follow common optimal designs or, when pre-harmonization is not possible, be designed so closely equivalent post-harmonization is facilitated.

Third, general management and operational tools should be developed to be utilized across cross-national projects. Too much work in cross-national research is project-, even survey-, specific. This is inefficient and increases both the cost and time it takes to complete projects. Fortunately, tools to both manage cross-national surveys and conducted essential steps are being developed and these should be employed across projects. Their use will improve the performance of each individual, cross-national project and increase comparability across such projects.

Some of the best work has been developed by two European Union programs: 1) DASISH – Data Service Infrastructure for the Social Sciences and Humanities, 2012-2014 (dasish.eu/about_dasish) and 2) SERISS – Synergies for Europe's Research Infrastructure in the Social Sciences, 2015-2019 (seriss.eu). SERISS tools developed or being developed include the Translation Management Tool, Fieldwork Management and Monitoring System, Data Harmonisation Platform, Questionnaire Design and Development Tool, and Questionnaire Variables Database.

Also, the work of the project on Survey Data Recycling: New Analytic Framework, Integrated Database, and Tools for Cross-national Social, Behavioral and Economic Research is advancing the science of data harmonization (dataharmonization.org).

The art and science of doing large-scale, cross-national survey research will be notably improved if coordination across projects and programs increases and comparative, survey research will advance both methodologically and substantively as a result of such collaboration.

Tom W. Smith is the Senior Fellow and Director of NORC's Center for the Study of Politics and Society, Ogburn Stouffer Center for the Study of Social Organizations, at the University of Chicago. He is the Director of the General Social Survey (GSS) and a co-founder of the International Social Survey Program (ISSP).

Looking For Ways to Characterize Countries Politically and Economically Using Longitudinal Data

by Claire Durand, Paul Pelletier and David Wutchiett

This research note follows a preceding one (Durand, Valois and Peña Ibarra 2016/2017), which presented a project aimed at combining survey datasets that included questions related to trust in institutions. Our project has progressed very well, featuring a database with 756 datasets from the Barometers, LAPOP and World Values Surveys combined from 98 countries from the regions of Central and South America, Sub-Saharan Africa, West Asia and North Africa (WANA), and Asia. Research using an earlier version of the database with 635 surveys (Durand, Peña Ibarra, Charest and Pelletier 2017) has shown that region, as defined above, explains around 40% of the between-country variance in institutional trust. Why does trust vary between regions? The aim of this research note is to explore possible answers to this question.

Methods

In order to understand why trust varies between regions, we need to examine whether we can find indicators of socio-economic and political characteristics that differ between regions and appear related to trust. In looking for these indicators, we face two problems. First, it is difficult to find information that is equally relevant and available across the different regions of the world and for each country, and second, that are consistently available over the period of evaluation. Many socio-economic and political indicators are available only for countries of the "Western World" or only for part of the period. Since most characteristics of interest, including trust, change over time, we need to take this into account in our analyses.

Data

On institutional trust

The database includes more than 1M respondents who reported on their level of trust in various institutions, altogether more than 100 specific institutions. Since the different survey projects do not use the same answer scales, we harmonized the data so that scales of four anchors were translated into a common 7-point scale, where 4,3,2,1 (from high trust to no trust) are recoded into 1, 3, 5 and 7. We produced a data file where each answer to a trust question from a respondent corresponds to one line. This data file has 13M answers on trust, coded for the institution to which the question pertains. The original institutions are recoded into 14 categories that range from political institutions – state and government, political parties, elections, international organizations – to administrative institutions – public administration, army, police, judiciary system – to civil society institutions – media, church, trade unions, NGO – and those of the economic system – banks, enterprises.

For this paper, we focus on trust in the State or Government. All the responses to trust in the Government, the State, the Parliament, and the President are included in this category. Before grouping them into the same category, we checked that the means and standard deviations for the specific institutions were similar within project and region.

On the Characteristics of Countries

We were interested in synthetic indicators that were both relevant according to the literature on determinants of trust and were available for most countries and most years. We tried to identify all databases where data regarding the socio-political and economic situation of countries were available. These include the Quality of Governance database (QOG, qog.pol.gu.se/data), the V-DEM database (v-dem.net), the Polity sixth edition (systemicpeace.org), and Solt's Standardized World Income Inequality database (SWIID, fsolt.org/swiid). We took data on the same countries the database on trust covers and starting with 1995 to be able to compare similar periods for trust and for the other characteristics. To measure countries' economic situation we selected the Solt Gini indices and GDP. To measure countries' social situation we selected indicators of a country's proportion of urban population and indicators of the population's ethnic and religious diversity. To measure countries' political situation we selected the Polity2 index of political regime.

Analyses

We needed to characterize the countries according to the change in their characteristics over time rather than their situation at a given point in time. To do this, we explored clustering methods for longitudinal data.

The Process

The first step in any cluster analysis is to decide on the appropriate number of clusters. The most well-known index to inform this decision for longitudinal data is the Calinski-Harabatz (CH) index (for the description of CH and other indexes of the number of clusters, see Desgraupes 2013). However, other

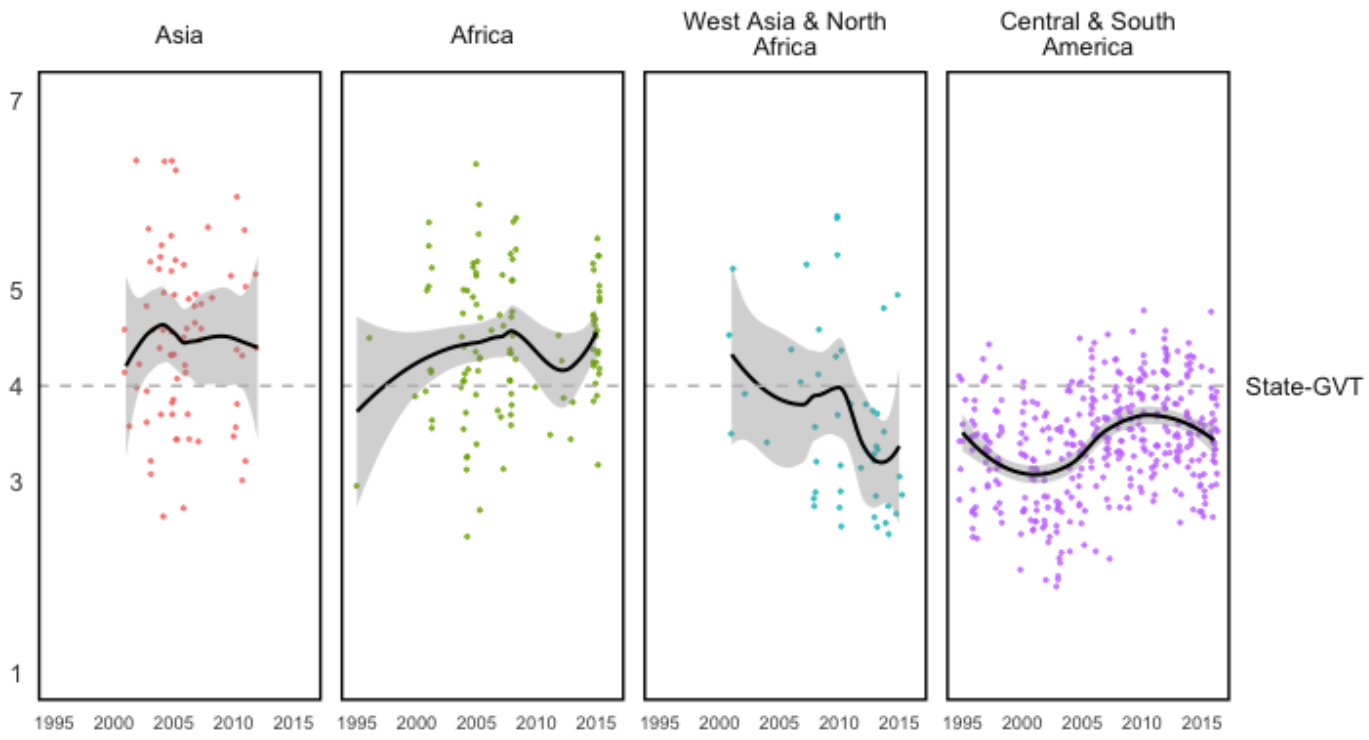
indices may also be used, which do not always suggest the same number of clusters. We performed two analyses, one using only the CH original index and another comparing five different indices: three types of CH index, the Ray-Turi and the Davies-Bouldin indices, respectively. This information helps with making the final decision as to the most appropriate number of clusters. However, other considerations must also be taken into account: there should be enough countries in each cluster, and, additionally, in our case, we would like to identify intra-regional homogeneity and inter-regional differences. Therefore, we would prefer four-cluster solutions that differentiate across the four regions we are studying: Central and South America, Sub-Saharan Africa, West Asia and North Africa (WANA), and Asia. Finally, the clusters must seem to "make sense" given our understanding of the concepts, methods, data and socio-political circumstances.

