

[Trust in Bureaucracy]*

Yuehong ‘Cassandra’ Tai *Pennsylvania State University*
Frederick Solt *University of Iowa*

[0/150 words]

Keywords: xxxxx, political trust, measurement

Intro

[Y]et the comparative aspects of public administration have largely been ignored, and as long as the study of public administration is not comparative, claims for a “science of public administration” sound rather hollow (Dahl 1947, pp. 8). Comparative studies provide novel insights and innovatory concepts but face methodological challenges (Pollitt 2011). Measurement equivalence is a critical one of these challenges. Specifically, non-equivalent measures across countries poses a threat to comparative public administration study, yielding biased results, wrong theoretical conclusions, and misleading policy implications (Jilke, Meuleman and Van de Walle 2015).

The lack of comparable data is especially prominent in comparative public administrative survey research. As a core topic in comparative administration, analyses on trust connect classic administration theories and behavioral perspectives (Van Ryzin 2011). However, studies on trust and public administration in across national analyses are susceptible to lack of comparable data and have to be restricted to a small number of country-level units, mainly focusing on OECD countries (see summary in Van de Walle and Migchelbrink 2022). Jilke, Meuleman and Van de Walle (2015) found that ignoring the incomparability of trust measures across countries can produce misleading conclusions. For example, without accounting for item bias, Swedes have higher trust in public institutions than citizens in Canada and the United States, which is exactly the opposite when non-equivalence caused by item bias is modeled.

The causes and consequences of trust in government, in legislative institutions, and judiciary on democratic governance, legitimacy of legitimacy, direction of public policy, and public compli-

*Corresponding author: frederick-solt@uiowa.edu. Current version: September 27, 2022.

ance in emergencies have been widely discussed over decades (Easton 1975; Chanley, Rudolph and Rahn 2000; Rogowski and Stone 2021; Goldstein and Wiedemann 2021). However, trust in public administration has long been ignored by scholarly inquiry (Rogowski and Stone 2021), although public trust is particular important for administrative agencies. Since the public cannot monitor and control agencies and civil servants directly, trust in bureaucracy is required to grant agencies and officials to act in the public’s interest (Thomas 1998). Agencies and civil servants are agents who implement policies, deliver public services and goods, and contact with citizens directly and frequently, public trust determines the public’s acceptance to delivered goods and compliance with public policies (Morelock 2021). With trust in bureaucracy, the public could support “the implementation of policy programs” (Kim 2005, p. 611). On contrast, without public trust, public officials could struggle in performing their tasks and attaining the public’s collaborative responses in emergencies (Yates 1982; Van Ryzin 2011).

Among few existing studies, data is limited in both regional scale and time periods (Morelock 2021; Choi 2018; Houston et al. 2016), which severely impedes causal inferences in dynamic relationship between trust in bureaucracy and quality of administration. Specifically, the lack of comparative data on trust on bureaucracy makes it impossible to test the competing theories about what affects trusting attitudes, government performance or quality of governance (Bouckaert 2012; Kettl 2000; Van de Walle and Migchelbrink 2022; Morelock 2021).

In this letter, we present the trust in civil servants/(bureaucracy/public administration) (TCS) dataset, which is based on the host of national and cross-national survey data available and recent advances in latent variable modeling of public opinion that allow us to make use of this sparse and incomparable data. It provides comparable estimates of the trust and confidence the public puts in civil servants and public administrators/administration across countries and over time. We validate the data by showing that these TCS scores are strongly correlated with responses to single survey items as well as with measures of [perceived corruption, unemployment, income inequality, and internal and external efficacy, the rule of law, government effectiveness]. We expect that the TCS data will become an invaluable source for broadly cross-national and longitudinal research on the causes and effects of trust in the civil service.

Examining the Source Data on Trust in Bureaucracy

National and cross-national surveys have asked questions on trust attitudes toward public administrations over the past half-century, but the resulting data are both sparse, that is, unavailable for many countries and years, and incomparable, generated by many different survey items. In all, we identified 16 such survey items that were asked in no fewer than five country-years in countries surveyed at least twice; these items were drawn from 107 different survey datasets.¹

Together, the survey items in the source data were asked in 136 different countries in at least two time points over 48 years, from 1973 to 2021, yielding a total of 1,768 country-year-item observations. Observations for every year in each country surveyed would number 6,528, and a complete set of country-year-items would encompass 104,448 observations. Compared to this complete set of country-year-items, the available data can be seen to be very, very sparse. From a more optimistic standpoint, we note there are 1,370 country-years in which we have at least *some* information about the trust in civil servants of the population, that is, some 49% of the 2,797 country-years spanned by the data we collected. But there can be no denying that the many different survey items employed renders these data incomparable and difficult to use together.

