[Trust in Bureaucracy]*

Yuehong 'Cassandra' Tai   *Pennsylvania State University*
Frederick Solt   *University of Iowa*

[0/150 words]

Intro

In this letter, we present the trust in civil servants (PGE) 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 public's attitudes on gender equality in the public sphere of politics and paid work across countries and over time. We show that these PGE scores are strongly correlated with responses to single survey items as well as with measures of women's participation in the workforce and in the boardroom. We expect that the PGE data will become an invaluable source for broadly cross-national and longitudinal research on the causes and effects of collective attitudes toward gender equality in the public sphere.

Examining the Source Data on Trust in Bureaucracy

National and cross-national surveys have often included questions tapping attitudes toward equality for women and men in the public sphere 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 137 different countries in at least two time points over 48 years, from 1973 to 2021, yielding a total of 1,769 country-year-item observations. Observations for every year in each country surveyed would number 6,576, and

---

*Corresponding author: frederick-solt@uiowa.edu. Current version: March 18, 2022.

[1]The complete list of trust in civil servants survey items is included in online Appendix A. A discussion of excluded survey items—those covering the related but distinct concepts of gender equality in the private sphere, how women should balance opportunities in the public sphere with their traditional duties in the private sphere, or issues of gender dominance such as sexual harassment—is found in online Appendix B.

a complete set of country-year-items would encompass 105,216 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 1,370 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 2,804 country-years spanned by the data we collected. But there can be no denying Claveria's (**?**) observation that the many different survey items employed renders these data incomparable and difficult to use together.
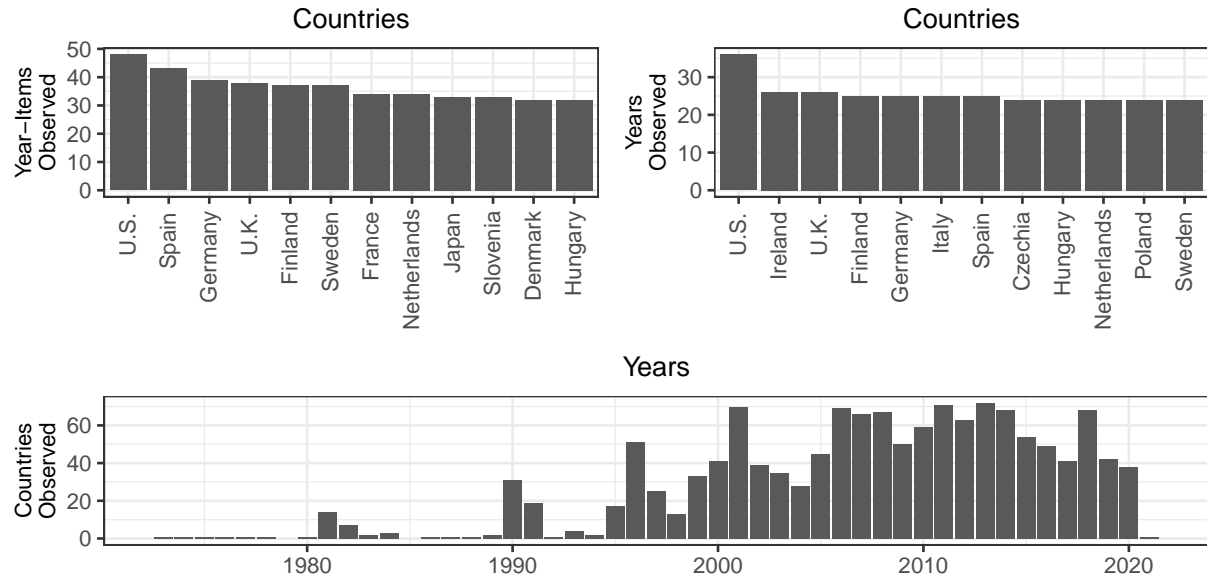