The next step is to determine the most appropriate procedure for clustering. We explored two different methods for longitudinal clustering available in the R packages: *traj and K-Means longitudinal (KML)*. We compared these methods for some of the measures and concluded that most of the time, they gave very similar results. However, when the methods produced different clusters, in general KML gave clusters that seemed more appropriate considering knowledge of institutional and political circumstances corresponding to the different countries. Therefore, we selected KML for the analyses presented here.

The final step consists of performing the analyses. The first series of analyses aimed at clustering countries based on change in trust in the State or Government. The second series aimed at clustering these same countries on socio-economic and political indicators expected to be related to trust. In this note, we present selected results regarding trust in the State or Government, the Solt Gini dispositional index, the GDP, the countries' proportion of urban population, indicators of ethnic and religious diversity, and political regime.

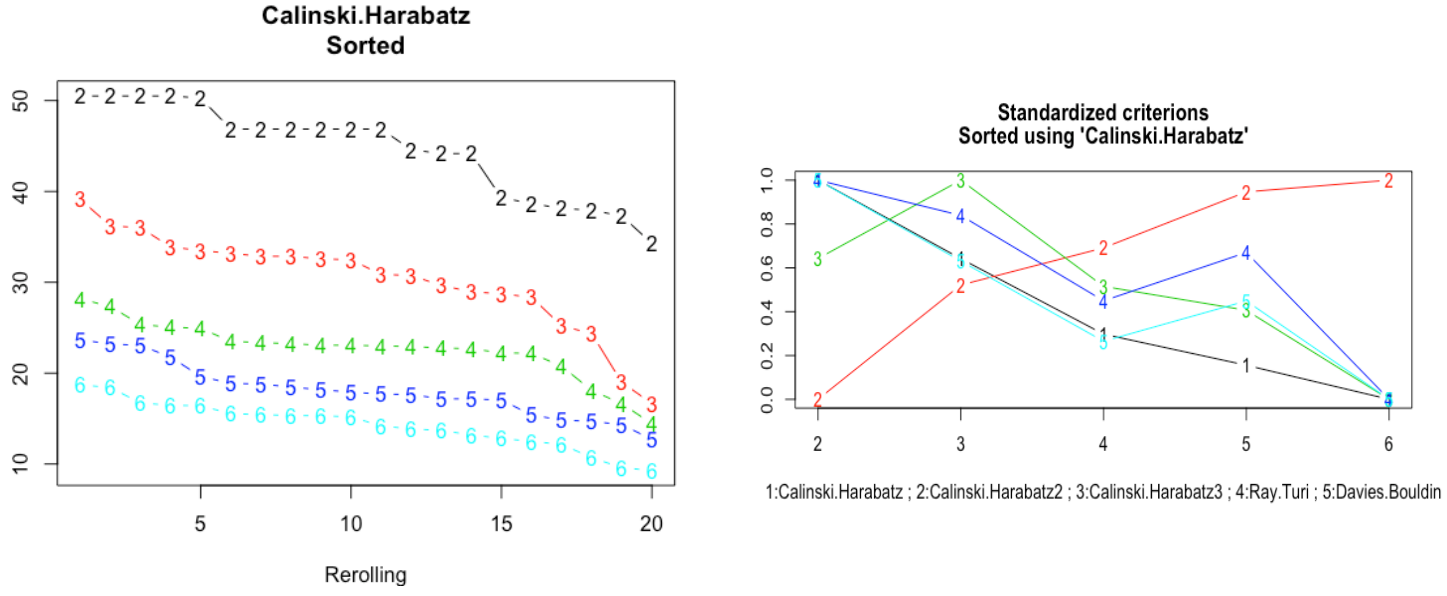
Indicators of institutional trust

As a first step, we examine the trends in trust in the State or Government for each region. In **Graph 1**, the dots represent the average trust for a given year in a given country and the lines are from local regressions smoothing across the various estimates. The graph shows that trust is particularly low in South and Central America and that it is declining sharply in the WANA region. Furthermore, there is much variation within region. The graph also shows that there are more data in South and Central America than in the other regions, where international survey projects have started later and in only some countries.



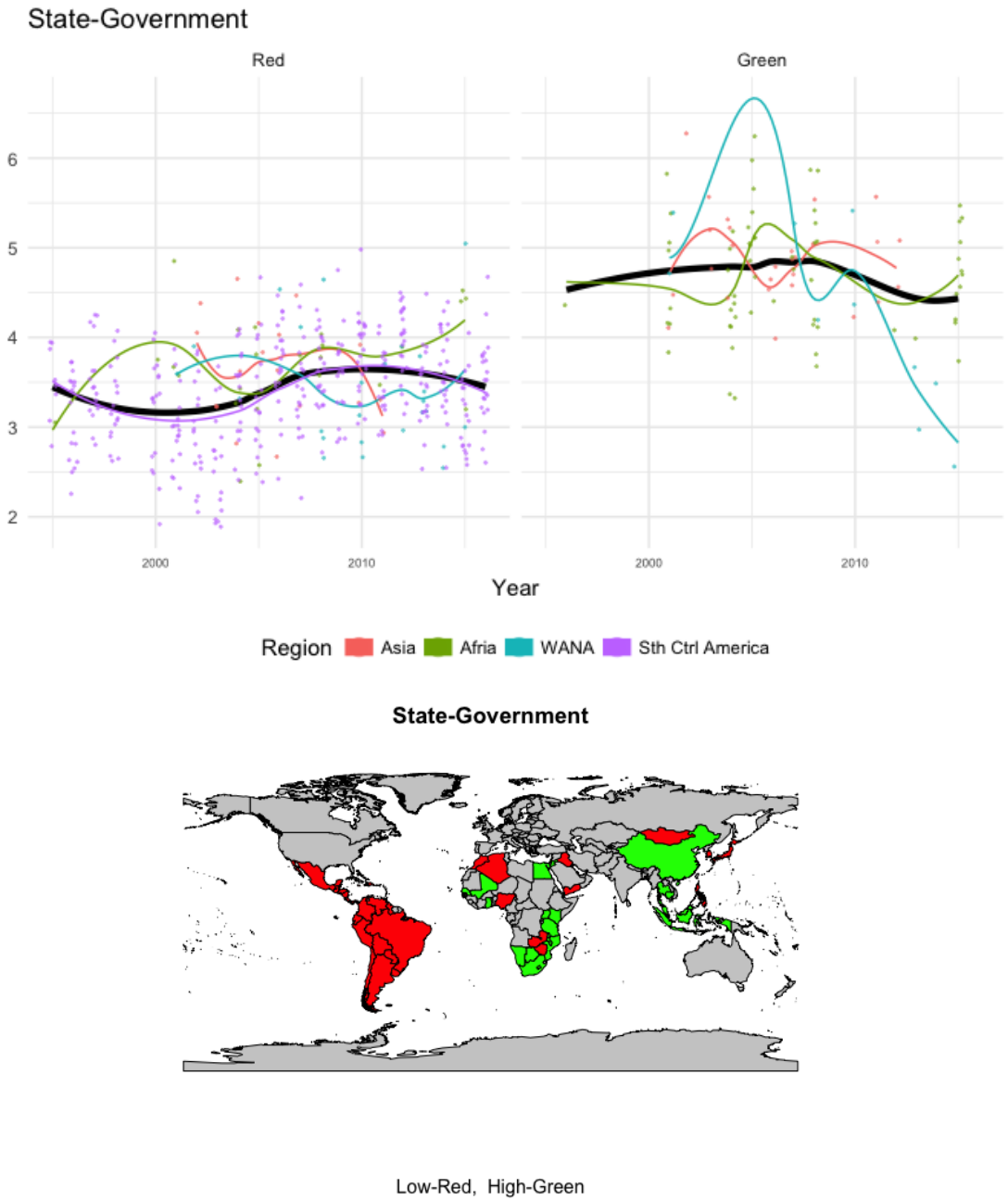
Graph 1. Trends in average trust in State or Government per region

Let us now turn to cluster analysis. For this first analysis, we show how the different criteria perform in terms of indicating the most appropriate number of clusters. We first run the Calinski-Harabatz test (see **Graph 2**). This test gives an indication of the ratio of between and within cluster variance. Higher estimates correspond to a maximization of between-cluster over within-cluster variance. The test clearly indicates that a two-cluster solution is the most appropriate.



