Trust in Bureaucracy: Public Opinion on Public Servants in Dynamic Comparative Perspective

October 06, 2024

Authors

- Yuehong Cassandra Tai, ORCID: https://orcid.org/0000-0001-7303-7443, Postdoctoral Fellow, Center for Social Data Analytics, Pennsylvania State University, yhcasstai@psu.edu
- Frederick Solt, corresponding author, ORCID: https://orcid.org/0000-0002-3154-6132, Associate Professor, Department of Political Science, University of Iowa, frederick-solt@uiowa.edu

Trust in Bureaucracy: Public Opinion on Public Servants in Dynamic Comparative Perspective

October 06, 2024

Abstract

blah blah blah

[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, 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 posts 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 restricted to a small number of country-level unites, 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 incaparability of trust measures cross countries can produce misleading conclusions. For example, without accounting for item bias, Swedens 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 compliance 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, 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 the quality of governance (Bouckaert 2012; Kettl 2000; Van de Walle and Migchelbrink 2022; Morelock 2021).

In this research, 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 apply our TCS data in a cross-national-time-series data to examine the competing theories on explaining trust in bureaucracy. 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 17 such survey items that were asked in no fewer than five country-years in countries surveyed at least twice; these items were drawn from 132 different survey datasets.[1]

Together, the survey items in the source data were asked in 98 different countries in at least two time points over 36 years, from 1973 to 2022, yielding a total of 1,814 country-year-item observations. Observations for every year in each country surveyed would number 3,528, and a complete set of country-year-items would encompass 59,976 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 there are 1,344 country-years in which we have at least *some* information about the trust in civil servants of the population, that is, some 54% of the 2,475 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.

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

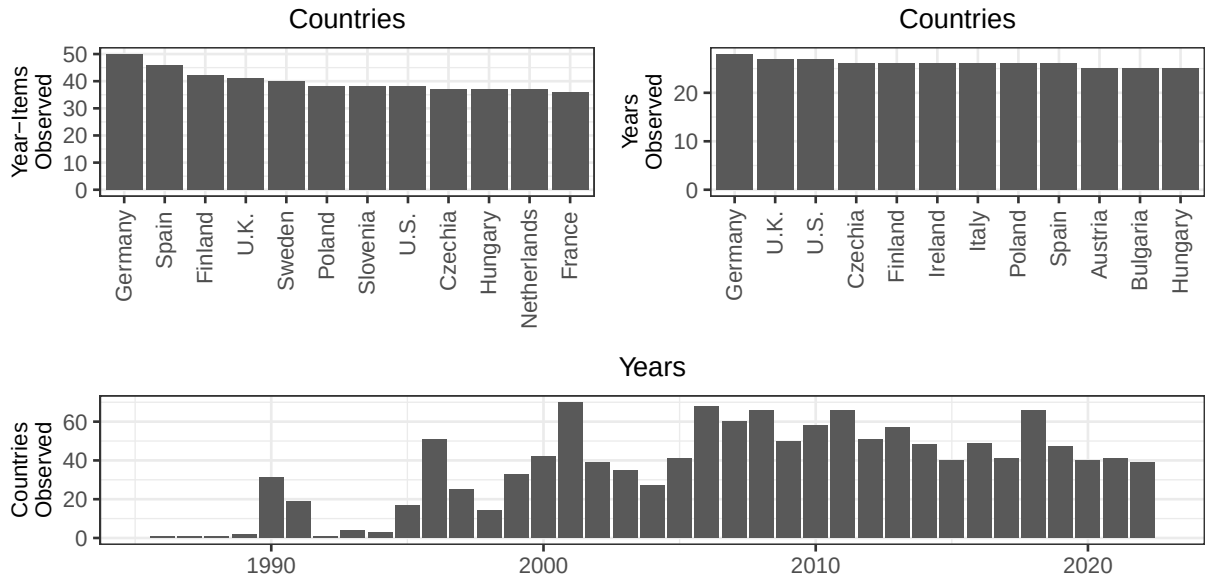

Figure 1: Countries and Years with the Most Observations in the Source Data

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 614 different country-years. That this constitutes only 25% 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 38 observations, is far and away

[2]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. 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.