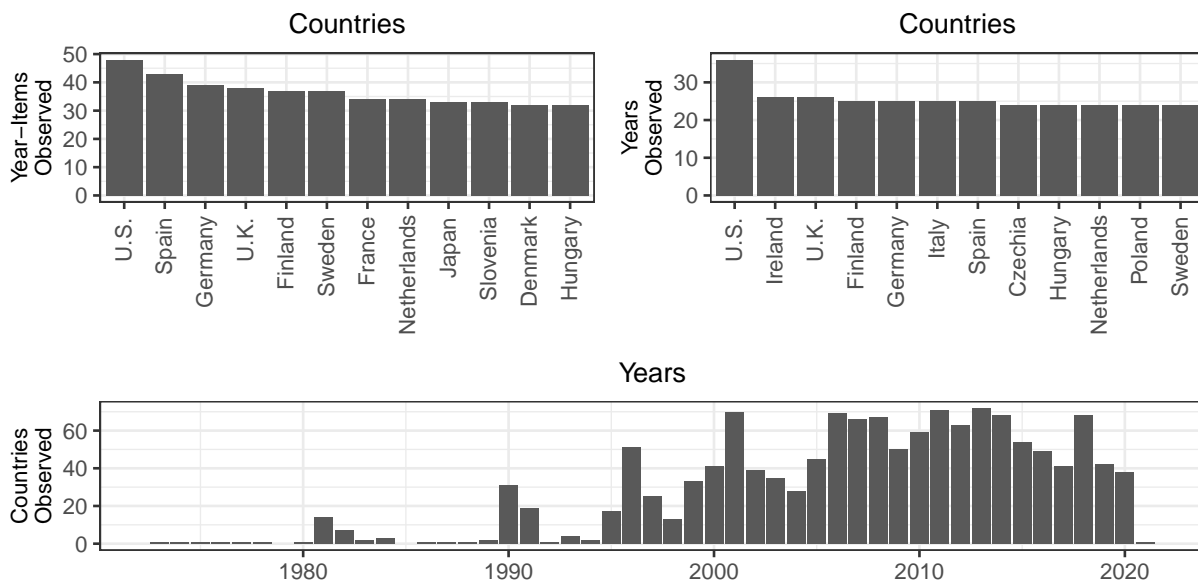


Figure 1: Countries and Years with the Most Observations in the TCS Source Data

¹The complete list of trust in civil servants/public administration survey items is included in online Appendix A.

Consider the most frequently asked item in the data we collected, which asks respondents whether they strongly agree, agree, disagree, or strongly disagree with the statement “I am going to name a number of institutions. For each one, could you tell me how much trust you have in them. Is it a great deal of trust, some trust, not very much trust or none at all? Civil service.”² Employed by the Arab Barometer, the Asia Europe Survey, the Asian Barometer, the British Social Attitudes Survey, the Latino Barometer, the East Asian Social Survey, the European Values Survey, the Italian National Election Study, the South Asian Barometer, and the World Values Survey, this question was asked in a total of 650 different country-years. That this constitutes only 23% of the country-years spanned by our data—and again, this is the *most common* survey item—again underscores just how sparse and incomparable the available public opinion data is on this topic.

The upper left panel of Figure 1 shows the dozen countries with the highest count of country-year-item observations. The United States, with 48 observations, is far and away the best represented country in the source data, followed by Spain, Germany, United Kingdom, and Finland. At the other end of the spectrum, 32 countries have only the minimum two observations required to be included in the source dataset at all. The upper right panel shows the twelve countries with the most years observed; this group is similar, but with Ireland, Italy, Czechia, and Poland joining the list and France, Japan, Slovenia, and Denmark dropping off. The bottom panel counts the countries observed in each year and reveals just how few relevant survey items were asked before 1990. Country coverage reached its peak in 2013, when respondents in 72 countries were asked items about trust in civil servants. In the next section, we describe how we are able to make use of all of this sparse and incomparable survey data to generate complete, comparable time-series TCS scores using a latent variable model.

Estimating trust in civil servants

A number of recent studies have developed latent variable models of public opinion based on cross-national survey data (see Claassen 2019; Caughey, O’Grady and Warshaw 2019; McGann, Dellepiane-Avellaneda and Bartle 2019; Kolczynska et al. 2020). To estimate trust in civil servants

²Question text may vary slightly across survey datasets, but not, roughly speaking, by more than the translation differences across languages found within the typical cross-national survey dataset. In this case, some questions ask about “the public administration” or “government officials” rather than “the civil service,” and some refer to “confidence” rather than “trust.” These words are often translated identically.

across countries and over time, we employ the latest of these methods that is appropriate for data that is not only incomparable but also sparse, the Dynamic Comparative Public Opinion (DCPO) model elaborated in Solt (2020b).³ The DCPO model is a population-level two-parameter ordinal logistic item response theory (IRT) model with country-specific item-bias terms. For a detailed description of the DCPO model, see Appendix B and Solt (2020b, 3-8); here, we focus on how it deals with the principal issues raised by our source data, incomparability and sparsity.