Graph 2. Testing number of clusters for trust in State or Government

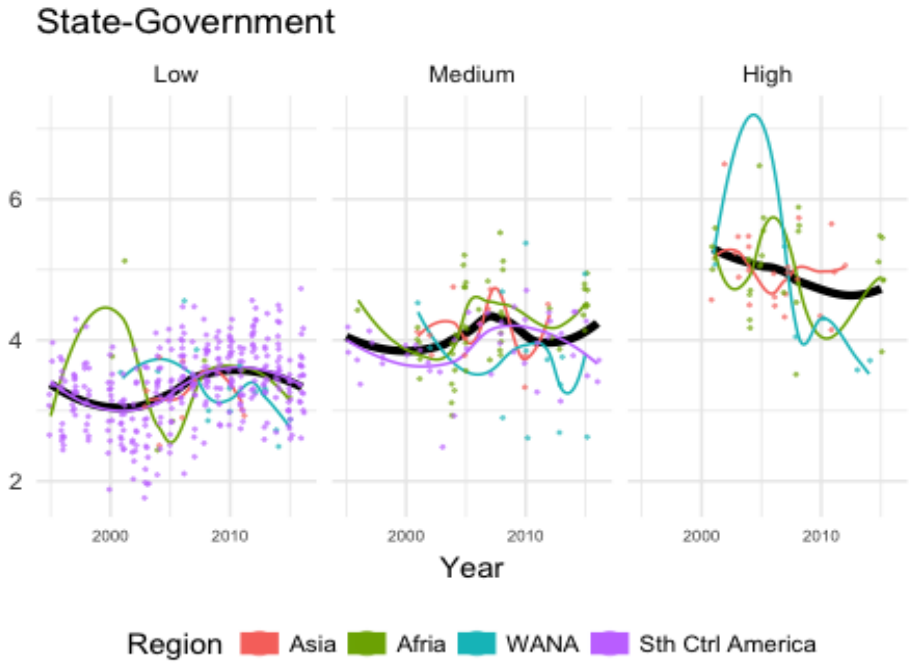
In addition, Graph 2 shows how different criteria compare. These criteria also seek to maximize between cluster variance but use different denominators for the ratio of between to within-cluster variance (CH2 and CH3), or use different indicators of the within-cluster distance between cases. Three criteria have their highest value at two clusters (the first column). However, the CH3 index suggests a three cluster solution while the CH2 index suggests that a six-cluster solution might be better.

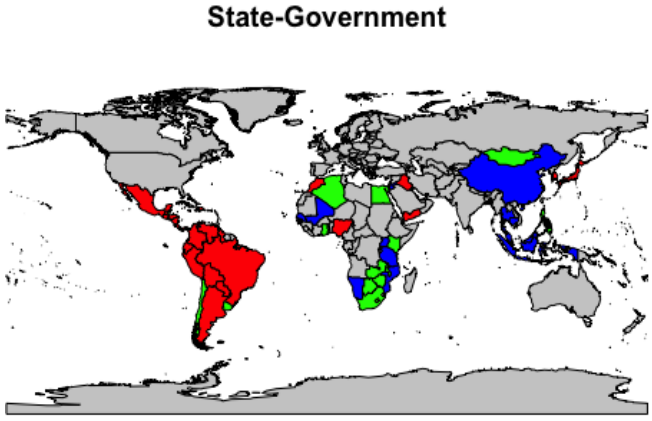


Graph 3. Two-cluster solution for trust in State or Government

With a two-cluster solution, we end up with one cluster of high trust and one of low trust. The first part of **Graph 3** shows the variation within each cluster. The large black line shows the mean trajectory, the thin coloured lines show mean regional trajectories and the dots show each country's yearly value of average trust in the State or Government. On the high trust cluster graph, the green line shows that the variation for the WANA region is very large, due partly to the lack of data.

As the map in Graph 3 illustrates, with this two-cluster solution, the countries of South and Central America are all in the low trust cluster (in red) and the countries of Asia are almost all in the high trust cluster. In Africa and West Asia, the countries are spread between the two clusters.





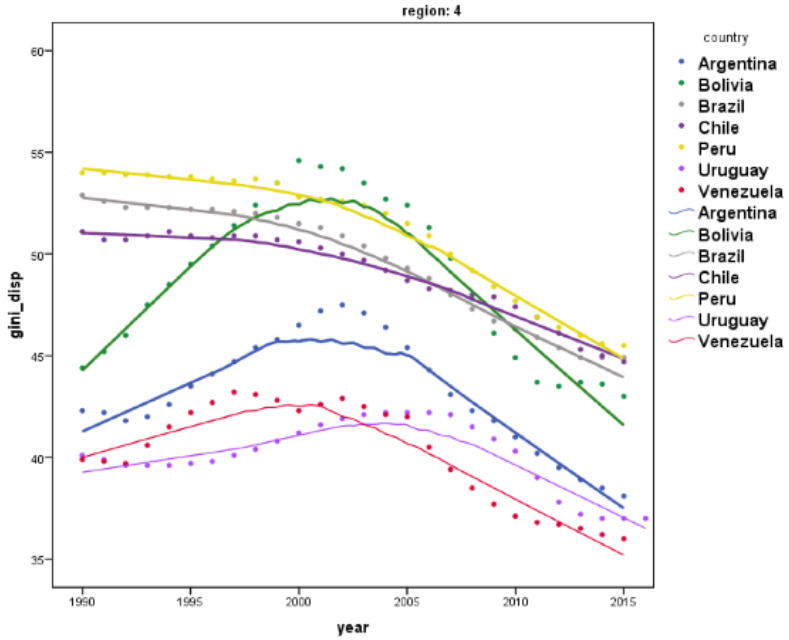
Low-Red, Medium-Green, High-Blue

Graph 4. Three-cluster solution for trust in State or Government

If we choose to keep three clusters instead of two, we end up with one cluster of varying low trust – whose trend is very similar to the trends seen in Graph 1 for South and Central America, one of varying average trust, and one of declining high trust (see **Graph 4**). These two last clusters cannot be easily related to what we see in Graph 1. The map in Graph 4 shows that, except for Chile and Uruguay, the countries of South and Central America are in the low trust cluster, in red, while most of Asia is in the high declining trust cluster, in blue. A mix of the different clusters characterizes Sub-Saharan Africa and the WANA region. It is also in these regions that we find most of the countries with average trust levels (in green). Whatever the solution, we do not end up with four clusters of trust concentrated in different regions. *South* and Central America is the most homogenous region.