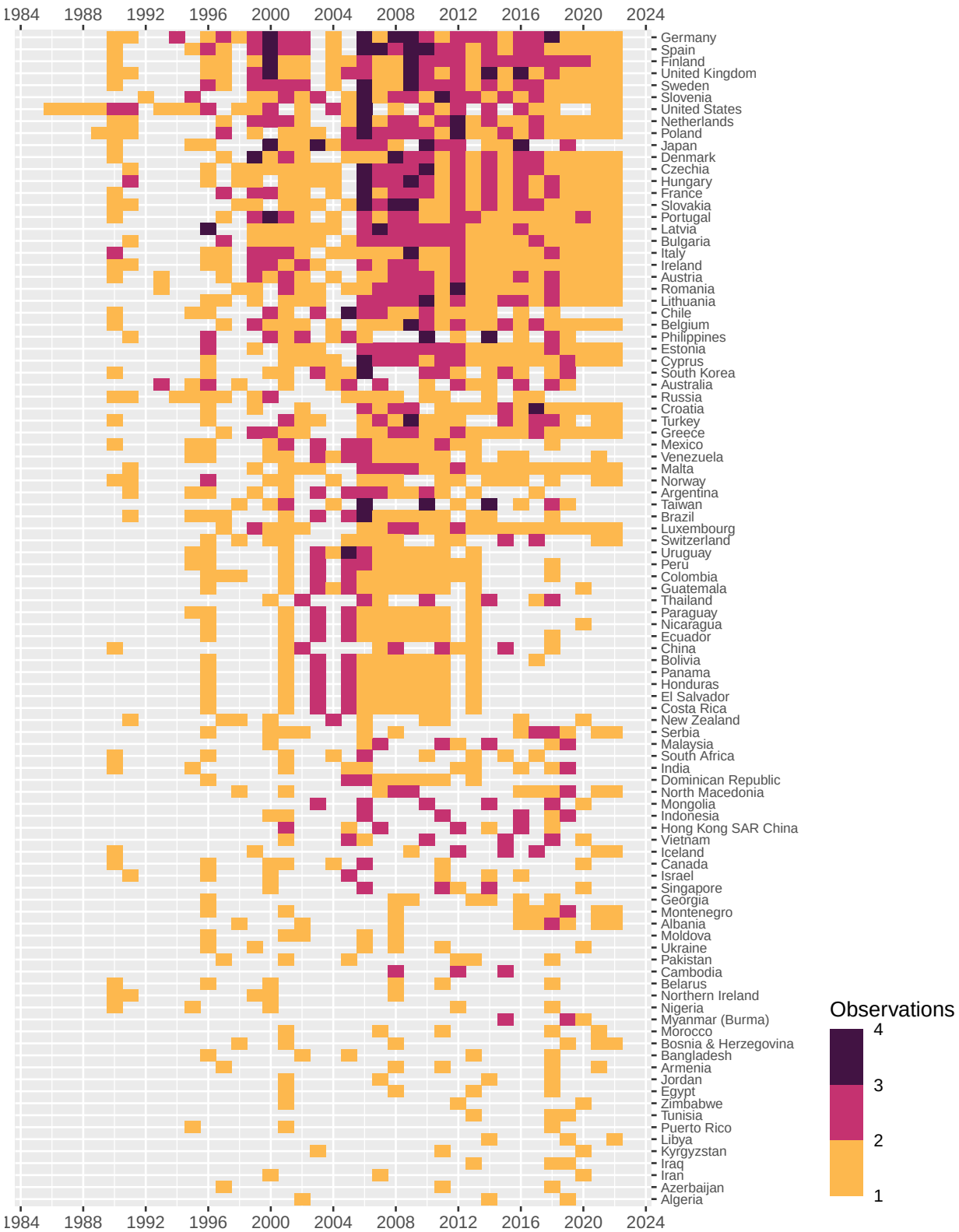

Figure 2: Source Data Observations by Country and Year

the best represented country in the source data, followed by Spain, Finland, United Kingdom, and Sweden. At the other end of the spectrum, one 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, Austria, and Bulgaria joining the list and Sweden, Slovenia, Netherlands, and France 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 2001, when respondents in 70 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

Several latent-variable models of public opinion based on cross-national survey data have been developed recently (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 (2020c).[3] The DCPO model is a population-level two-parameter ordinal logistic item response theory (IRT) model with country-specific item-bias terms. For more information on the DCPO model, see Appendix B and Solt (2020c, 3–8); in this section, 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

[3]The DCPO model provides a better fit to survey data than the models proposed in Claassen (2019) or Caughey, O'Grady, and Warshaw (2019; Solt 2020c). The model put forward in McGann, Dellepiane-Avellaneda, and Bartle (2019) depends on dense survey data unlike the sparse data on trust in civil servants just described. Building on all of these four works, Kolczynska et al. (2020) is the very most recent effort, but the multilevel regression and post-stratification (MRP) approach it offers depends both on dense survey data and on additional data describing population characteristics, so it too is inappropriate for our purposes here.

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 reveals more or less of our latent trait is easily seen with regard to 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 simply agreeing, which shows more trust than responding "disagree," which in turn is a more trusting response than "strongly disagree." But this is likely to also be 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 the time series for each country has gaps, and many country-years that are observed have available only one item—DCPO uses 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. The random walk priors smooth the estimates of trust in civil servants over time and allow—at the cost of greater measurement uncertainty—estimates to be generated even in years for which little or no data is available.

We estimated the model using the `DCPOtools` 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 dispersion parameters of the survey items indicate that all of our source data items load well on the latent variable (see Appendix A). The result is estimates, in all 2,475 country-years spanned by the source data, of the citizenry's aggregate trust in civil servants, what we call TCS scores. Figure 3 displays the most recent available TCS score