The DCPO model accounts for the incomparability of different survey questions with two parameters. First, it incorporates the *difficulty* of each question’s responses, that is, how much trust in civil servants is indicated by a given response. That each response evinces more or less of our latent trait is most easily seen with regard to the ordinal responses to the same question: strongly agreeing with the statement “you can generally trust the people who run our government to do what is right,” exhibits more trust in civil servants than responding “agree,” which in turn is more trusting than responding “disagree,” which is a more trusting response than “strongly disagree.” But this is also true across questions. For example, expressing “great trust” in civil servants “to look after your interests” likely expresses even more trust than just strongly agreeing that civil servants can be trusted to do what is right. Second, the DCPO model accounts for each question’s *dispersion*, its noisiness with regard to our latent trait. The lower a question’s dispersion, the better that changes in responses to the question map onto changes in trust of civil servants. Together, the model’s difficulty and dispersion estimates work to generate comparable estimates of the latent variable of trust in civil servants from the available but incomparable source data.

To address the sparsity of the source data—the fact that there are gaps in the time series of each country, and even many observed country-years have only one or few observed items—DCPO uses simple local-level dynamic linear models, i.e., random-walk priors, for each country. That is, within each country, each year’s value of trust in civil servants is modeled as the previous year’s estimate plus a random shock. These dynamic models smooth the estimates of trust in civil servants over time and allow estimation even in years for which little or no survey data is available, albeit at the

³Solt (2020b) demonstrates that the DCPO model provides a better fit to survey data than the models put forward by Claassen (2019) or Caughey, O’Grady and Warshaw (2019). The McGann, Dellepiane-Avellaneda and Bartle (2019) model depends on dense survey data unlike the sparse data on trust in civil servants described in the preceding section. Kolczynska et al. (2020) is the very most recent of these five works and builds on each of the others, but the MRP approach developed in that piece is suitable not only when the available survey data are dense but also when ancillary data on population characteristics are available, so it is similarly inappropriate to this application.

expense of greater measurement uncertainty.

We estimated the model using the `DCP0tools` package for R (Solt 2020a), running four chains for 1,000 iterations each and discarding the first half as warmup, which left us with 2,000 samples. The \hat{R} diagnostic had a maximum value of 1.01, indicating that the model converged.

The result is estimates, in all 2,775 country-years spanned by the source data, of mean trust in civil servants, what we call TCS scores. Figure 3 displays the most recent available TCS score for each of the 136 countries and territories in the dataset.

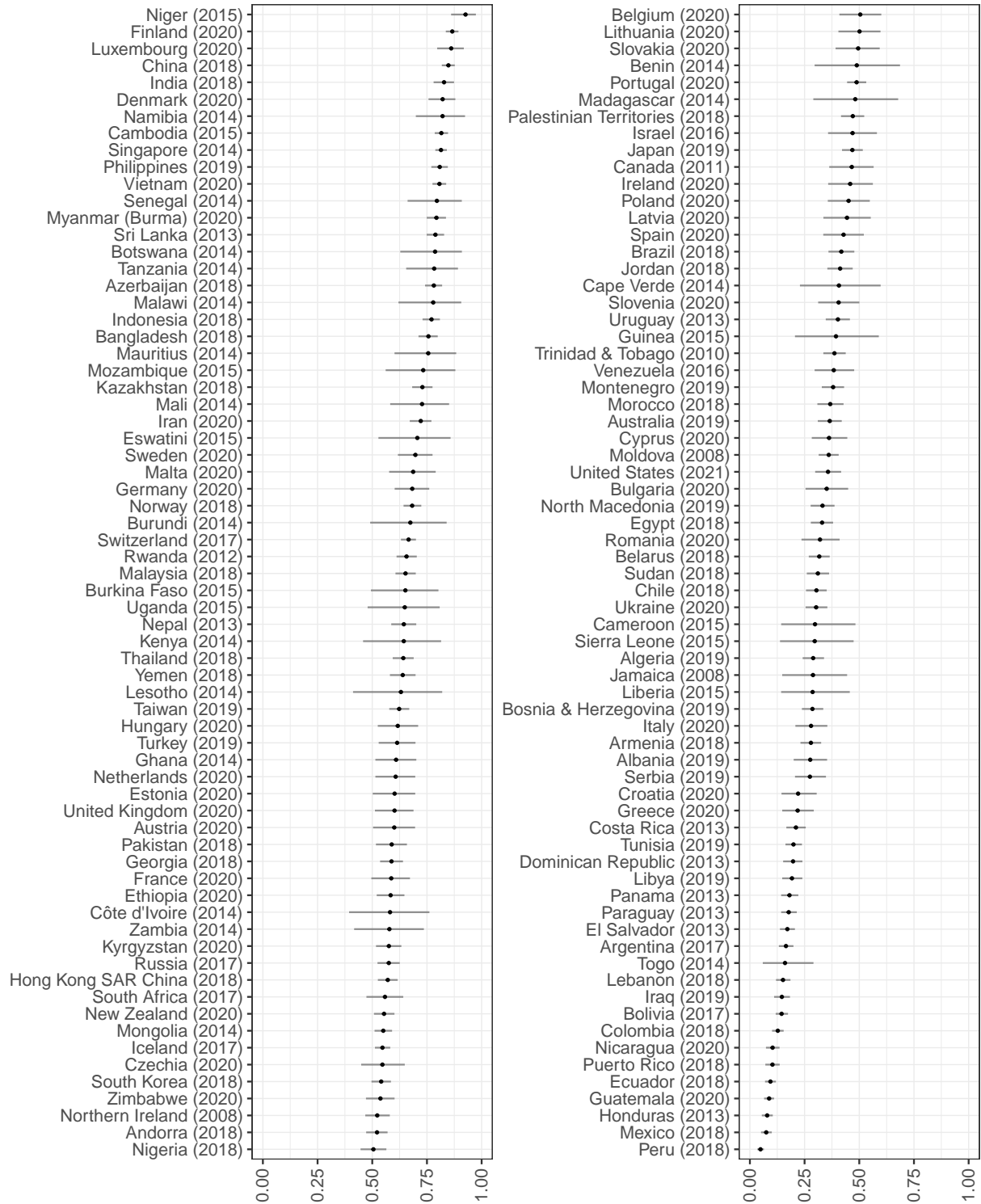
[The Scandinavian countries, Finland and Denmark which have high quality of governance, and Asian countries, China, India and Singapore which have a history of meritocracy, are at the top of the list.] The latest scores for Peru, Mexico, Honduras, Guatemala, and Ecuador have them as the places where the public has the lowest trust toward civil servants.