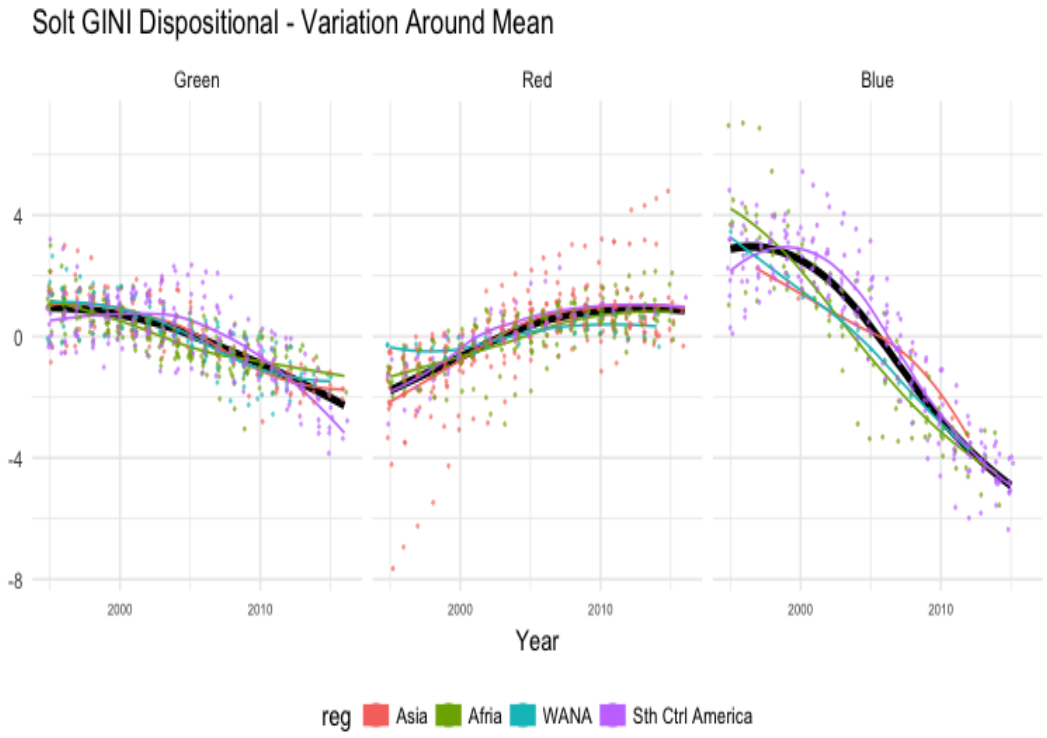
Indicators of the economic situation

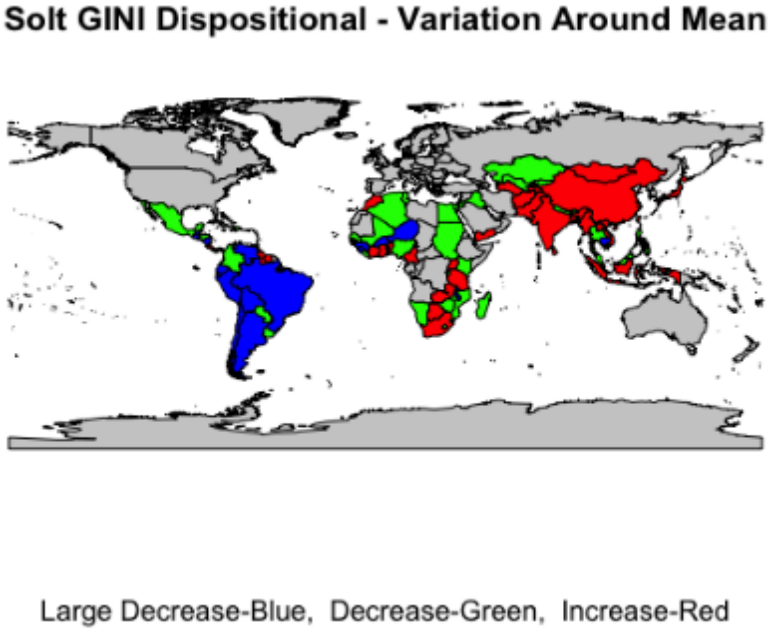
Solt proposes two indices of income inequality. The *dispositional index* is an index of inequalities post tax, post transfer, contrary to the *market index* that is pre-tax, pre-transfer. We first present in **Graph 5** the trajectories of the dispositional index from 1990 to 2016 for a selection of Central and South American countries, to give a concrete idea of what we are trying to better understand. The graph shows three groups of trajectories. A first group – Peru, Brazil and Chile – has a high GINI index at the beginning of the period, followed by a small but constant decrease until the end of the period. The second group – Argentina, Uruguay and Venezuela – starts with a low GINI index that increases until the 2000's and decreases afterwards. Finally, there is a specific curve for Bolivia, a country where inequalities increased dramatically according to the GINI index, peaking shortly before the election of the left-wing president Evo Morales in 2006. Under his mandate, inequality, as measured by the GINI index, decreased substantially.



Graph 5. Solt GINI for selected countries of Central and South America

These paths differ not only in their GINI level but also in their trajectory. These patterns and relationships are what we aim to recover in the next analyses. The indices for the number of clusters showed that solutions with either two, four or five clusters were equally appropriate. However, whatever the number of clusters, this analysis tended to give clusters that were determined mostly by the level of inequalities and not by the trends over time.

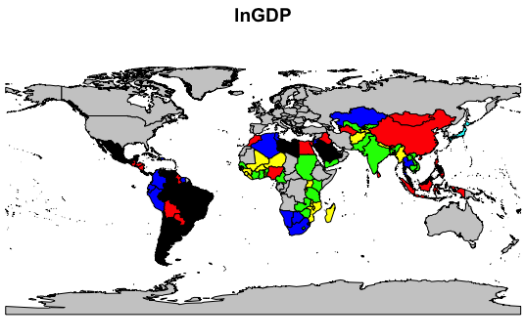




Graph 6. Three-cluster solution for Solt GINI Dispositional – Variation around the mean

To focus on the shape of change over time, we conducted another analysis where we replaced the raw values by the deviation around each country's mean value. **Graph 6** shows the results. We found that a three-cluster solution was most appropriate. The analysis produces three differing shapes, one curve showing an increasing trend and the two others declining, one slowly and the other more rapidly. We see substantial homogeneity within these clusters.

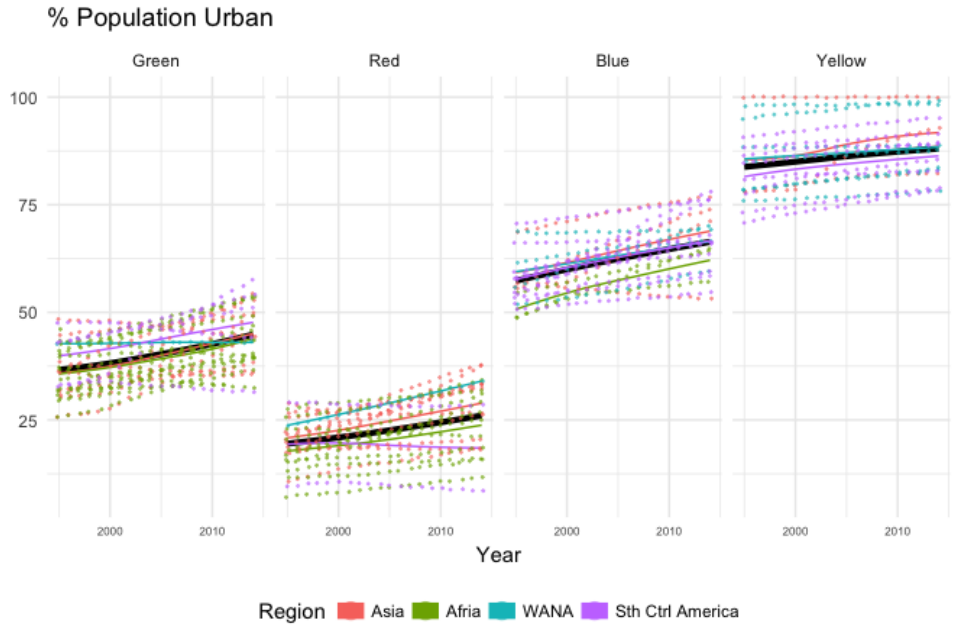
The map in Graph 6 shows that the large declining trajectory cluster (blue) is mostly present within South America whereas the increasing inequality trajectory (in red) disproportionately appears within Southeast Asia and Sub-Saharan Africa. The gradual declining trajectory (in green) is present across the four regions. The same analyses using the GINI market index give similar results. Therefore, we have two sets of clustering results for the GINI indexes, one determined mostly by levels of inequalities, the other by the shape of change over time. However, it seems that the quadratic trajectories that we have seen in Graph 5 are grouped with the high declining trajectory.

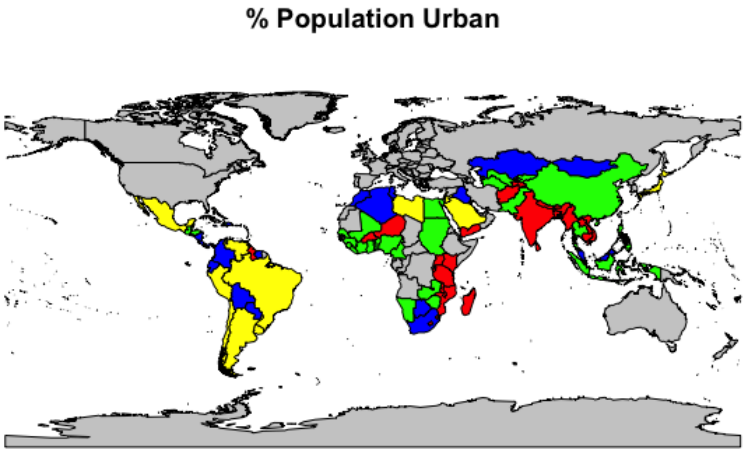


Lowest-Yellow, Low-Green,
Low Middle-Red, High Middle-Blue,
High-Black, Highest-Cyan

Graph 7. Six-cluster solution for Ln (GDP)

In the case of Ln (GDP), the shape of the trajectories is similar and increasing from 2000 on – more substantially after 2008 – for all clusters. Therefore, the clusters essentially differentiate countries according to the level of GDP. We rely on the six-cluster solution, presented in **Graph 7**. The low GDP clusters are concentrated in Sub-Saharan Africa, Southeast Asia, India and Bangladesh, while mid-range GDP countries were identified in the WANA region, Southeast Asia and Central and South America. The high GDP clusters comprise the Gulf States, Japan, South Korea and Singapore in Asia and the larger Latin American countries such as Brazil, Argentina and Mexico. There is no within region homogeneity in terms of clusters of GDP levels.