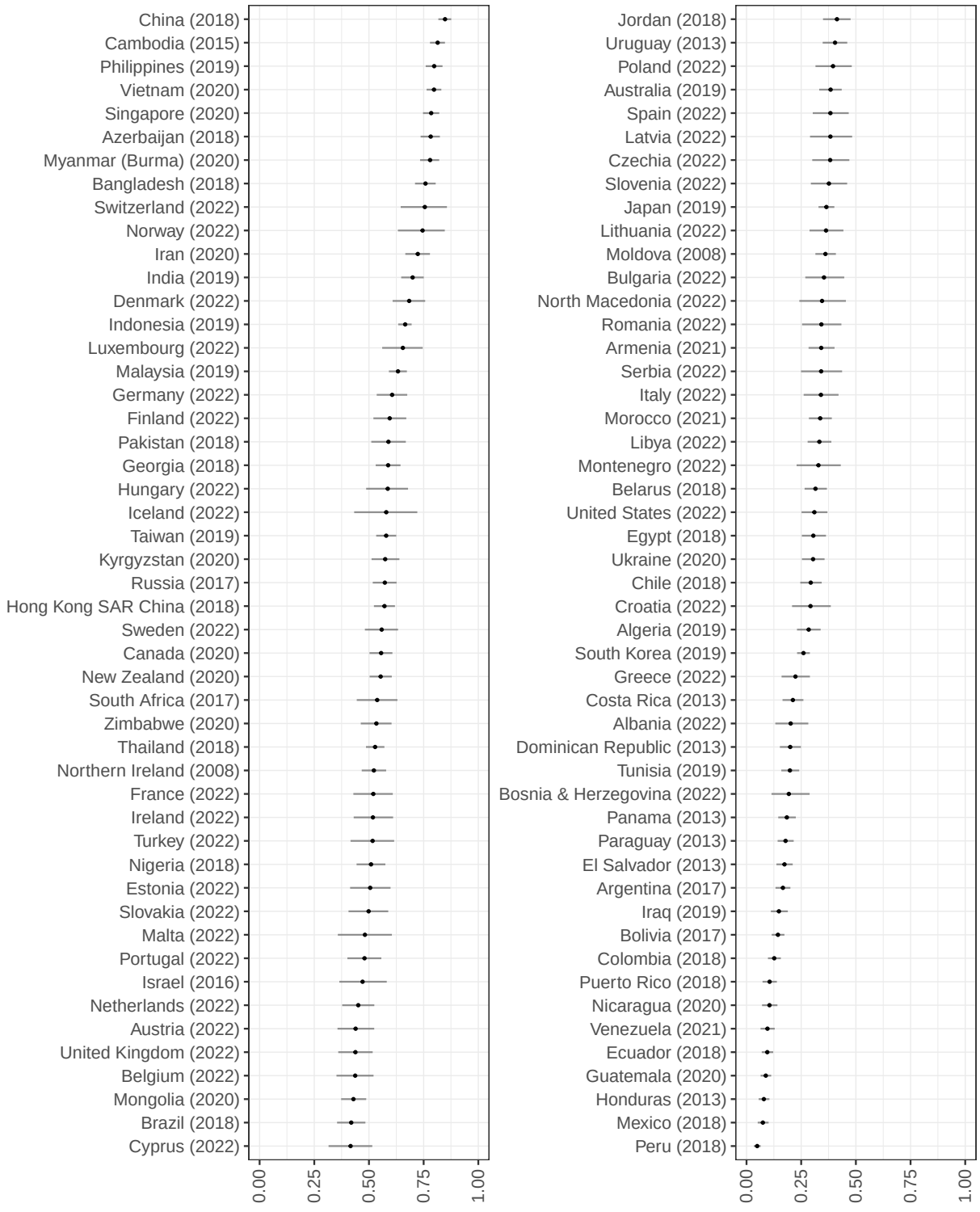

Figure 3: TCS Scores, Most Recent Available Year

for each of the 98 countries and territories in the dataset.

Asian countries, especially those with a history of meritocracy, dominate the top of the list. The least corrupt counties, like Switzerland, Norway, Denmark and Finland, also rank highly. On the other hand, 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. The bottom-ranked countries are either among the most corrupt, like Venezuela, or have high crime rates, like Peru and Honduras.

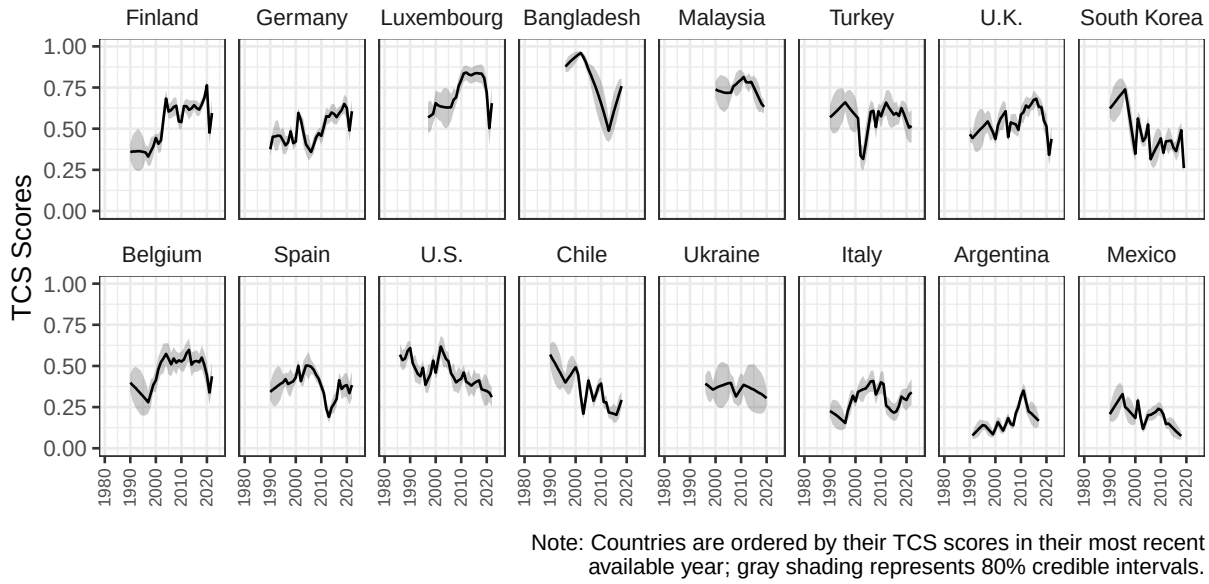

Figure 4: TCS Scores Over Time Within Selected Countries

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 comparative studies of countries and regions too often neglected (see Wilson and Knutsen 2020). Figure 4 also shows that while TCS has risen prominently in some countries, such as Finland and Germany, trusting attitudes have maintained high or low over time in others, like Luxembourg and Mexico, or fallen steadily, as in U.S. and Chile. They have advanced and retreated as in Spain or have declined and recovered as in Turkey. The differences within countries over time and differences across countries present a challenge to explain the causes and consequences of trust in bureaucracies.

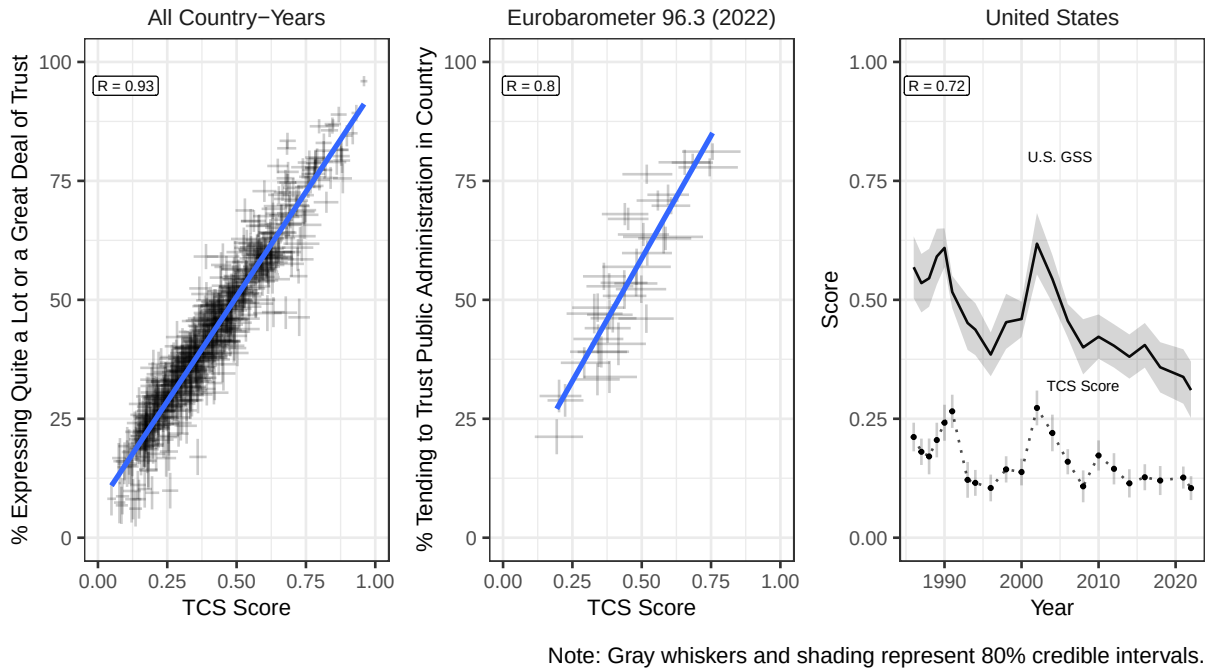

Figure 5: Convergent Validation: Correlations Between TSC Scores and Individual TSC Source-Data Survey Items