We show the changes of TCS over time in sixteen countries in Figure 4. As displayed in Figure 3, the dataset covers a wide range of geographic breadth, allowing the comparative studies on countries and regions too often neglected (see ?). Figure 4 also shows that while TCS has risen prominently in some countries, such as Finland and U.K., trusting attitudes have maintained high or low over time in others, like China and Turkey, or fallen steadily, as in U.s. and Mexico. They have advanced and retreated as in Spain or have declined and recovered as in Argentina . It is worthwhile to further explore the causes and consequences of these trends.

Validating Trust in Civil Servants

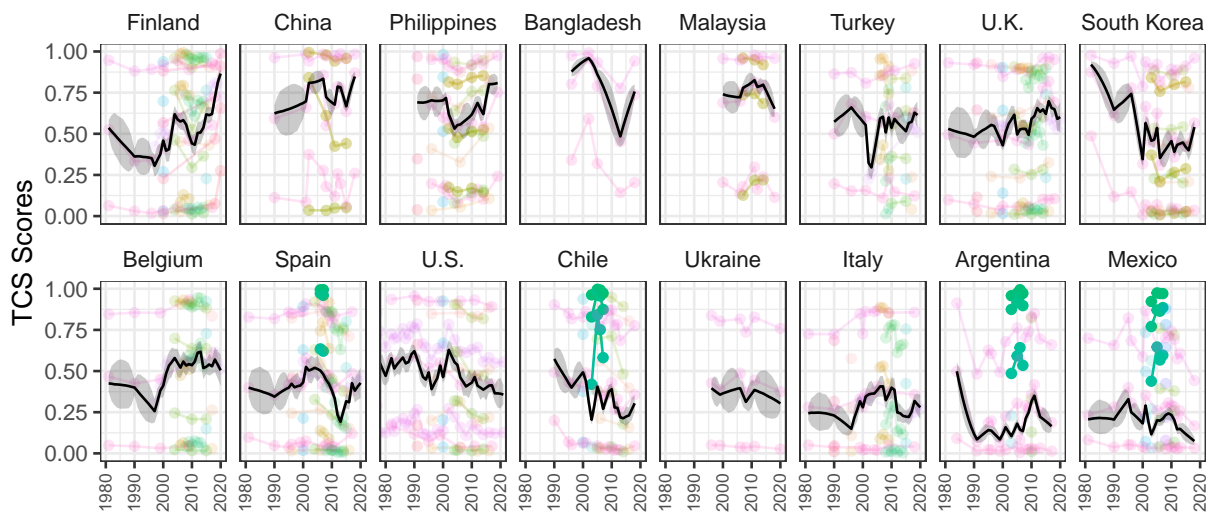
Convergence Validations

Construct Validations



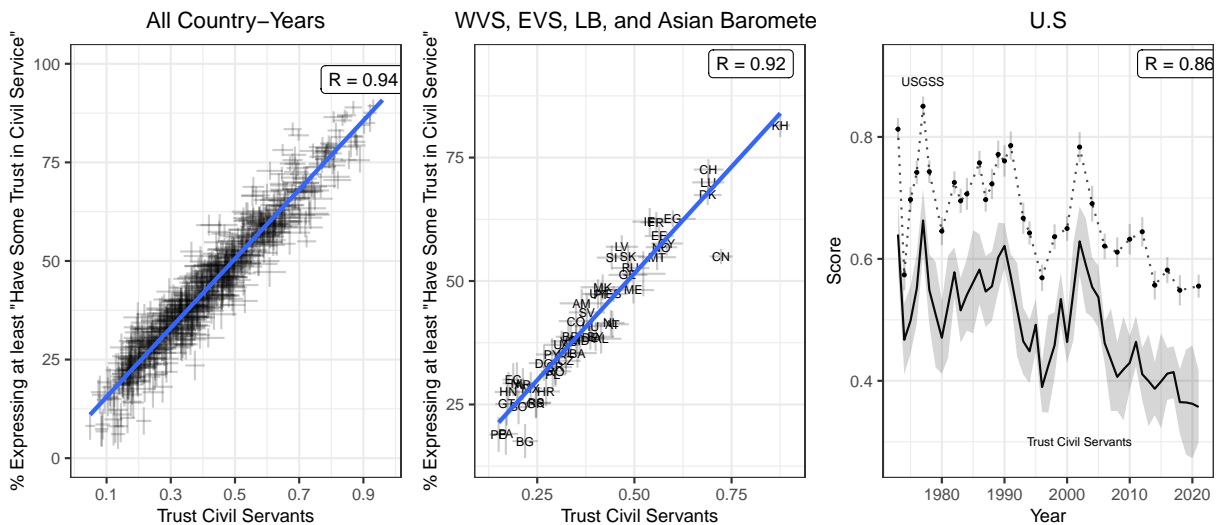
Note: Gray whiskers represent 80% credible intervals.

Figure 3: TCS Scores, Most Recent Available Year



Note: Countries are ordered by their TCS scores in their most recent available year; gray shading represents 80% credible intervals.

Figure 4: TCS Scores Over Time Within Selected Countries



Note: Gray whiskers and shading represent 80% credible intervals.

Figure 5: Internal Validation for Trust Civil Servants.

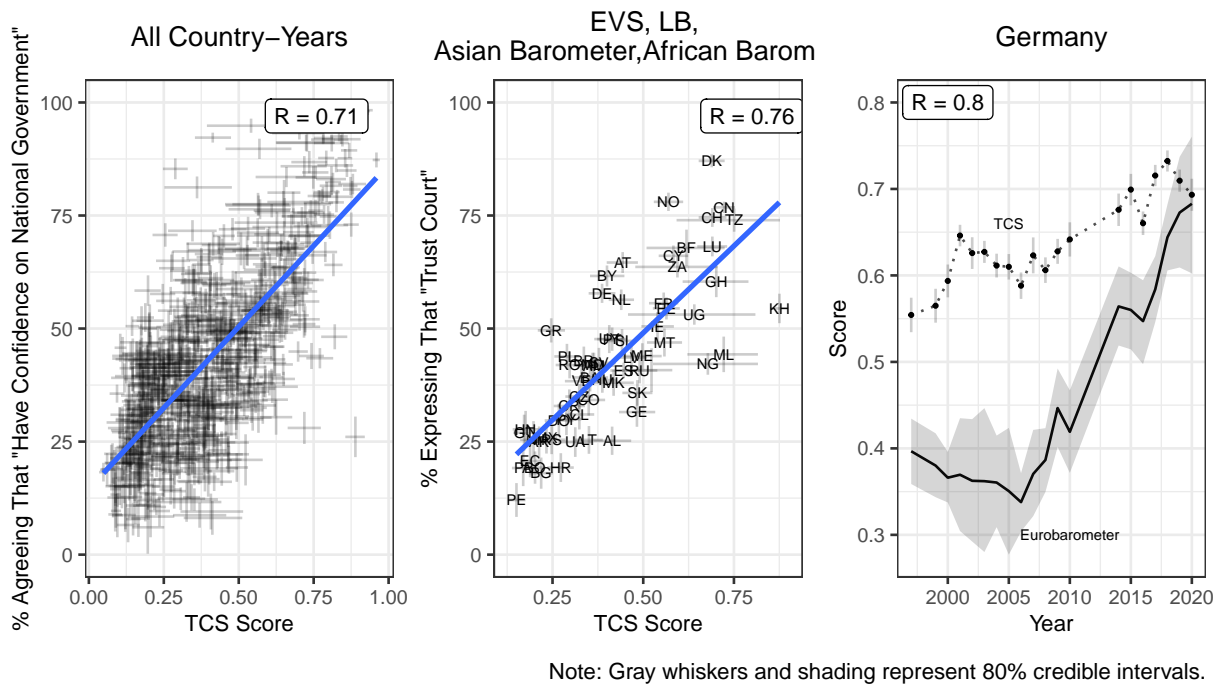


Figure 6: Construct Validation: Correlations Between TCS Scores and Individual Trust in Institutions Survey Items