Low-Red, Low Middle-Green, High Middle-Blue, High-Yellow

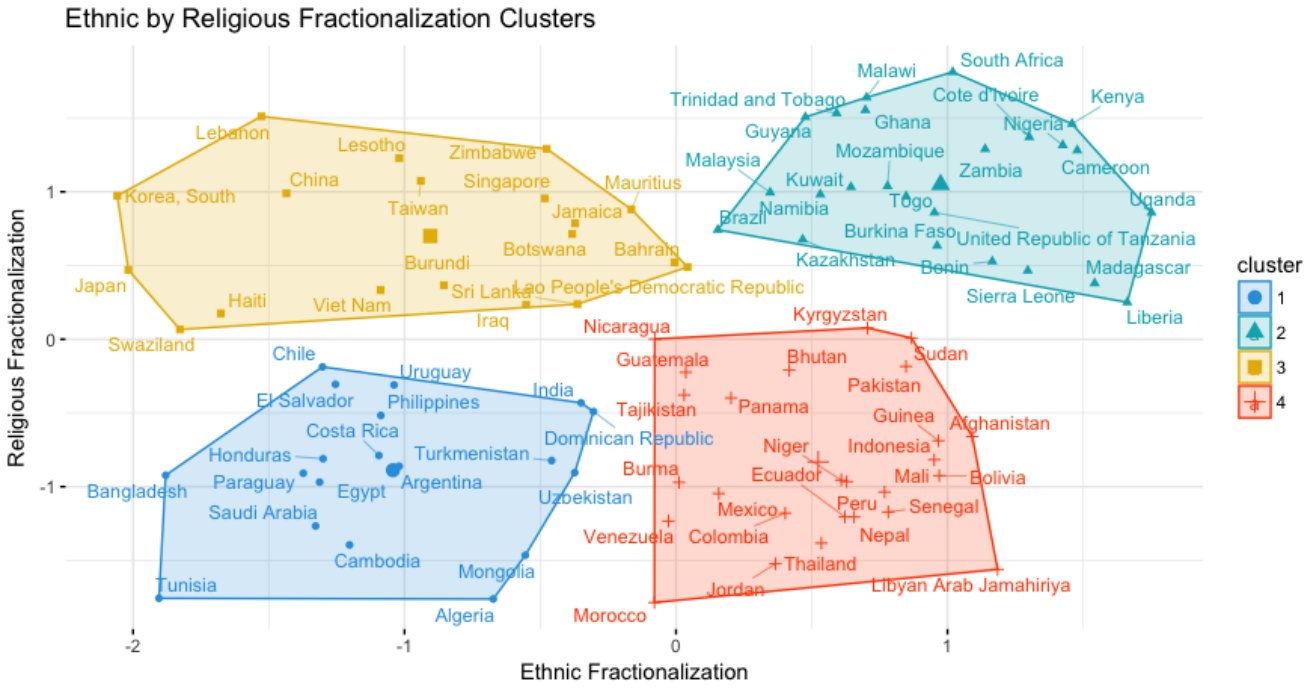
Graph 8. Four-cluster solution for proportion of urban population

Indicators of the social situation

One indicator of the countries' social situation is the proportion of their urban population. As with GDP, urbanization is always growing and never regressing; the rate of progression is largely similar across all regions. Therefore, clustering takes into account mostly the level of urbanization. The different indicators of the appropriate number of clusters do not converge but give similar scores for retaining four clusters. As shown in **Graph 8**, a medium-high to high degree of urbanization characterizes the countries of South and Central America. The situation is quite different in the other regions, where each of the four clusters of urbanization are present.

Diversity is also expected to be associated with trust. Researchers posit that high diversity fosters low trust, particularly interpersonal trust. This hypothesis may also apply at the institutional level: it may be more difficult for an institution to foster trust from very diverse citizens with different views and needs than from citizens with a more homogenous culture.

The data available to measure diversity are problematic. Data are available for most countries and years in the QOG database but the information does not vary over time. Given this situation, we instead clustered the countries according to their ethnic and religious fractionalization (Alesina, Devleeschauwer, Easterly, & al., 2003) at the reported constant levels. The evaluation criteria for the number of clusters suggested that the most appropriate solution was a four-cluster solution (Graph 9). One cluster grouped together countries of low diversity, another, countries of high diversity, this according to both indicators; a third cluster grouped low ethnic, high religious diversity countries and a final one included high ethnic, low religious diversity countries.

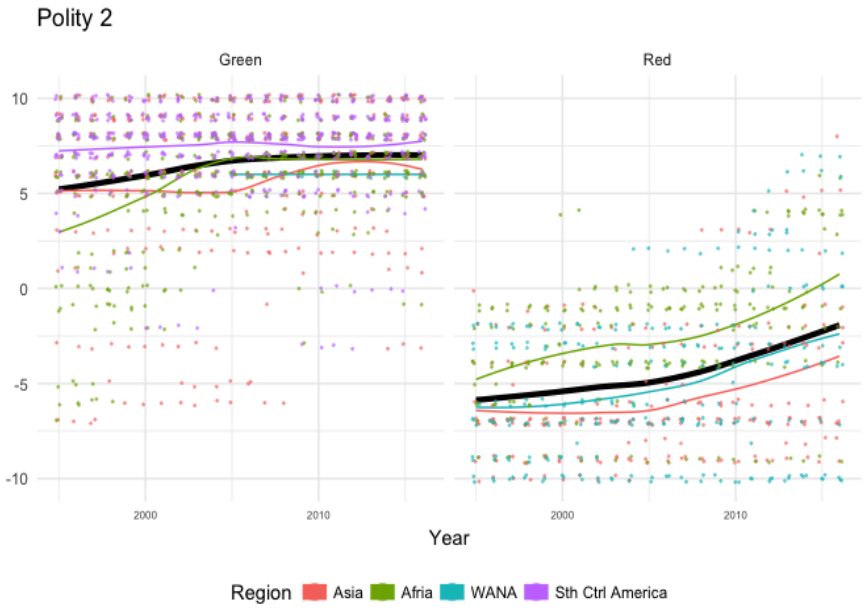


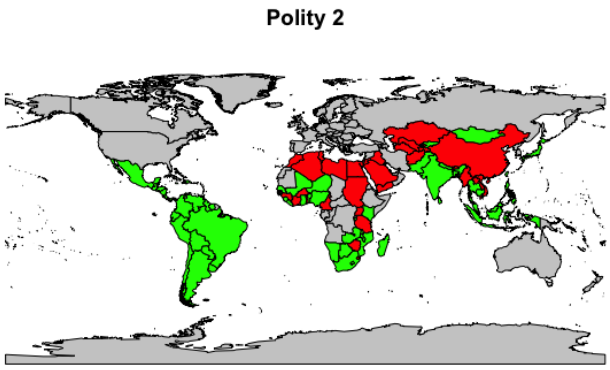
Graph 9. Four-cluster solution for ethnic and religious fractionalization

As we can see in **Graph 9**, each cluster includes countries from most of the regions. Therefore, the level and type of diversity of the population is not a regional characteristic.

Indicator of the political situation

We were looking for indicators based on objective features of the political systems, indicators that do not rely on perceptions. The Polity 2 index elaborated by the Polity IV project respects this criterion. The index ranges from -10 to +10 on a scale from autocracy to democracy and is based on features of the political systems – intensity of political competition, regulation of political participation, competitiveness of executive recruitment, openness of executive recruitment, and the constraint it placed on its chief executive. As seen in **Graph 10**, the index does not change much over time for the different countries although the more authoritarian countries tend to become more democratic.





Democracy-Green, Autocracy-Red

Graph 10. Two-cluster solution for Polity 2

The KML analysis with two clusters gives one cluster for the more democratic countries and another for the more autocratic countries. The map in Graph 10 shows that all of South and Central America together with most of the countries of Africa and Asia are in the democratic cluster (in green) while more autocratic regimes characterize the WANA region and parts of Asia and Africa.

Discussion and conclusion

In summary, some characteristics – diversity of the population, political regime – are quite stable. Therefore, they could be related to different levels of trust but they can hardly explain change. In addition, diversity does not characterize specific regions. Other characteristics – GDP, urbanization – increase at the same pace in the different countries. They could also be related to the level of trust in the State or Government, but not to differential change. Finally, some indicators – the GINI indices – vary between countries both in their level and in the shape of their trajectory over time. For these characteristics, we performed two groupings: one for the level and one for the shape. Again, there is no homogeneity within region. There are at least two and sometimes three different clusters present in each region. In general, the best statistical solution is not a four-cluster solution and, when it is the case, there is no homogeneity within regions.

Do these cluster analyses help understand regional differences? Let us focus on Latin America, a region where most of the countries group together in the same cluster of trust in the State or Government. For the period under study, this same region is characterized by democratic regimes together with medium to-high levels of GDP and of urbanization. It is also rather homogenous in terms of trends in the GINI indices. Most countries are in the declining inequalities cluster. However, the region is very heterogeneous in terms of ethnic and religious diversity. What does it mean in terms of interpretation? Can we conclude that democracy and urbanization are associated with low trust? Or is it a regional particularity that is related to historical factors? What influences what? How can we explain that an improving economic situation seems concurrent with low trust?