Before using these estimates in analysis, we validate our trust civil service score through convergent validation and construct validation, since validation tests of cross-national latent variables are crucially important (see, e.g., Hu et al. 2023). Figure 5 shows the measure's validity in tests of convergent validation that tests whether a measure is empirically associated with alternative indicators of the same concept (Adcock and Collier 2001, 540). We started with 'internal' convergent validation test (see, e.g., Caughey, O'Grady, and Warshaw 2019, 689; Solt 2020c, 10) by comparing our TCS score with individual items from source-data to generate them.

The left panel in Figure 5 shows a scatterplot of country-years in which the TCS scores are plotted against the percentage of respondents who expressed "a quite a lot"or "a great deal" of trust in response to the question: "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."

The middle panel plots our TCS score against the percentage who responded "Tend to trust." to the question, "I would like to ask you a question about how much trust you have in certain institutions. For each of the following institutions, please tell me if you tend to trust it or tend not to trust it: Public administration in (OUR COUNTRY)" in the Eurobarometer 96.3 January-February 2022 module. This question is asked in the most countries compared to other single cross-national surveys.

Finally, the right panel compares the trend of the longest item that has been asked since 1973 in U.S.General Social Survey, "I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them? Executive branch of the federal government." to the trend of the TCS scores. The TCS scores effectively reflect changes of trust in executive branch over time.

All three demonstrations show strong correlations between TCS and individual items, estimated taking into account the uncertainty in the measures.

Figure 6 present three 'external' convergent validation tests, comparing TCS scores to responses to survey items that were *not* included in the source data: items that asked respondents' confidence and trust in national government, parliament, and judiciary in their countries. There is a longstanding debate about the dimentionality of political trust Y. C. Tai (2022). Although trust in civil servants has been theoretically classified into the same dimentionality with either all these three types of trustHooghe (2011), or one of them Y. C. Tai (2022), our validations empirically show the importance of taking trust in civil servants separately an seriously.

In the left panel, we plot TCS score against data from seven rounds of World Value Survey, which asked respondents how much they trust their national government. The center plot shows data from European Values Surveys asking respondents' confidence in parliament. The right presents the percentage of respondents who expressed at least some trust in judiciary in their country in latinobarometro. Our measure positively correlated with all of them, with

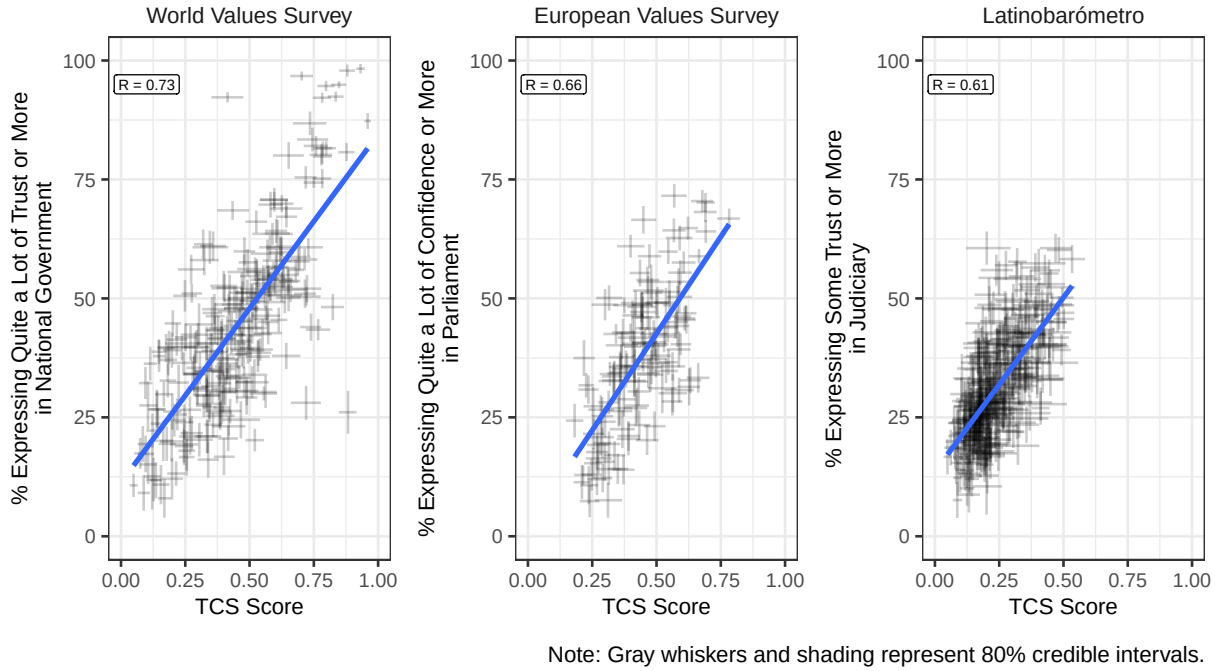

Figure 6: Construct Validation: Correlations Between TCS Scores and Trust in Institutions Survey Items

a stronger correlation with trust in national government and mild correlation with trust in parliament and judiciary.

The general strong relationship between TCS measures and institutional trust confirms the theoretical expectation about institutional trust, but the variation in TCS' relationship with trust in different institutions cross regions is worthwhile to further investigate to refine the dimensions in political trust.

We then conduct construct validation in Figure 7.

