Public Gender Egalitarianism: A Dataset of Dynamic Comparative Public Opinion Toward Egalitarian Gender Roles in the Public Sphere*

Introduction

Focusing on Public Opinion Toward Gender Equality in the Public Sphere

We classified the available survey items on gender egalitarian attitudes into four categories. The first of these categories consists of questions asking respondents' views of gender equality in the traditionally masculine public sphere of education, paid work, and politics. The second category, in turn, encompasses questions focusing on gender equality in the traditionally feminine private sphere of housework and childcare. A third category we identified comprises questions asking respondents how women should balance opportunities in the public sphere with their traditional duties in the private sphere, such as whether mothers in the workforce can have similarly warm relationships with their children as mothers who are not; it is telling, though not surprising, that the complementary set of questions, on how *men* should balance responsibilities in the private sphere with their traditional roles in the public sphere, is only rarely included in surveys.[1] The fourth and final category includes respondents' views on various forms of women's domination by men, from whether wives should adopt their husbands' surnames through the recognition that various forms of sexual harassment are not "flattering" to the justifiability of intimate partner violence committed by husbands.

To avoid any potential multidimensionality in the Public Gender Egalitarianism data, we include only survey items corresponding to the first category, gender egalitarianism in the public sphere.[2]

*Corresponding author: frederick-solt@uiowa.edu. Current version: December 23, 2020.

[1]One laudable example of this mostly unasked sort of question, apparently first included in Australia's 1989 National Social Science Survey and slowly becoming more common, is the item querying respondents the extent to which they agree with the statement, "Family life often suffers when men concentrate too much on their work."

[2]Items that fell into both the first category and another, such as the `tradroles` questions—those with stems of the form "A man's job is to earn money; a woman's job is to look after the home and family"—which by explicitly counterposing the traditional gender roles suggests women unlike men should not earn money in the public sphere (while also suggesting that men unlike women should not look after home and family in the private sphere) were included in the PGE source data. The complete list of gender egalitarianism items is included in the Appendix.

Thus, the "Public" in the name of the dataset does double duty, referring both to the fact that it measures the *public's* attitudes on gender equality and to its attitudes specifically on this *public* aspect of gender egalitarianism. In all, we identified 49 of these survey items that were asked in no fewer than five country-years in countries surveyed at least twice; these items were drawn from 85 different survey datasets. We describe these source data in detail in the next section.

Examining the Source Data on Public Gender Egalitarianism

Together, the survey items in the source data were asked in 123 different countries in at least two time points over 48 years, from 1972 to 2020, yielding a total of 2,913 country-year-item observations. Considering that observations for every year in each country surveyed would number 5,904, a complete set of country-year-items would encompass 289,296 observations. Viewed from this perspective of complete data, the available data can be seen to be very, very sparse. From a more optimistic standpoint, we note there there are 1,168 country-years in which we have at least *some* information about the public gender egalitarianism of the population, that is, some 44% of the 2,651 country-years spanned by the data we collected. But there can be no gainsaying that the many different survey items employed renders these data difficult to use together.