We analyzed these same data using multilevel analysis. The preliminary analyses focused on average institutional trust, a variable that includes trust in the State or Government. The results do not validate an impact of the inequalities – as measured by the GINI dispositional index or its within-country deviation over time – on average institutional trust. However, they show a positive impact of ln (GDP) and a negative impact of the proportion of urban population on average trust. More detailed analyses show that the Polity2 index of political regime is negatively associated with trust in the State or Government, as well as with trust in the political parties. However, it is positively associated with trust in the electoral process. This would lead to the conclusion that, in more democratic regimes, the process by which leaders are chosen is still trusted but the institutions involved in the political battles and the eventual leaders elected following the process are not that much trusted.

These variables explain about 15% of the between-country variance, as compared with 40% for the regions. Therefore, we think we are heading in the right direction but more analyses still need

to be conducted. The next steps will be to explore the KML shape algorithm to see if it can grasp the quadratic trends present in our data and to explore new indicators of the regional situation. In particular, we are looking for better indicators of inequalities and of the social and political situation.

Claire Durand is professor and David Wutchiett is a doctoral student, both at the Department of Sociology, University of Montreal. Paul Pelletier is a consulting statistician. The presentations related to this article are available on ResearchGate or on Claire Durand's website mapageweb.umontreal.ca/durandc

References

Alesina, Alesin, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg, R. (2003). "Fractionalization." *Journal of Economic Growth, 8*, 155-94.

Desgraupes, Bernard. 2013. *Clustering Indices*. Paris : University Paris Ouest Lab Modal'X cran.r-project.org/web/packages/clusterCrit/vignettes/clusterCrit.pdf

Durand, Claire, Isabelle Valois and Luis Patricio Peña Ibarra. 2016/2017. "Harmonize or Control? The Use of Multilevel Analysis to Analyze Trust in Institutions in the World." *Harmonization: Newsletter on Survey Data Harmonization in the Social Sciences* 2 (2): 4-12.

Durand, Claire, Luis Patricio Peña Ibarra, Anne-Sophie Charest and Paul Pelletier. 2017. "How to Explain Variation in Institutional Trust in the World, a 4-Level Longitudinal Model." Paper presented at the WAPOR 70th conference, Lisbon, Portugal, July 16th. mapageweb.umontreal.ca/durandc/Recherche/Publications/confiance/How_to_explain_4level.pdf

Harmonizing Time-Use Data

by Ewa Jarosz

Economic data capture national and domestic production by using its financial equivalent. However, a great share of what people do does not leave any monetary trace. Childcare, domestic work, helping a neighbour or elderly household member, working in unregistered jobs or selling goods on the black market do not get captured in economic data. They can only be measured at the point of input, that is by the amount of time individuals dedicate to perform those tasks. National-level time-use surveys allow to estimate the amount of work performed everyday by stay-at-home mothers or by community activists. But they do much more. Because time-use surveys allow to describe in detail individual's days, they also permit measuring differences in lifestyle practices and existing inequalities with regard to such issues as domestic division of labour between men and women, class differences in duration and type of childcare, health behaviours, and many more.

First time-use surveys were collected in 1913 in the UK and the US (Bevans, 1913; Pember-Reeves, 1913). Only a few years later, following Lenin's directive, this type of data were collected in the Soviet Union to measure social and industrial progress (Sacks, 1977). The first multinational time-use study was conducted by Alexander Szalai in 1965 and included the data from single cities in Poland, Hungary, Yugoslavia, Czehoslovakia, Germany (GDR), France, Belgium, United States, and the USSR (Szalai, 1966). Interest in time allocation was further fuelled by the work of Gary Becker, and his theory on allocation of time within a household (1965). Today time data are collected across the world and are widely used by international organizations and researchers to measure social development, inequality, informal economy, among other. Due to their prevalence they also allow for cross-national comparisons once they are harmonized. Ex-post survey data harmonization (e.g. Granda, Wolf and Hadorn 2010) is necessary, since the original surveys record data in many formats and using different categories of activities or even different variables.

Time-use surveys provide a record of daily activities carried out by a sample of respondents.

Arguably the best example of harmonized time use data is the Multinational Time Use Study (MTUS; Gershuny and Fisher, 2014) administered by the Centre for Time Use Research at the University of Oxford. MTUS includes harmonized time-use data from nearly 30 countries, altogether covering the period between 1965 and 2015. It is regularly updated to include newly released dataset. Data are available to registered users in the form of raw data (timeuse.org/mtus/access) or as an extract from the data which can be created using MTUS-X extract builder. MTUS-X is administered by the Minnesota Population Center and the Maryland Population Research Center in collaboration with the Centre for Time Use Research (mtusdata.org/mtus). User guides for all datasets mentioned above are available online.

Time-use surveys are a specific type of data providing a record of daily activities carried out by a sample of respondents. Usually the diary record is accompanied by one or more datasets providing background information on the respondent (individual-level data), household (household level data), and the day of the diary (such as when the diary was completed or what was the weather on the diary day). The diary data themselves come in different matrix shapes, and often the first stage of the harmonization process requires unifying matrices across the datasets. The second stage is to create common variables that capture the types of activities that respondents recorded doing, and other variables describing activity setting, such as location (where the activity took place), social context (with whom), or any additional activity characteristics, such as whether a device was used (ICT), or how much a respondent enjoyed what he or she did. Finally, all background variables need to be harmonized for the purpose of cross-sectional or longitudinal comparisons. In time-use data, the process of ex-post harmonization of background variables is the same as in the case of other types of surveys, so I will not discuss it here. Instead, I will focus on the process of harmonization of time use

sequences, which poses specific challenges to the researcher and needs to be carried out in a way specific to this particular type of data.

Data format

There are two basic formats of matrix arrangement: wide and long. Wide format means that all slots of activities (primary and secondary) are recorded as separate variables in columns and each diary is recorded in a single row of the matrix. Long format means that there is a primary activity variable in one column, secondary activity in another, but the data are stacked, that is recorded in several rows for each diary, each row representing another activity recorded in the diary. In this scenario each slot also has a unique number within the diary to identify its place within a sequence.

There are also two basic ways of recording activities in time: recoding them in slots or in episodes. Slots are time periods of equal duration. This is a typical way of recording activities within the Harmonized European Time Use Survey (HETUS) framework (ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-RA-08-014). In this type of set-up, the diary consists of 144 10-minutes slots that respondents fill in. An alternative approach is to record time via episodes. If the diary is based on episodes, it is entirely to the respondent to specify how long an activity took. Episodes are of different duration (sometimes as short at 1 minute), but they all sum up to 1440 minutes per day (24 h).

Together, the matrix form and the way activities are recorded (in slots or episodes) give four basic combinations of how time data might be arranged (see Table 1). If the data to be harmonized come in different formats, which is highly likely, harmonization requires matrix transposition and –if relevant – conversion of slots into episodes. Transposing the matrix is a simple procedure, available in all software packages. Converting time slots into episodes enables comparisons with regard to such measures as the number of episodes of a particular activity per day, or mean duration of an episode of a given activity. Conversion into the same unit is routinely done for any time-use comparison. If the original datasets come in different formats (slots or episodes), conversion is always done from slots into episodes. It is not possible to convert episodes into slots, because episodes are of different durations, which are usually not dividable by 10.

Table 1: Basic forms of organization of time diary data

Wide episodic	Wide slot
Long episodic	Long slot

Are slots and episodes fully comparable once rendered into the same unit? No. This is because of the fact that they represent, in practice, slightly different modes of data collection. In episode-based diaries respondents are free to record any activity, including very short ones that last below 10 minutes (a typical duration of a single slot). A slot-based diary is more likely to capture only activities that took 10 minutes or more. Furthermore, a slot-based diary forces respondents to round-up the actual time an activity took to match the slot unit. So very short activities might get extended to 10 minutes, and activities lasting for, say 15 minutes might be recorded as either a 10-minutes activity or 20-minutes activity. Slot-based diaries are therefore more orderly but less exact. A single dataset is always recorded either in slots or in episodes; it is not possible that these formats are mixed.

Activities, locations, and who was around

Activity sequences[1] include main activity sequence and accompanying activity sequence (also known as secondary activity) - usually there is only one accompanying activity though it is possible to collect more activities, and that is the case for the recent UK time-use survey, which collects three accompanying activities (UK Time Use Survey from 2014/2015; Gershuny and Sullivan, 2017). The harmonization process requires that all sequences are comparable, that is, if a sequence including two activities (main and secondary), and four activities (main, secondary, and two accompanying ones) is to be harmonized, some level of detail will be lost. Because most of the accompanying activity sequences is empty (in case of secondary activity it is around 80% of slots or episodes), that is no additional activity is recorded, the costs of such deletion are not high. Furthermore, it is very rare of time-use surveys to record more than two sequences (main and secondary activity). What usually gets harmonized is therefore primary activity, secondary activity and variables known as 'activity settings' that is location, people present, sometimes use of ICT during an activity, or other activity-specific information, depending on the purpose of harmonization.