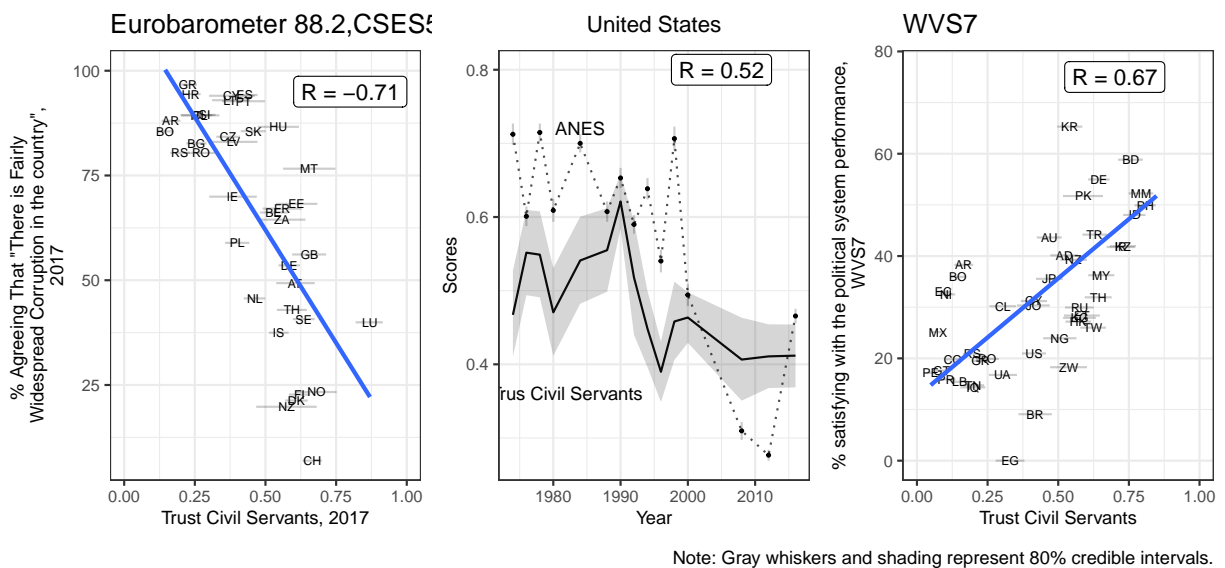


Figure 7: Construct Validation for Trust Civil Servants .

Appendix A: Survey Items Used to Estimate Trust in Civil Servants

National and cross-national surveys have often included questions tapping trusting attitudes over the past half-century, but the resulting data are both sparse, that is, unavailable for many countries and years, and incomparable, generated by many different survey items. In all, we identified 16 such survey items that were asked in no fewer than five country-years in countries surveyed at least twice; these items were drawn from 107 different survey datasets. These items are listed in the table below, along with the dispersion (α) and difficulty (β) scores estimated for each from the DCPO model. Question text may vary slightly across survey datasets, but not, roughly speaking, by more than the translation differences across languages found within the typical cross-national survey dataset. Lower values of dispersion indicate questions that better identify publics with a higher level of trust from those with lower. Items have one less difficulty score than the number of response categories. Survey dataset codes correspond to those used in the `DCPOtools` R package; they appear in decreasing order of country-years contributed.

Together, the survey items in the source data were asked in 136 different countries in at least two time points over 48 years, from 1973 to 2021, yielding a total of 1,768 country-year-item observations. The number of items observed in the source data for each country-year is plotted in Figure 2 below. The TCS scores of country-years with more observed items are likely to be estimated more precisely. The estimates for country-years with fewer (or no) observed items rely more heavily (or entirely) on the random-walk prior and are therefore less certain. Figure ?? displays the distribution of the source data in yet another way, showing the (binned) number of country-years in which both members of each pair of items is observed.

Appendix B: The DCPO Model

A number of recent studies have developed latent variable models of public opinion based on cross-national survey data (see Claassen 2019; Caughey, O’Grady and Warshaw 2019; McGann, Dellepiane-Avellaneda and Bartle 2019; Kolczynska et al. 2020). To estimate trust in civil servants across countries and over time, we employ the latest of these methods that is appropriate for data that is not only incomparable but also sparse, the Dynamic Comparative Public Opinion (DCPO)

model elaborated in Solt (2020b).⁴ The DCPO model is a population-level two-parameter ordinal logistic item response theory (IRT) model with country-specific item-bias terms.