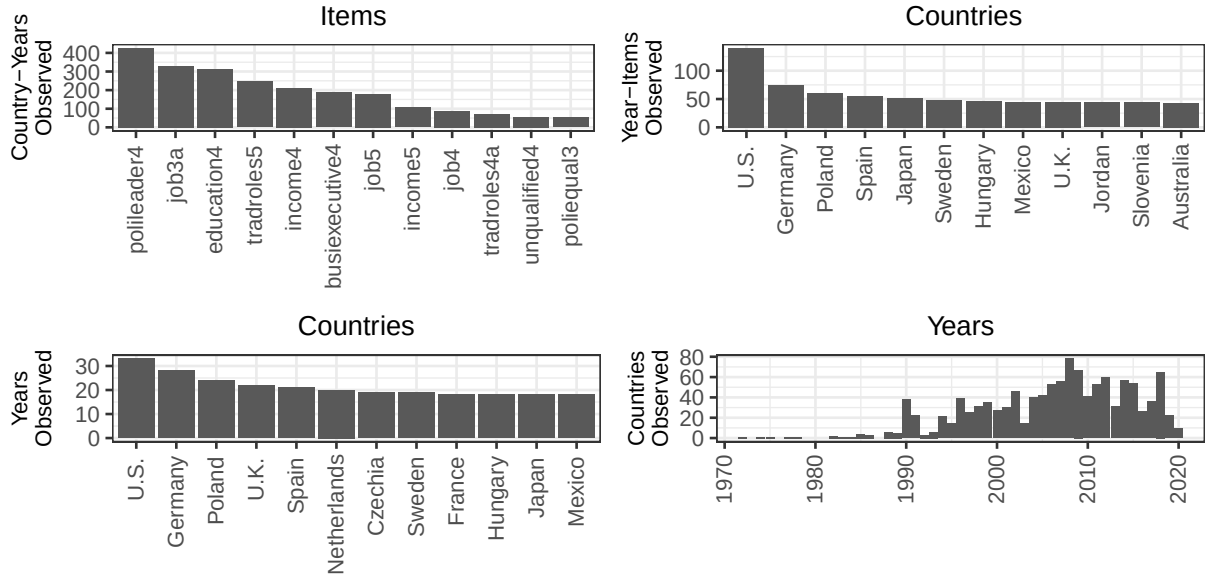

Figure 1: Items, Countries, and Years with the Most Observations in the PGE Source Data

The left panel of Figure 1 displays in how many country-years each of the twelve most-commonly asked survey items are available. The `polileader4` item, 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," was the most frequently asked question in the data we collected. 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 426 different country-years. That this constitutes only 16% of the country-years spanned by our data—and remember, `polileader4` is the *most common* survey item—again underscores just how sparse and incomparable the available public opinion data is on this topic.

Which countries are the most data-rich? The right panel of Figure 1 shows the dozen countries with the highest count of country-year-item observations. The United States, with 140 observations, is far and away the best represented country in the source data, followed by Germany, Poland, Spain, and Japan. At the other end of the spectrum, three countries—Cambodia, Sri Lanka, and Suriname—have only the minimum two observations required to be included in the source dataset at all.

Estimating Public Gender Egalitarianism

There has been a recent blossoming of scholarship developing 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 public gender egalitarianism across countries and over time, we draw on 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 presented in Solt (2020*c*).[3]

The DCPO model is a population-level two-parameter ordinal logistic item response theory (IRT) model with country-specific item bias. That is, η_{ktqr}, 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

[3]Solt (2020*b*) 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 public gender egalitarianism described in the preceding section. Kolczynska et al. (2020) is the very most recent of the five works and builds on each of the others, but the MRP approach developed in that piece is suitable only when the available survey data are dense and ancillary data on population characteristics are also available, so it is similarly inappropriate to this application.

r, is estimated as varying with the mean and standard deviation of the unbounded latent trait of public opinion, $\bar{\theta'}_{kt}$ and σ_{kt}; with the dispersion, α_q, of the question and the difficulty, β_{qr}, of the question's response category; as well as with a term to capture how responses to particular questions in particular questions may be biased, δ_{kq}: $\eta_{ktqr} = \text{logit}^{-1}(\frac{\bar{\theta'}_{kt}-(\beta_{qr}+\delta_{kq})}{\sqrt{\alpha_q^2+(1.7*\sigma_{kt})^2}})$. A beta-binomial distribution is used to model the observed survey data given η_{ktqr}, 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. Finally, the logistic function is used to transform $\bar{\theta'}_{kt}$ to the unit interval and so give the bounded mean of latent public opinion, $\bar{\theta}_{kt}$, which is our parameter of interest here (see Solt 2020*c*, 3-8).

We used the `DCPO` package for R (Solt 2020*a*) to estimate this model on the public gender egalitarianism source data described above. The result is estimates, in all 2,651 country-years spanned by the source data, of mean public gender egalitarianism, what we call PGE scores. Figure 2 displays the most recent available PGE score for each of the 123 countries in the dataset.