Construct validation assesses whether a given indicator is empirically correlated with other indicators in a way that conforms to theoretical expectations (Adcock and Collier 2001, 542). Corruption is often often argued as a likely contributor to trust in civil servants and administration Van de Walle and Migchelbrink (2022). The left panel compares perceived widespread of corruption, measured as the percentage of those saying most or state authorities are involved in corruption in seven WVS surveys, with the TCS scores. As theoretical

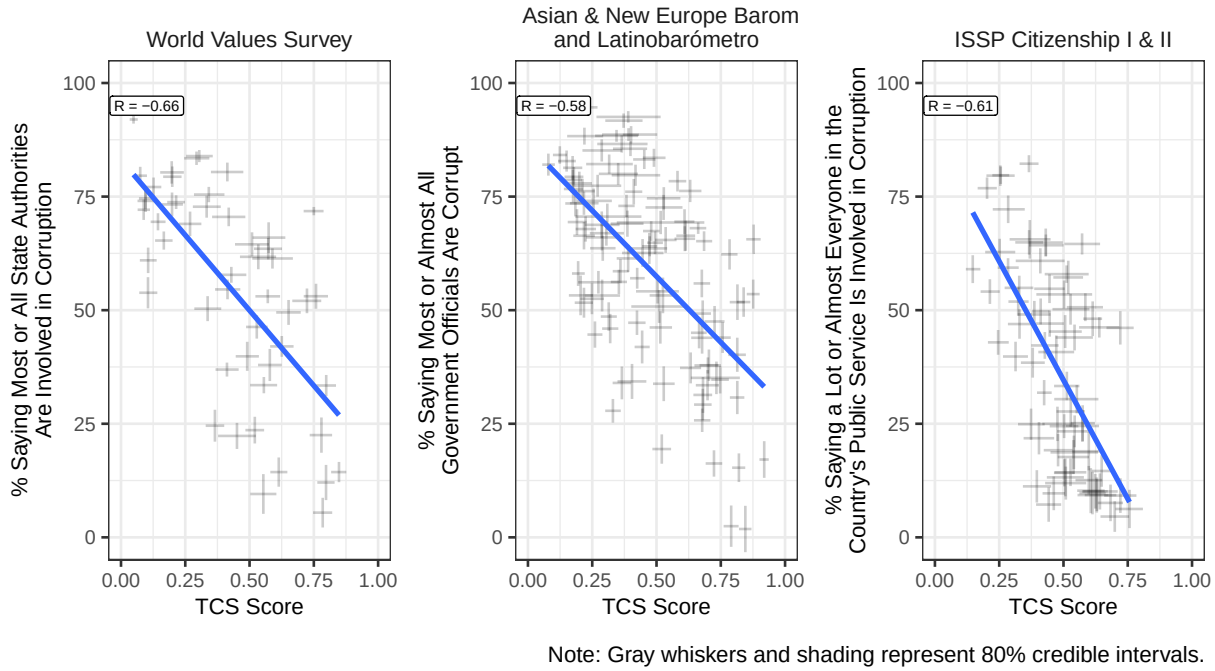

Figure 7: Construct Validation: Correlations Between TCS Scores and Corruption of Public Servants Survey Items

expectation, there is a clear negative relationship between the spread of perceived corruption and TCS: when there is widespread perception of corruption in authorities, the public tends to distrust civil servants. The center and right panel of Figure 7 also present negative relationships between TCS scores and the spread of perceived corruption of government officials cross different geographic regions. The center panel shows the data in developing or new democratic countries surveyed in Asian, New Europe Barometer, and Latinobarometro, and the right panel displays the data in countries surveyed in The International Social Survey Programme Citizenship module (2004, 2014).

To sum up, the evidence of construct validation of TCS scores agains the perceived widespread of corruption in Figure 7, together with the evidence of external validation in Figure 6 and convergent validation in Figure 5, demonstrates the validity of the TCS scores as measures of the public's trust in civil servants.

A longstanding puzzle in public administration is understanding what explains trust in bureaucracies. A dominant theme is the belief that higher levels of government performance lead to greater trust in civil servants, based on the assumption that better performance correlates with higher trust and that lower trust toward bureaucrats reflects dissatisfaction with government performance (Yang and Holzer 2006). One common approach to measuring performance is through macroeconomic outcomes, such as economic growth, unemployment rate, economic inequality, and inflation. However, the results from studies on macroeconomic outcomes are mixed. For example, Choi (2018) found that GDP per capita positively affects trust in bureaucracies, while Houston et al. (2016) did not find significant effects of GDP per capita and inflation rate on trust in civil servants. Instead, Houston et al. (2016) found that the unemployment rate negatively influences trust in civil servants. Contrary to previous studies that found some evidence for the role of government outcomes, Morelock (2021) found that none of the outcome indicators, including GDP per capita, inflation rate, unemployment, and the GINI index, had a significant effect on trust in civil servants.

Amidst these mixed results regarding macroeconomic outcomes, a growing body of literature emphasizes the role of government quality—or process—in explaining trust in bureaucracies. Van Ryzin (2011) found that the quality of government, measured by the World Bank's Worldwide Governance Indicators, plays a more crucial role than government outcomes measured by the UN's Human Development Index, which had a negative effect in his model. Morelock (2021)'s study also highlights the positive role of government effectiveness, although Houston et al. (2016)'s study find inconsistent role of government effectiveness. A relatively consistent finding across studies is the significant role of corruption. Van de Walle and Migchelbrink (2022) concluded that the perceived absence of corruption is more impactful on trust in bureaucracies than performance evaluations. The critical influence of perceived corruption and corruption control on public trust in civil servants is also supported by Houston et al. (2016) and Morelock (2021). However, since authors used different measures and modeling strategies, such as whether to include both outcomes and quality indicators in one model, it remains unclear how consistent these results would be if a uniform

modeling strategy were used.

Given these mixed results from studies on outcomes and quality in relation to trust, researchers acknowledge common limitations: limited country coverage, reliance on cross-sectional rather than dynamic analysis, and difficulties in obtaining comparable measures across countries or regions Van de Walle and Migchelbrink (2022). These limitations in comparative data and the lack of time-series data hinder a deeper understanding of the relationship between government outcomes, quality, and trust in bureaucracies.