DCPO models the total number of survey responses expressing at least as much trust in civil servants as response category r to each question q in country k at time t , y_{ktqr} , out of the total number of respondents surveyed, n_{ktqr} , using the beta-binomial distribution:

$$a_{ktqr} = \phi \eta_{ktqr} \quad (1)$$

$$b_{ktqr} = \phi(1 - \eta_{ktqr}) \quad (2)$$

$$y_{ktqr} \sim \text{BetaBinomial}(n_{ktqr}, a_{ktqr}, b_{ktqr}) \quad (3)$$

where ϕ represents an overall dispersion parameter to account for additional sources of survey error beyond sampling error and η_{ktqr} is the expected probability that a random person in country k at time t answers question q with a response at least as positive as response r .⁵

This expected probability, η_{ktqr} , is in turn estimated as follows:

$$\eta_{ktqr} = \text{logit}^{-1}\left(\frac{\bar{\theta}'_{kt} - (\beta_{qr} + \delta_{kq})}{\sqrt{\alpha_q^2 + (1.7 * \sigma_{kt})^2}}\right) \quad (4)$$

In this equation, β_{qr} represents the difficulty of response r to question q , that is, the degree of trust in civil servants the response expresses. The δ_{kq} term represents country-specific item bias: the extent to which all responses to a particular question q may be more (or less) difficult in a given country k due to translation issues, cultural differences in response styles, or other idiosyncrasies that render the same survey item not equivalent across countries.⁶ The dispersion of question q , its noisiness in relation to our latent variable, is α_q . The mean and standard deviation of the

⁴Solt (2020b) demonstrates that the DCPO model provides a better fit to survey data than the models put forward by Claassen (2019) or Caughey, O'Grady and Warshaw (2019). The McGann, Dellepiane-Avellaneda and Bartle (2019) model depends on dense survey data unlike the sparse data on trust in civil servants described in the preceding section. Kolczynska et al. (2020) is the very most recent of these five works and builds on each of the others, but the MRP approach developed in that piece is suitable not only when the available survey data are dense but also when ancillary data on population characteristics are available, so it is similarly inappropriate to this application.

⁵The ordinal responses to question q are coded to range from 1 (expressing the least trust in civil servants) to R (expressing the most trust in civil servants), and r takes on all values greater than 1 and less than or equal to R .

⁶Estimating δ_{kq} requires repeated administrations of question q in country k , so when responses to question q are observed in country k in only a single year, the DCPO model sets δ_{kq} to zero by assumption, increasing the error of the model by any country-item bias that is present. Questions that are asked repeatedly over time in only a single country pose no risk of country-specific item bias, so δ_{kq} in such cases are also set to zero.

unbounded latent trait of trust in civil servants are $\bar{\theta}'_{kt}$ and σ_{kt} , respectively.

Random-walk priors are used to account for the dynamics in $\bar{\theta}'_{kt}$ and σ_{kt} , and weakly informative priors are placed on the other parameters.⁷ The dispersion parameters α_q are constrained to be positive and all survey responses are coded with high values indicating more trust in civil servants to fix direction. The difficulty β of “disagree” (on the four-point, “strongly agree” to “strongly disagree” scale) to the statement “On the whole, men make better political leaders than women do” is set to 1 to identify location, and for each question q the difficulties for increasing response categories r are constrained to be increasing. The sum of δ_{kq} across all countries k is set to zero for each question q :

$$\sum_{k=1}^K \delta_{kq} = 0 \quad (5)$$

Finally, the logistic function is used to transform $\bar{\theta}'_{kt}$ to the unit interval and so give the bounded mean of latent trust in civil servants, $\bar{\theta}_{kt}$, which is our parameter of interest here (see Solt 2020b, 3-8).

The DCPO model accounts for the incomparability of different survey questions with two parameters. First, it incorporates the *difficulty* of each question’s responses, that is, how much trust in civil servants is indicated by a given response. That each response evinces more or less of our latent trait is most easily seen with regard to the ordinal responses to the same question: strongly agreeing with the statement “both the husband and wife should contribute to household income,” exhibits more trust in civil servants than responding “agree,” which in turn is more egalitarian than responding “disagree,” which is a more egalitarian response than “strongly disagree.” But this is also true across questions. For example, strongly disagreeing that “on the whole, men make better business executives than women do” likely expresses even more egalitarianism than strongly agreeing merely that both spouses should have paying jobs. Second, the DCPO model accounts

⁷The dispersion parameters α_q are drawn from standard half-normal prior distributions, that is, the positive half of $N(0, 1)$. The first difficulty parameters for each question, β_{q1} , are drawn from standard normal prior distributions, and the differences between β s for each r for the same question q are drawn from standard half-normal prior distributions. The item-bias parameters δ_{kq} receive normally-distributed hierarchical priors with mean 0 and standard deviations drawn from standard half-normal prior distributions. The initial value of the mean unbounded latent trait for each country, $\bar{\theta}'_{k1}$, is assigned a standard normal prior, as are the transition variances $\sigma_{\bar{\theta}'}^2$ and σ_{σ}^2 ; the initial value of the standard deviation of the unbounded latent trait for each country, σ_{k1} , is drawn from a standard lognormal prior distribution. The overall dispersion, ϕ , receives a somewhat more informative prior drawn from a gamma(4, 0.1) distribution that yields values that are well scaled for that parameter.

for each question’s *dispersion*, its noisiness with regard to our latent trait. The lower a question’s dispersion, the better that changes in responses to the question map onto changes in trust in civil servants. Together, the model’s difficulty and dispersion estimates work to generate comparable estimates of the latent variable of trust in civil servants from the available but incomparable source data.