Creating common measures of activities across different surveys raises its own difficulties. Ex-post harmonization involves recoding both the primary and secondary activity. Different activities get recoded into a new, harmonized variable, with unified activity codes. These codes most likely will be less detailed than those provided in the original (source) variables. One challenge lies in whether the location code should also be incorporated into the activity variable, for example, should eating a work lunch in the canteen be coded as work lunch, or eating, or eating at work? Similarly, eating and drinking in a pub might be coded as eating and drinking or being in a pub – where location provides the actual name for an activity. At the very least, users of time-use data harmonized ex-post need access to accurate documentation of the harmonization process and a codebook of the target variables clarifying which activities (and locations) are included in the new activity codes. A more complex solution could

[1] Sequences are chains of events in which the order of events is intentional and might be used to produce other variables describing the sequence itself. In fact there are several sequences that run parallel – main activity sequence, secondary activity, and sequences of various activity settings. In practice these are not disentangled but used as characteristics of events. Each event in time is described by what was done at that time (main activity), what else, if anything was done simultaneously (secondary activity), who else was present (co-presence), where the activity was performed (location) and other event-specific information. Length of the sequence is defined by the number of events and their summary duration.

be that of recording the harmonization process via metadata, and make the metadata available in the harmonized dataset (see Slomczynski and Tomescu-Dubrow, forthcoming).

Nearly all diaries also collect information on who else was present when a given activity was carried out. Harmonizing this information is usually straightforward, as it only requires assigning particular 'people present' codes to each episode for which co-presence was reported. For example, the MTUS recode produces a variable indicating whether 1/ respondent was alone, 2/ a partner was present, 3/ a child under 18 years old was present. Original data might include very detailed information about the presence of different household members, extended family and other people, but unless such details are required for specific analysis, they are not included in the harmonized variables. Each dataset harmonized within the MTUS project comes with a detailed file describing how particular variables were created.

Ex-post harmonization of time use data is technically complex. Loss of detail is a necessary part of the process. Nonetheless, harmonization efforts are worthwhile since time sequences are a complex and comprehensive source of information of what a day in a life of a population or individuals within this population looks like.

Ewa Jarosz is a Research Fellow in the Centre for Time-Use Research, Department of Sociology, University of Oxford. She is also affiliated with Laboratory for Comparative Social Research at the Higher School of Economics in Moscow, Russia, and with the Institute of Philosophy and Sociology at the Polish Academy of Sciences.

References

Becker, G., (1965). *A theory of the allocation of time*, The Economic Journal, 75, pp. 493-517

Bevans, G., (1913). *How Working Men Spent Their Time*. New York: Columbia University Press

Gershuny J., Fisher K. (2014) Multinational Time Use Study. In: Michalos A.C. (eds) *Encyclopaedia of Quality of Life and Well-Being Research*. Springer, Dordrecht

Gershuny, J., Sullivan, O. (2017). *United Kingdom Time Use Survey, 2014-2015*. Centre for Time Use Research, University of Oxford. [data collection]. UK Data Service. SN: 8128, http://doi.org/10.5255/UKDA-SN-8128-1

Granda, Peter, Christof Wolf, and Reto Hadorn. (2010). Harmonizing Survey Data. In: Janet A. Harkness, M. Braun, B. Edwards, T. Johnson, Lars Lyberg, Peter Ph. Mohler, B.-E., Pennell and Tom W. Smith, eds. *Survey Methods in Multinational, Multicultural and Multiregional Contexts*. Hoboken, NJ: John Wiley & Sons (pp. 315-332).

Pember-Reeves, M. (1913/1979). *Round about a Pound a Week*. London: Virago.

Sacks, M. (1977). Unchanging Times: A Comparison of the Everyday Life of Soviet Working Men and Women between 1923 and 1966. *Journal of Marriage and Family*, 39(4), 793-805. doi:10.2307/350483

Slomczynski, Kazimierz M and Irina Tomescu-Dubrow. Forthcoming. Basic Principles of Survey Data Recycling. Ch. 43 in *Advances in Comparative Survey Methodology: Multinational, Multiregional and*

Multicultural Contexts (3MC), T.P. Johnson, B-E Pennell, I. A. L. Stoop, & B. Dorer (eds), Wiley Hoboken, New Jersey

Szalai, A., (1966), The Multinational Comparative Time Budget Research Project: A Venture In International Research Cooperation, *American Behavioral Scientist*, Vol 10, Issue 4, pp. 1 – 31.

Varieties of Democracy (V-Dem) Data on Democracy

by Marina Povitkina, on behalf of the Varieties of Democracy team

Varieties of Democracy (V-Dem, v-dem.net/en) is a new approach to conceptualizing and measuring democracy. V-Dem project has assembled a multidimensional and disaggregated dataset that reflects the complexity of the concept of democracy as a system of government that exceeds the presence of elections. The V-Dem project distinguishes between seven high-level principles of democracy: *electoral, liberal, participatory, deliberative, egalitarian, majoritarian and consensual*, and collects data to measure these principles.

V-Dem is one of the largest social science data collection efforts with a database containing over 19 million data points. By May 2018, the dataset covers 201 countries from 1789 to 2017 with annual updates to follow. The V-Dem dataset is available for download, free of charge at v-dem.net/en/data/data-version-8. Users are also able to explore the V-Dem data through online analysis tools, with which they can create and save their own graphs at v-dem.net/en/analysis/analysis.

In addition to its 450+ indicators covering many aspects of countries' political systems, V-Dem aggregates indicators into mid- and high-level indices designed to measure different dimensions of democracy. High-level indices measure the types of democracy, conceptualized by V-Dem, which are *electoral, liberal, participatory, deliberative, egalitarian, majoritarian and consensual* types. Mid- level indices capture concepts such as *clean elections, women's political empowerment,* and *civil society participation*.

The V-Dem team

The V-Dem project is a collaborative international effort that unites thousands of social scientists researching democracy and countries' political systems. The V-Dem project has recruited approximately 3,000 experts from almost every country in the world to provide knowledge on democracy in countries over time. The project is managed by leading scholars in the democracy research, who work closely with experts.

Methodology

V-Dem draws on theoretical and methodological expertise from its worldwide team to produce data in the most objective and reliable way possible. Approximately half of the indicators in the V-Dem dataset are based on factual information obtained from official documents such as constitutions and government records. The other half consists of more subjective assessments from experts on topics

like political practices and compliance with *de jure* rules. The assessment involves 3,000 country experts from almost every country in the world with typically no less than five experts rating each country on different concepts for each year. Having multiple experts per country provides the opportunity to statistically account for both uncertainties about estimates and potential biases that experts may evince.

In general, expert-coded data raise concerns regarding comparability across time and space. Rating complex concepts requires judgment, which may vary across experts and cases. Moreover, because even equally knowledgeable experts may disagree, it is imperative to report measurement error to the user.

To address these concerns, V-Dem uses a custom-built Bayesian measurement model that minimizes coder error and addresses issues of comparability across countries and over time. The basic logic behind the model is that an unobserved latent trait exists, but we are only able to see imperfect manifestations of this trait. By taking all of these manifest items (in this case, expert ratings) together, we are able to provide an estimate of the trait. In the dataset, V-Dem presents the user with a best estimate of the value for an observation *(the point estimate)*, as well as an estimate of uncertainty *(the credible regions*, a Bayesian corollary of confidence intervals).

V-Dem uses a custom-built Bayesian measurement model that minimizes coder error and addresses issues of comparability across countries and over time.

To facilitate cross-country comparability, V-Dem has encouraged country experts to code multiple countries using two techniques. The first is **bridge coding**, in which an expert codes the same set of questions for one extra country for the same time period as the original country they coded. This technique is particularly useful when the two countries have divergent regime histories: it provides experts the opportunity to code the full range of the ordinal question scale, thus providing more information as to where an expert's thresholds are. By extension, this information also provides a better sense of the thresholds of this coder's colleagues who only coded one of the countries she coded.

The second technique is **lateral coding**. Its purpose is to gain detailed information regarding an individual expert's thresholds by asking her to code many different aspects in countries over a certain period of time that are coded by a variety of other experts. By comparing her coding to those of many other experts, V-Dem is able to gain a greater sense of how a given coder systematically diverges from other experts' coding; conversely, V-Dem also gains information on how those other experts diverge from this given coder.