With our time-series cross-national data on trust in bureaucracies, we combine both outcome and quality indicators to examine the factors influencing this trust. For outcome indicators, we followed previous studies and used GDP per capita, inflation, and unemployment from 1984 to 2022 as measures of macroeconomic performance. GDP per capita and inflation data were sourced from the International Monetary Fund, while unemployment data were collected from the World Bank, which uses modeled International Labour Organization estimates. To measure income inequality, we relied on Solt (2020b)'s Standardized World Income Inequality Database (SWIID), specifically using the index of inequality in disposable income for its superior coverage and comparability.

For quality of government indicators, we used the Corruption Perceptions Index from Transparency International, covering the years 1995 to 2022, to capture the perceived level of corruption. We also focused on the Government Effectiveness indicator from the World Bank's Worldwide Governance Indicators, as it encompasses the overall quality of public services, the civil service, and policy formulation and implementation.

To further leverage our trust data, we collected the number of intentional homicides at the country-year level from the United Nations Office on Drugs and Crime, capturing outcomes in public safety, given that macroeconomic outcomes do not represent government performance in other fields Morelock (2021). Additionally, to account for the effect of democratic development on trust, we included the Liberal Democracy Index from the V-Dem dataset by Coppedge et al. (2023) and Pemstein et al. (2023).

We adopted a Bayesian multilevel model with varying intercepts for each country and each

year. The varying intercepts for each country account for heteroskedasticity across countries, while those for each year account for 'time shocks' that impact all countries simultaneously (Shor et al. 2007). To capture both short-term and historical effects of predictors, we employed the 'within-between random effects' specification. This approach incorporates the time-invariant country mean and the time-varying difference between each country-year value and this country mean, following Bell and Jones (2015) and Woo et al. (2023). Measurement uncertainty in the data for trust in bureaucracies, income inequality, and the Corruption Perceptions Index was also incorporated into the analysis (see Y. 'Cassandra'. Tai, Hu, and Solt 2024). The model was estimated using the brms R package (Bürkner 2017).

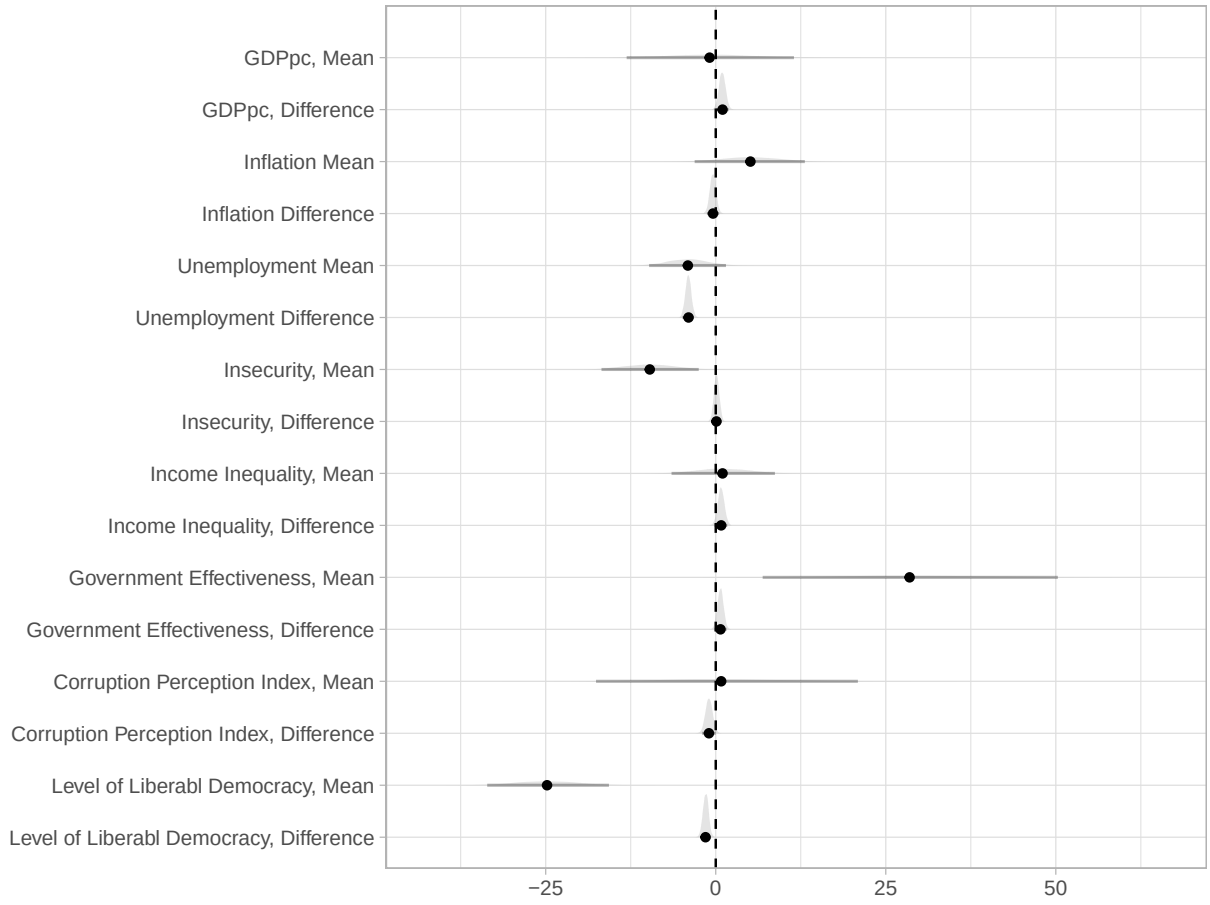

Figure 8: Predicting Trust in Bueacracies

The results are presented in Figure 8. In terms of economic outcomes, the increase in GDP per capita is associated with a higher level of trust in civil servants in the short term. A two-standard-deviation year-to-year change increases trust by 1 (95% c.i.: 0 to 1.9) point (95% c.i.:0 to 1.9). Similarly, we find that a two-standard-deviation year-to-year increase in unemployment rate in a short run decreases trust in civil servants by 4 (95% c.i.: -4.9 to -3.2) point (95% c.i.:-4.9 to -3.2). Regarding outcomes in public safety, government's performance in public security measured by the number of intentional homicides at the country-year level has a "historical", long-run negative effect on trust in civil servants. A two-standard deviation increase in a country's mean public insecurity is associated with 9.7 (95% c.i.: -16.8 to -2.5) points less trust (95% credible interval: -16.8 to -2.5 points) across all countries. However, we do not find evidence for the either "historical" or short-run effect of inflation and income inequality on trust in civil servants.