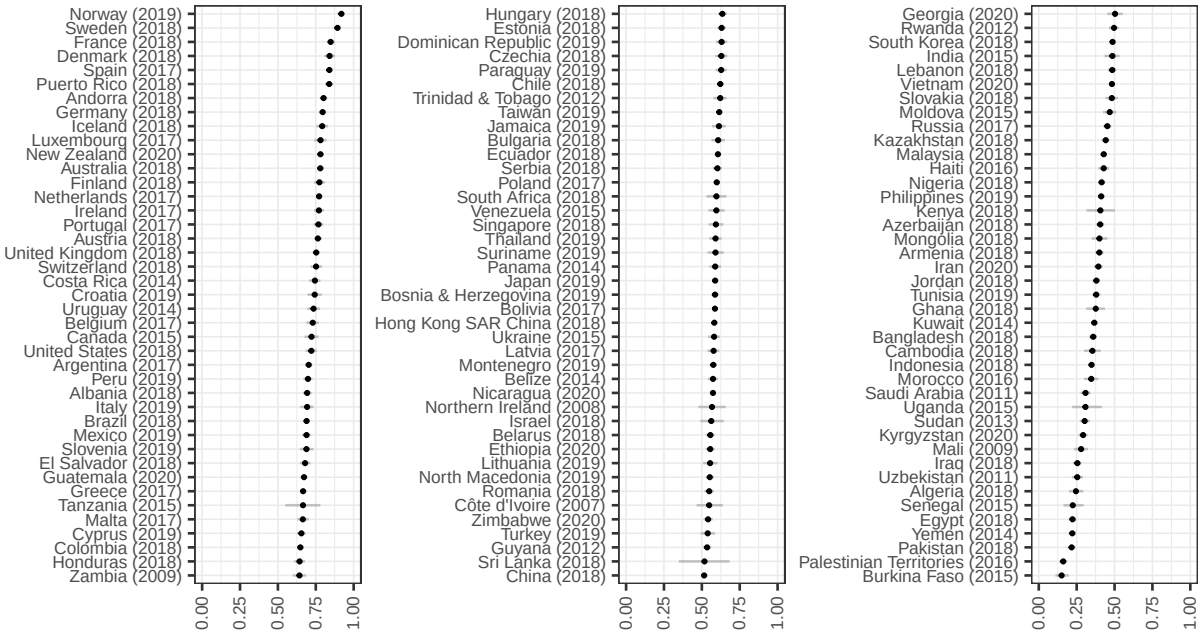

Figure 2: PGE Scores, Most Recent Available Year

The Scandinavian countries, France, and Spain are at the top of this list, followed by 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 Egypt, Yemen, Pakistan, the Palestinian Territories, and Burkina Faso have them as the places where public opinion is least favorable to gender equality in the public sphere.

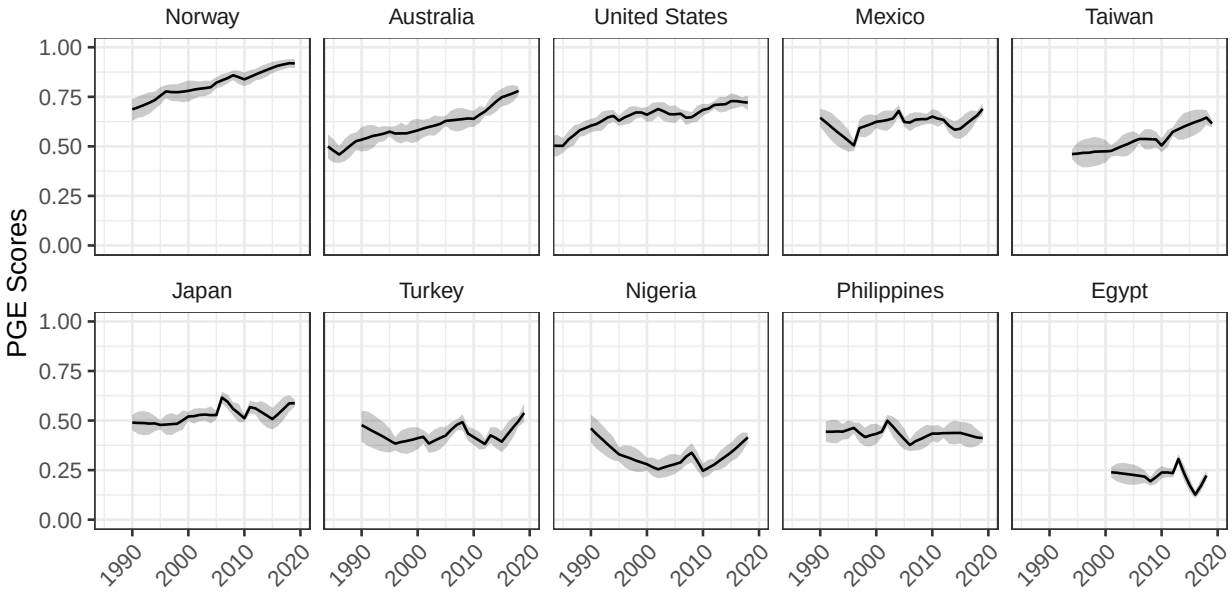

Figure 3: PGE Scores Over Time, Selected Countries

Figure 3 shows the time series of the PGE scores for ten countries.

Validating Public Gender Egalitarianism

We consider now the validity of the PGE scores. Like Caughey, O'Grady and Warshaw (2019, 684-685), we provide evidence of our measure's validity with convergent validation and construct validation.