To address the sparsity of the source data—the fact that there are gaps in the time series of each country, and even many observed country-years have only one or few observed items—DCPO uses simple local-level dynamic linear models, i.e., random-walk priors, for each country. That is, within each country, each year’s value of trust in civil servants is modeled as the previous year’s estimate plus a random shock. These dynamic models smooth the estimates of trust in civil servants over time and allow estimation even in years for which little or no survey data is available, albeit at the expense of greater measurement uncertainty.

References

- Bouckaert, Geert. 2012. "Trust and public administration." *Administration* 60(1):91–115.
- Caughey, Devin, Tom O'Grady and Christopher Warshaw. 2019. "Policy Ideology in European Mass Publics, 1981–2016." *American Political Science Review* 113(3):674–693.
- Chanley, Virginia A, Thomas J Rudolph and Wendy M Rahn. 2000. "The origins and consequences of public trust in government: A time series analysis." *Public opinion quarterly* 64(3):239–256.
- Choi, SeoYoun. 2018. "Bureaucratic characteristics and citizen trust in civil service in OECD member nations." *International Area Studies Review* 21(2):114–133.
- Claassen, Christopher. 2019. "Estimating Smooth Country–Year Panels of Public Opinion." *Political Analysis* 27(1):1–20.
- Dahl, Robert A. 1947. "The Science of Public Administration: Three Problems." *Public Administration Review* 7(1):1–11.
- Easton, David. 1975. "A re-assessment of the concept of political support." *British journal of political science* 5(4):435–457.
- Goldstein, Daniel AN and Johannes Wiedemann. 2021. "Who do you trust? The consequences of partisanship and trust for public responsiveness to COVID-19 orders." *Perspectives on Politics* pp. 1–27.
- Houston, David J, Nurgul R Aitalieva, Andrew L Morelock and Chris A Shults. 2016. "Citizen trust in civil servants: A cross-national examination." *International Journal of Public Administration* 39(14):1203–1214.
- Jilke, Sebastian, Bart Meuleman and Steven Van de Walle. 2015. "We need to compare, but how? Measurement equivalence in comparative public administration." *Public Administration Review* 75(1):36–48.
- Kettl, Donald F. 2000. *The global public management revolution: A report on the transformation of governance*. Brookings Institution Press.
- Kim, Seok-Eun. 2005. "The role of trust in the modern administrative state: An integrative model." *Administration & Society* 37(5):611–635.
- Kolczynska, Marta, Paul-Christian Bürkner, Lauren Kennedy and Aki Vehtari. 2020. "Trust in State Institutions in Europe, 1989-2019." SocArXiv. <https://osf.io/preprints/socarxiv/3v5g7/>.
- McGann, Anthony, Sabastian Dellepiane-Avellaneda and John Bartle. 2019. "Parallel Lines? Policy Mood in a Plurinational Democracy." *Electoral Studies* 58:48–57.
- Morelock, Andrew L. 2021. "In Bureaucrats We Trust? Good Governance and Trust in Civil Servants." *Public Administration Quarterly* 45(3):315–337.
- Pollitt, Christopher. 2011. "Not odious but onerous: Comparative public administration." *Public Administration* 89(1):114–127.

- Rogowski, Jon C and Andrew R Stone. 2021. “How political contestation over judicial nominations polarizes Americans’ attitudes toward the Supreme court.” *British Journal of Political Science* 51(3):1251–1269.
- Solt, Frederick. 2020*a*. “DCPO: Dynamic Comparative Public Opinion.” Available at the Comprehensive R Archive Network (CRAN). <https://CRAN.R-project.org/package=DCPO>.
- Solt, Frederick. 2020*b*. “Modeling Dynamic Comparative Public Opinion.” SocArXiv. <https://osf.io/preprints/socarxiv/d5n9p>.
- Thomas, Craig W. 1998. “Maintaining and restoring public trust in government agencies and their employees.” *Administration & society* 30(2):166–193.
- Van de Walle, Steven and Koen Migchelbrink. 2022. “Institutional quality, corruption, and impartiality: The role of process and outcome for citizen trust in public administration in 173 European regions.” *Journal of Economic Policy Reform* 25(1):9–27.
- Van Ryzin, Gregg G. 2011. “Outcomes, process, and trust of civil servants.” *Journal of Public Administration Research and Theory* 21(4):745–760.
- Yates, Douglas. 1982. *Bureaucratic democracy: The search for democracy and efficiency in American government*. Harvard University Press.