Figure 1: Countries and Years with the Most Observations in the PGE Source Data

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 "On the whole, men make better political leaders than women do." Employed by the Americas Barometer, the Arab Barometer, the Eurobarometer, the Latinobarómetro, the Pew Research Center, 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 remember, 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, NA countries—Andorra, Benin, Botswana, Burundi, Cameroon, Cape Verde, Côte d'Ivoire, Eswatini, Ethiopia, Guinea, Jamaica, Kazakhstan, Kenya, Lebanon, Lesotho, Liberia, Libya, Madagascar, Malawi, Mauritius, Mozambique, Namibia, Nepal, Niger, Palestinian Territories, Russian Federation, Rwanda, Senegal, Sierra Leone, Sri Lanka, Togo, Trinidad & Tobago, and Yemen—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 gender egalitarianism in the public sphere. 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 PGE 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 **????**). 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
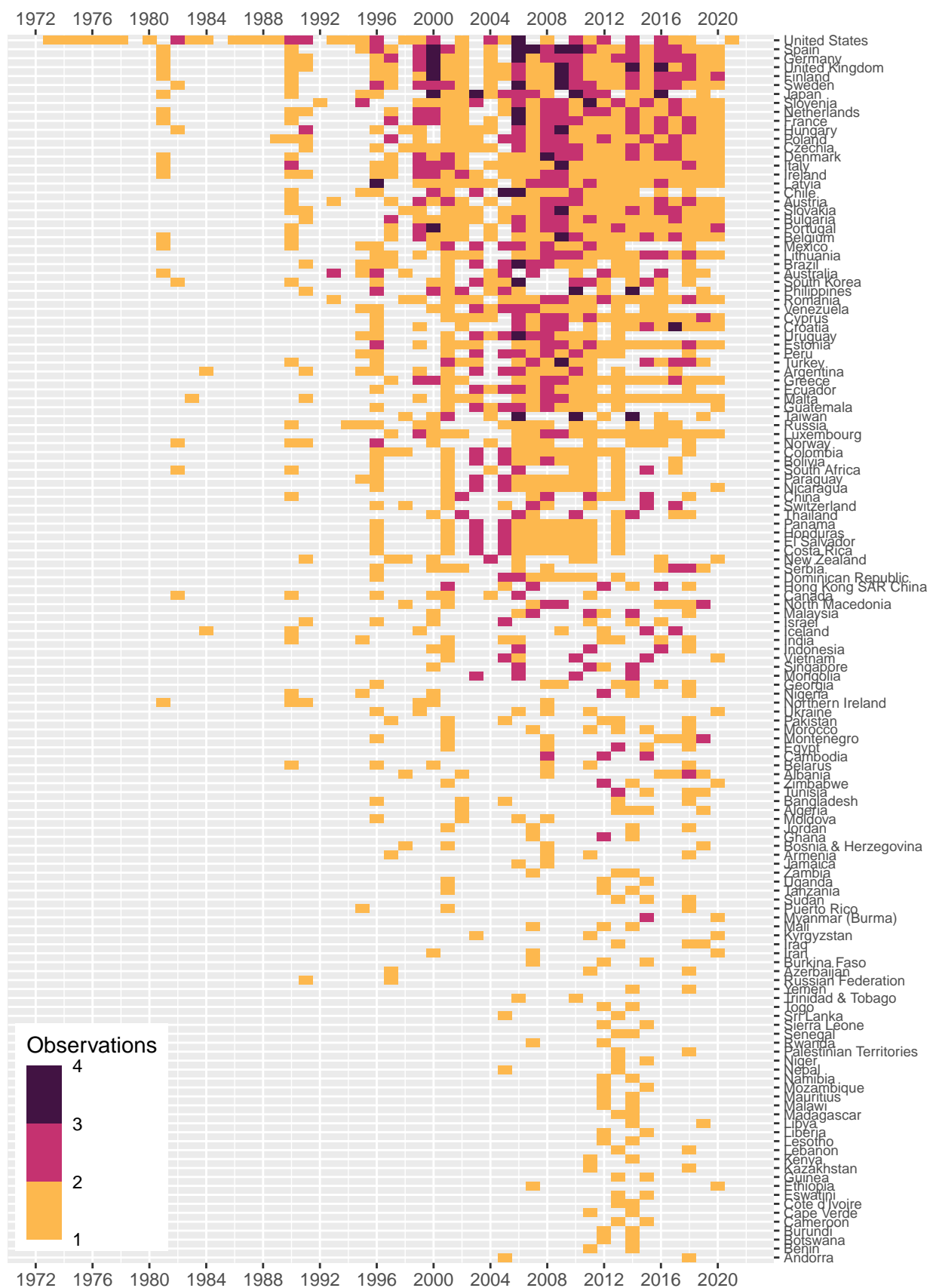
Figure 2: Source Data Observations by Country and Year

but also sparse, the Dynamic Comparative Public Opinion (DCPO) model elaborated in **?**.[2] 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 **?**, 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 expense of greater measurement uncertainty.

We estimated the model using the `DCPOtools` package for R (**?**), running four chains for 1,000