Finally, V-Dem employs **anchoring vignettes** to further improve the estimates of expert-level parameters and thus the concepts they measure. Anchoring vignettes are descriptions of hypothetical

cases that provide all the necessary information to answer a given question. Since there is no contextual information in the vignettes, they provide a detailed information about how individual experts understand the scale itself.

V-Dem also draws on the academic expertise of the leading social science research methodologists to develop theoretically informed techniques for aggregating indicators into mid- and high-level indices. In this sense, V-Dem is at the cutting edge of developing new and improved methods of social science measurement (Coppedge et al 2018). For more information regarding V-Dem's methodology, see Pemstein et al (2018), Marquardt and Pemstein (forthcoming), Pemstein, Tzelgov and Wang (2015) or consult Policy Brief #17 by Maxwell, Marquardt, and Lührmann (2018) for a concise summary.

Relevance

The nuanced V-Dem data can be used by academics studying the nature, causes, and consequences of democracy in quantitative research, as well as by governments, development agencies, and NGOs for critical decisions such as selecting country program priorities, informing program design, and monitoring the impact of their programs (Mechkova and Sigman 2018).

The data provide a comprehensive image on every aspect of countries' political life. The dataset can be used for time series analysis, as it provides estimates stretching far back over time and captures historical developments. It can also be used for time-series cross-sectional analysis or cross-sectional analysis, as it deals with cross-country comparability issues more than any other existing dataset on countries' political characteristics. Examples of articles using the V-Dem data can be found in the V-Dem Institute's collection of Working Papers v-dem.net/en/news-publications/working-papers and in the Social Science Research Network (SSRN) depository.

Funding

V-Dem has been or is currently funded by: Riksbankens Jubileumsfond, Knut & Alice Wallenberg Foundation, The Ministry of Foreign Affairs - Sweden, the Swedish Research Council, the European Commission/EuroAID, Marianne & Marcus Wallenberg Foundation, International IDEA, the Ministry of Foreign Affairs - Denmark, the Danish Research Council, the Canadian International Development Agency, NORAD/the Norwegian Research Council, Aarhus University, Fundação Francisco Manuel dos Santos, Mo Ibrahim Foundation, Knut & Alice Wallenberg Foundation and the Quality of Government Institute, with co-funding from University of Gothenburg and University of Notre Dame.

Note: *The information about the project is taken from the Policy Brief #5, 2016 by Valeriya Mechkova and Rachel Sigman and Policy Brief #17, 2018 by Laura Maxwell, Kyle L. Marquardt and Anna Lührmann with the permission from the authors.*

Marina Povitkina is a postdoctoral researcher at the Center for Collective Action Research at the University of Gothenburg, Sweden. She is affiliated with the Quality of Government Institute and the Varieties of Democracy Institute at the University of Gothenburg.

References

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Agnes Cornell, Sirianne Dahlum, Haakon Gjerløw, Adam Glynn, Allen Hicken, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Juraj Medzihorsky, Moa Olin, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Johannes von Römer, Brigitte Seim, Rachel Sigman, Jeffrey Staton, Natalia Stepanova, Aksel Sundström, Eitan Tzelgov, Yi-ting Wang, Tore Wig, Steven Wilson, and Daniel Ziblatt. 2018. *"V-Dem [Country-Year/Country-Date] Dataset v8"*. Varieties of Democracy (V-Dem) Project.

Marquardt, Kyle L. and Daniel Pemstein. Forthcoming. "IRT Models for Expert-Coded Panel Data." *Political Analysis*.

Maxwell, Laura, Marquardt, Kyle and Lührmann, Anna. 2018. *The V-Dem Method for Aggregating Expert-Coded Data. V-Dem Policy Brief*, No. 17, 2018, University of Gothenburg, Varieties of Democracy Institute.

Mechkova, Valeriya and Sigman, Rachel. 2016. *Varieties of Democracy (V-Dem). V-Dem Policy Brief, No. 5*, October 2016, University of Gothenburg, Varieties of Democracy Institute.

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang, Joshua Krusell and Farhad Miri. 2018. *"The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data"*. University of Gothenburg, Varieties of Democracy Institute: Working Paper No. 21, 3d edition.

Pemstein, Daniel, Eitan Tzelgov and Yi-ting Wang. 2015. "Evaluating and Improving Item Response Theory Models for Cross-National Expert Surveys." *University of Gothenburg, Varieties of Democracy Institute:* Working Paper No. 1.

Examples of Marina Povitkina's articles using V-Dem data are:

Povitkina, Marina. 2018. "The Limits of Democracy in Tackling Climate Change." *Environmental Politics* 27(3): 411–32.

Persson, T.A., and M. Povitkina. 2017. "'Gimme Shelter': The Role of Democracy and Institutional Quality in Disaster Preparedness." *Political Research Quarterly* 70(4): 833–47.

Conferences & Workshops

The 2018 CSDI International Workshop in Limerick, Ireland (March 26-28) featured Survey Data Recycling (SDR) Work

by Irina Tomescu-Dubrow

The International Workshop on Comparative Survey Design and Implementation (CSDI, csdiworkshop.org) was founded in 2002 in Brussels at a meeting of researchers active in cross-national or cross-cultural survey research. The main goal of CSDI is to improve comparative survey design, implementation and related analysis. Its annual workshops provide a forum and platform of collaboration for scholars involved in research relevant for comparative survey methods.

At the 2018 CSDI meeting in Limerick, Ireland, members of the Survey Data Recycling (SDR) Team from the Polish Academy of Sciences and The Ohio State University organized the session *Data Quality and Comparability in Survey Data Harmonization*. Its focus were methods of evaluating and improving the comparability of existing survey datasets, strategies for establishing functional equivalence across surveys, and strategies for data analysis in the framework of *ex post* harmonization of comparative survey projects.

Presenters in the *Data Quality and Comparability in Survey Data Harmonization* session included WAPOR President **Claire Durand** and **Paul Pelletier**, University of Montréal, **Verena Ortmanns**, GESIS-Leibniz Institute for the Social Sciences, **Ilona Wysmułek**, Institute of Philosophy and Sociology PAN, and **Marta Kołczyńska, Irina Tomescu-Dubrow** and **Kazimerz M. Slomczynski**, Institute of Philosophy and Sociology PAN and CONSIRT.

CSDI presentations are available at csdiworkshop.org

Acknowledgement: We thank the National Science Foundation, USA (PTE Federal award 1738502) and the Institute of Philosophy and Sociology, Polish Academy of Sciences for financial and organization support for this session.

News

Short Course at GESIS: Open Science and Open Data. Lecturers are Dr. Sebastian Netscher, Dr. Anja Perry, and Anna K. Schwickerath. Place: Cologne, Germany. Dates of the course: 20.08 - 21.08.2018. The workshop focuses on the idea of Open Science and Open Data, taking data creation as well as data re-use into account. On the one hand, it introduces the FAIR principles to guide

researchers in creating re-usable research data. On the other hand, the workshop discusses the re-use of already existing research data and relevant aspects to keep in mind, working with intellectual property of others. With regard to the re-use of existing data, it introduces a (free) tool that helps researchers in the process of harmonizing their data. The workshop furthermore discusses aspects of legal and analytical re-usability as well as of replicability of research findings, in terms of Open Codes, e.g. in the context of data harmonization. For more information, including the syllabus, visit training.gesis.org.

Harmonization would like to hear from you!

We created this Newsletter to share news and help build a growing community of those who are interested in harmonizing social survey data. We invite you to contribute to this Newsletter. Here's how:

1. Send us content!

Send us your announcements (100 words max.), conference and workshop summaries (500 words max.), and new publications (250 words max.) that center on survey data harmonization in the social sciences; Send us your short research notes and articles (500-1000 words) on survey data harmonization in the social sciences. We are especially interested in advancing the methodology of survey data harmonization. Send it to: Joshua K. Dubrow, dubrow.2@osu.edu.

2. Tell your colleagues!

To help build a community, this *Newsletter* is open access. We encourage you to share it in an email, blog or social media.

Support

This newsletter is a production of Cross-national Studies: Interdisciplinary Research and Training program, of The Ohio State University (OSU) and the Polish Academy of Sciences (PAN). The catalyst for the newsletter was a cross-national survey data harmonization project financed by the Polish National Science Centre in the framework of the Harmonia grant competition (2012/06/M/HS6/00322). This newsletter is now funded, in part, by the US National Science Foundation (NSF) under the project, "Survey Data Recycling: New Analytic Framework, Integrated Database, and Tools for Cross-national Social, Behavioral and Economic Research" (SDR project - PTE Federal award 1738502). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The SDR project is a joint project of OSU and PAN. For more information, please visit dataharmonization.org.

Copyright Information