With regards of government quality-or process, a higher government effectiveness score was found to have a strong positive long-term effect on trust in civil servants. A two-standard deviation increase in a country's mean effectiveness score is associated with 28.5 (95% c.i.: 6.9 to 50.3) points more trust (95% credible interval: 6.9 to 50.3 points) across all countries. The development of democracy is associated with less trust in civil servants, both in the long run and in the short term. Critical citizens in more democratic countries could trust civil servant critically and have a higher demand for them (Norris 1999). However, perceived Corruption measured by Corruption perception index was not associated with trust in civil servants across countries at 95% confident interval at the both long-term and short-run and estimates are close to zero.

We find evidence for both arguments on government outcomes and government quality. Among predictors, compared to economic and public security outcomes, such as year-to-year GDP per capita, unemployment and countries' mean insecurity, government quality or process, measured by countries' mean government effectiveness and mean democratic development, has more substantive effects in long run.

Conclusion

References

Adcock, Robert, and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95 (3): 529–46.

Anderson, Christopher J, and Yuliya V Tverdova. 2003. "Corruption, Political Allegiances, and Attitudes Toward Government in Contemporary Democracies." *American Journal of Political Science* 47 (1): 91–109.

Bell, Andrew, and Kelvyn Jones. 2015. "Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data." *Political Science Research and Methods* 3 (1): 133–53.

Bouckaert, Geert. 2012. "Trust and Public Administration." *Administration* 60 (1): 91–115.

Bürkner, Paul-Christian. 2017. "brms: An R Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software* 80: 1–28.

Caughey, Devin, Tom O'Grady, and Christopher Warshaw. 2019. "Policy Ideology in European Mass Publics, 1981–2016." *American Political Science Review* 113 (3): 674–93.

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): 239–56.

Choi, SeoYoun. 2018. "Bureaucratic Characteristics and Citizen Trust in Civil Service in OECD Member Nations." *International Area Studies Review* 21 (2): 114–33.

Claassen, Christopher. 2019. "Estimating Smooth Country–Year Panels of Public Opinion." *Political Analysis* 27 (1): 1–20.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, et al. 2023. "V-Dem [Country-Year/Country-Date] Dataset V13." Varieties of Democracy (V-Dem) Project. https://doi.org/10.23696/vdemds23.

Dahl, Robert A. 1947. "The Science of Public Administration: Three Problems." *Public Administration Review* 7 (1): 1–11.

Easton, David. 1965. "A Systems Analysis of Political Life."

———. 1975. "A Re-Assessment of the Concept of Political Support." *British Journal of Political Science* 5 (4): 435–57.

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*, 1–27.

Hooghe, Marc. 2011. "Why There Is Basically Only One Form of Political Trust." *The British Journal of Politics & International Relations* 13 (2): 269–75.

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): 1203–14.

Hu, Yue, Yuehong Cassandra Tai, Hyein Ko, Byung-Deuk Woo, and Frederick Solt. 2023. "Support for Democracy Is Multidimensional: Why Unidimensional Latent Variable Measures of Democratic Support Are Invalid." SocArXiv. https://osf.io/preprints/socarxiv/rym8g/.

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): 611–35.

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/`

Marien, Sofie, and Marc Hooghe. 2011. "Does Political Trust Matter? An Empirical Investigation into the Relation Between Political Trust and Support for Law Compliance."

European Journal of Political Research 50 (2): 267–91.

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): 315–37.

Norris, Pippa. 1999. "Introduction: The Growth of Critical Citizens?" *Critical Citizens: Global Support for Democratic Government*, 1–27.

———. 2011. *Democratic Deficit: Critical Citizens Revisited*. Cambridge University Press.

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang, Juraj Medzihorsky, Joshua Krusell, Farhad Miri, and Johannes von Römer. 2023. "The v-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data." V-Dem Working Paper No. 21, 8th edition. University of Gothenburg: Varieties of Democracy Institute.

Pollitt, Christopher. 2011. "Not Odious but Onerous: Comparative Public Administration." *Public Administration* 89 (1): 114–27.

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): 1251–69.

Rothstein, Bo, and Dietlind Stolle. 2008. "The State and Social Capital: An Institutional Theory of Generalized Trust" 40 (4): 441–59.

Shor, Boris, Joseph Bafumi, Luke Keele, and David Park. 2007. "A Bayesian Multilevel Modeling Approach to Time-Series Cross-Sectional Data." *Political Analysis* 15 (2): 165–81.

Solt, Frederick. 2020a. "DCPO: Dynamic Comparative Public Opinion." Available at the Comprehensive R Archive Network (CRAN). `https://CRAN.R-project.org/package=DCPO`.

———. 2020b. "Measuring Income Inequality Across Countries and over Time: The Standardized World Income Inequality Database." *Social Science Quarterly* 101 (3): 1183–99.

———. 2020c. "Modeling Dynamic Comparative Public Opinion." SocArXiv.

https://osf.io/ preprints/socarxiv/d5n9p.

Tai, Yuehong Cassandra. 2022. "The Role of Trust in Governance: Public Health and Tax Compliance in Cross-National Comparative Perspective." PhD thesis, The University of Iowa.

Tai, Yuehong 'Cassandra', Yue Hu, and Frederick Solt. 2024. "Democracy, Public Support, and Measurement Uncertainty." *American Political Science Review* 118 (1): 512–18.

Thomas, Craig W. 1998. "Maintaining and Restoring Public Trust in Government Agencies and Their Employees." *Administration & Society* 30 (2): 166–93.

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): 745–60.

Wilson, Matthew Charles, and Carl Henrik Knutsen. 2020. "Geographical Coverage in Political Science Research." *Perspectives on Politics* FirstView: 1–16.

Woo, Byung-Deuk, Hyein Ko, Yuehong Cassandra Tai, Yue Hu, and Frederick Solt. 2023. "Measuring Tolerance of Homosexuality in the Mass Public Across Countries and over Time."

Yang, Kaifeng, and Marc Holzer. 2006. "The Performance–Trust Link: Implications for Performance Measurement." *Public Administration Review* 66 (1): 114–26.