---

[2]**?** demonstrates that the DCPO model provides a better fit to survey data than the models put forward by **?** or **?**. The **?** model depends on dense survey data unlike the sparse data on trust in civil servants described in the preceding section. **?** 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.
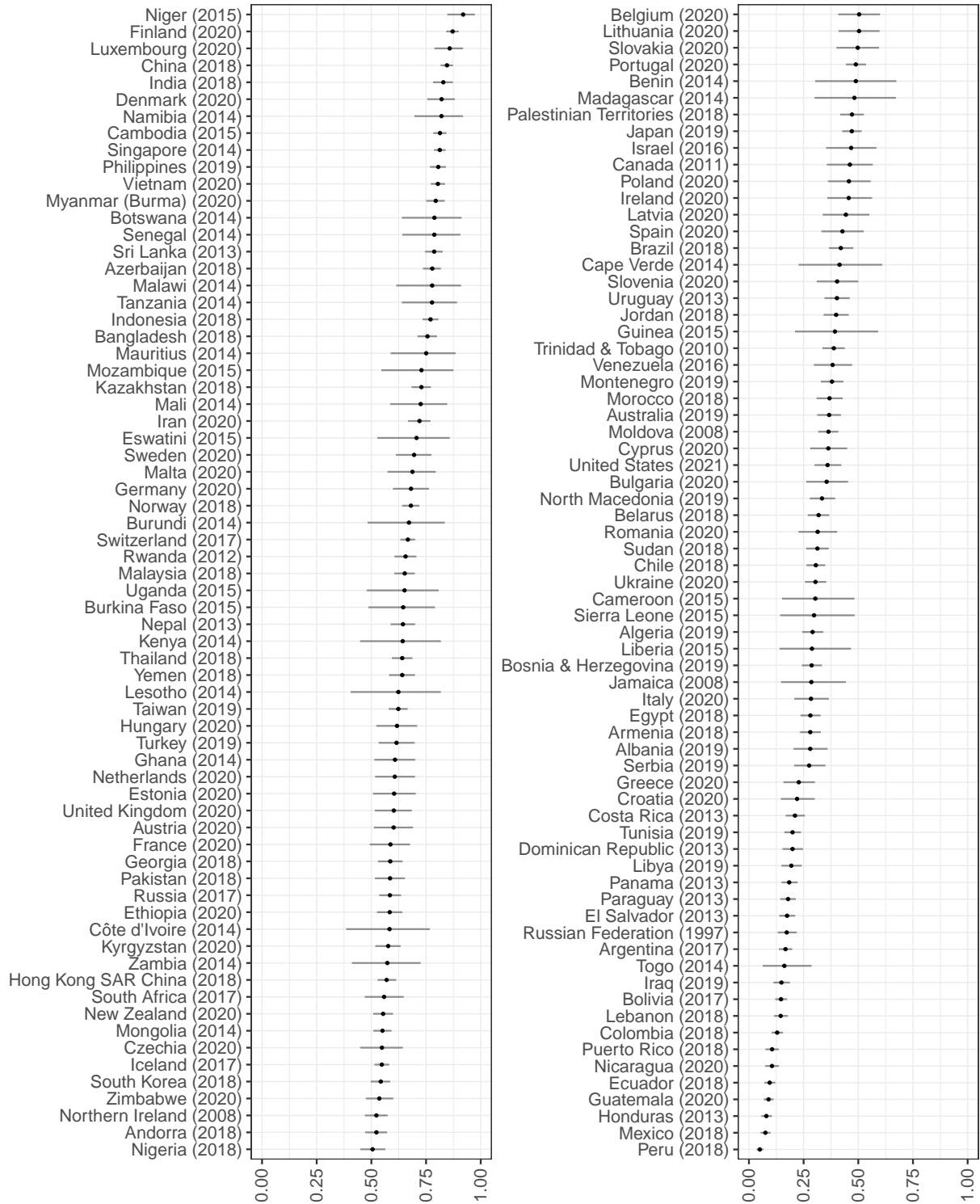
Note: Gray whiskers represent 80% credible intervals.

Figure 3: TCS Scores, Most Recent Available Year

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 1.02, indicating that the model converged.

The dispersion parameters of the survey items indicate that all of them load well on the latent variable (see Appendix A). The result is estimates, in all 2,782 country-years spanned by the source data, of mean public gender egalitarianism, what we call PGE scores. Figure 3 displays the most recent available PGE score for each of the 137 countries and territories in the dataset.

The Scandinavian countries and France are at the top of this list, along with Puerto Rico, which has had women of both of its major parties serve as chief executive and as recently as 2020 had a woman from each party holding the two most prominent elected offices on the island. The latest scores for Peru, Mexico, Honduras, Guatemala, and Ecuador have them as the places where public opinion is least favorable to gender equality in the public sphere.

Appendix A: Survey Items Used to Estimate Public Gender Egalitarianism and Their Distribution

National and cross-national surveys have often included questions tapping attitudes toward equality for women and men in the public sphere 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 more public gender egalitarianism from those with less. 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 137 different countries in at least two time points over 48 years, from 1973 to 2021, yielding a total of 1,769 country-year-item observations. The number of items observed in the source data for each country-year is plotted in Figure 2 below. The PGE 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 **????**). 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 **?**.[3] 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.[4]

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

---

[3]**?** demonstrates that the DCPO model provides a better fit to survey data than the models put forward by **?** or **?**. The **?** model depends on dense survey data unlike the sparse data on trust in civil servants described in the preceding section. **?** 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.

[4]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.

[5]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, 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.[6] 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 **?**, 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

---

[6]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 $\sigma_{\bar{\theta}'}^2$ 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.

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