Yates, Douglas. 1982. *Bureaucratic Democracy: The Search for Democracy and Efficiency in American Government*. Harvard University Press.

Trust in Bureaucracy: Public Opinion on Public Servants in Dynamic Comparative Perspective

October 06, 2024

Appendices

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 17 such survey items that were asked in no fewer than five country-years in countries surveyed at least twice; these items were drawn from 132 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 98 different countries in at least two time points over 36 years, from 1973 to 2022, yielding a total of 1,814 country-year-item observations. The number of items observed in the source data for each country-year is plotted in Figure @ref(fig:obs_by_cy) 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.

Table A1: Indicators Used in the Unidimensional Latent Variable Model of Democratic Support

Survey Item Code	Country-Years	Question Text	Response Categories	Dispersion	Difficulties	Survey Dataset Codes
trust4	614	And how much trust do you have in... Civil service / public administration	1 A great deal of trust / 2 Quite a lot of trust / 3 Not a lot of trust / 4 No trust at all	0.88	-1.39, 1.00, 3.86	evs, wvs, ases, lb, bsa, asianb, eass, itanes, kgss, sasianb, arabb
trust2	348	I would like to ask you a question about how much trust you have in certain institutions. For each of the following institutions, please tell me if you tend to trustit or tend not to trust it? Civil service	1 Tend to trust / 2 Tend not to trust	1.13	0.79	eb, cceb
runswell4	197	How would you judge the current situation in each of the following? The way public administration runs in	1 Very good / 2 Rather good / 3 Rather bad / 4 Very bad			eb
right5	109	Most of the time we can trust people in government to do what is right.	1 Strongly agree / 5 Strongly disagree	0.85	-1.11, 0.65, 2.00, 4.66	issp, usgss
best5	94	Most government administrators	1 Strongly agree / 5 Strongly disagree	1.24	-1.61, 0.53, 2.32, 5.65	issp, usgss, kgss
image4	89	Could you please tell me for each of the following, whether the term brings to mind something very positive, fairly positive, fairly negative or very negative. Public administration	1 Very positive / 2 Fairly positive / 3 Fairly negative / 4 Very negative	0.83	-1.80, 0.30, 3.03	eb
trustmun4	81	Generally speaking, the public administration of [CITY NAME] can be trusted	1 Strongly agree / 4 Strongly disagree			feb, lb
trustff4	73	Please look at this card and tell me how much confidence you have in each of the following groups, institutions or persons mentioned on the list: a lot, some, a little or no confidence? Firefighters	1 A lot / 2 Some / 3 A little / 4 None			lb
right4	60	You can generally trust the people who run our government to do what is right.	1 Strongly agree / 4 Strongly disagree	0.54	-0.58, 1.18, 3.44	asianb
trust3	32	Trust in Ministries and Government Agencies	1 Very much / 2 Some / 3 Not very much	1.02	-0.02, 3.02	usgss, jgss
trusteuro2	30	If you would trust information they provide on the changeover to the euro: Public administration	1 trust / 2 do not trust			feb
trustpollution5	26	How much trust do you have in each of the following groups to give you correct information about causes of pollution? Government departments	1 A great deal of trust / 2 Quite a lot of trust / 3 Some trust / 4 Not much trust / 5 Hardly any trust			issp

(continued)

Survey Item Code	Country-Years	Question Text	Response Categories	Dispersion	Difficulties	Survey Dataset Codes
trust5	23	Confidence in the Civil Service?	1 Complete confidence / 2 A great deal of confidence / 3 Some confidence / 4 Very little confidence / 5 No confidence at all	0.54	-0.81, 0.41, 1.88, 3.38	issp, gles, fsdtrust, fsdeva, bsa
trust11	16	Now, thinking about institutions like Parliament, please use the scale of 0 to 10 to indicate how much trust you have for each of the following, where 0 is no trust and 10 is a great deal of trust:	0 No trust / 10 A great deal of trust	0.56	-1.36, -1.01, -0.55, -0.08, 0.32, 1.04, 1.50, 2.15, 3.11, 3.96	cid, fsdelection, bes
right4a	10	In general, do you feel that the people in government are too often interested in looking after themselves, or do you feel that they can be trusted to do the right thing nearly all the time?	1 Usually look after themselves / 2 Sometimes look after themselves / 3 Sometimes can be trusted to do the right thing / 4 Usually can be trusted to do the right thing	0.81	0.48, 1.49, 2.93	aes
interests7	8	To what extent do you trust each of these political institutions to look after your interests? Civil servants	1 No trust / 7 Great trust	0.40	-1.14, 0.20, 0.84, 1.53, 2.86, 3.46	neb
trustmun7	4	Please tell me for each institution or organisation how much trust you place in it. The municipal administration	1 Absolutely no trust at all / 23456 / 7 A great deal of trust			allbus

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 (2020c).[1] 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:

$$a_{ktqr} = \phi\eta_{ktqr} \tag{1}$$

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

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

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.[2]

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

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

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.[3]

[1] Solt (2020c) 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.

[2] The 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.

[3] Estimating 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,

The dispersion of question q, its noisiness in relation to our latent variable, is α_q. The mean and standard deviation of the 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.[4] 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:

$$\sum_{k=1}^{K} \delta_{kq} = 0 \tag{5}$$

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 2020c, 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 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,

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.

[4]The 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 βs 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 $\sigma^2_{\bar{\theta'}}$ and σ^2_{σ}; 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.

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

Caughey, Devin, Tom O'Grady, and Christopher Warshaw. 2019. "Policy Ideology in European Mass Publics, 1981–2016." *American Political Science Review* 113 (3): 674–93.

Claassen, Christopher. 2019. "Estimating Smooth Country–Year Panels of Public Opinion." *Political Analysis* 27 (1): 1–20.

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/`

McGann, Anthony, Sabastian Dellepiane-Avellaneda, and John Bartle. 2019. "Parallel Lines? Policy Mood in a Plurinational Democracy." *Electoral Studies* 58: 48–57.

Solt, Frederick. 2020a. "DCPO: Dynamic Comparative Public Opinion." Available at the Comprehensive R Archive Network (CRAN). `https://CRAN.R-project.org/package=DCPO`.

———. 2020b. "Modeling Dynamic Comparative Public Opinion." SocArXiv. `https://osf.io/ preprints/socarxiv/d5n9p`.