Figure 4 displays our demonstrations of convergent validation. Convergent validation refers to showing that a measure is empirically associated with alternative indicators of the same concept (Adcock and Collier 2001, 540). Here, we compare PGE scores to responses to individual survey items that were included in the source data, that is, we provide an "internal" validation test (see, e.g., Caughey, O'Grady and Warshaw 2019, 689; Solt 2020*c*, 10). In the left panel, we examine the `polileader4` question mentioned above, the most common item in the source data across all country-years. Then, in the center panel, we look at the question that provides most data-rich cross-section, the `poliequal3` item in Pew Global's Spring 2007 survey. Finally, in the right panel,

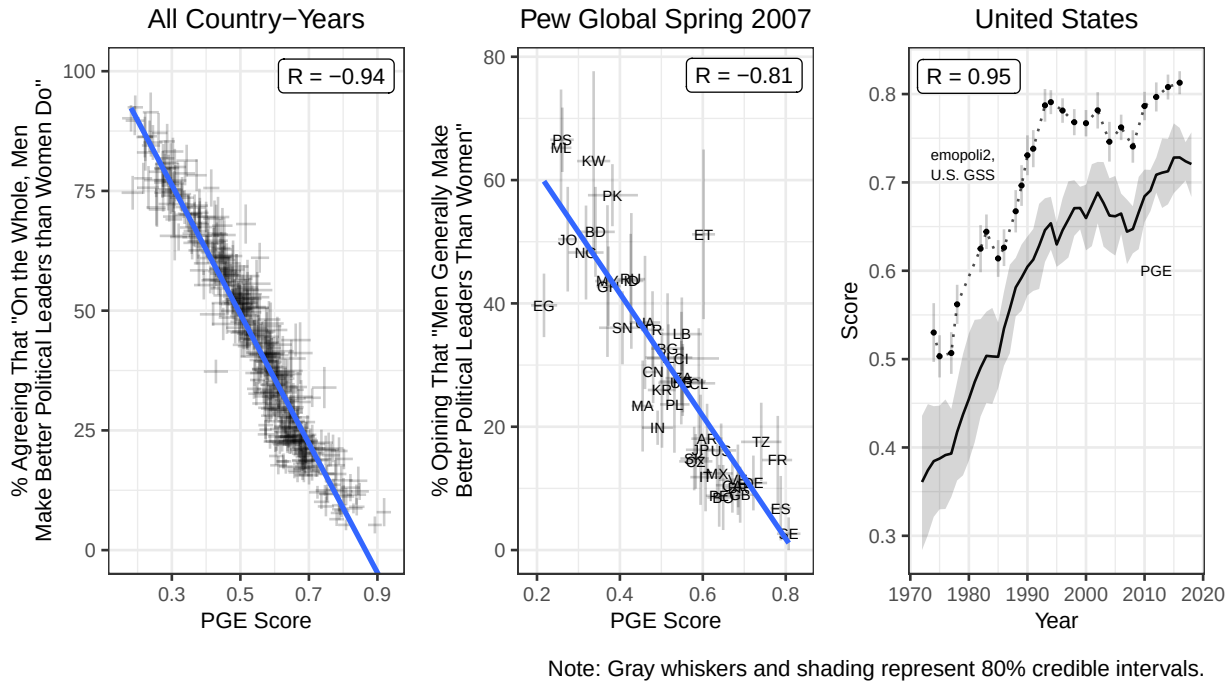

Figure 4: Convergent Validation: Correlations Between PGE Scores and Individual PGE Source Data Survey Items

to evaluate how well the PGE score captures change over time, we focus on the item with the largest number of observations for a single country, the `emopoli2` survey item, which asks whether respondents agreed or disagreed that "most men are better suited emotionally for politics than are most women." In every case, the correlations—estimated taking into account the uncertainty in the measures[4]—are in the expected direction and very strong.

We move on, then, to construct validation, which refers to demonstrating, for some *other* concept believed causally related to the concept a measure seeks to represent, that the measure is empirically associated with measures of that other concept (Adcock and Collier 2001, 542). In Figure 5, we look to individual survey items from our third category of gender egalitarianism, that is, questions that ask how women should balance opportunities in the public sphere with their traditional duties in the private sphere. Assuming that attitudes that women should prioritize housework and childcare over education, politics, and paid employment—or convictions that there will be negative consequences

[4]The uncertainty in the PGE score and in the percentage in the population who would agree with the item does not substantially affect the correlation with the `polileader4` question, but failing to account for this uncertainty would overstate the correlation with the Pew `poliequal3` item, at R = -0.87, and the U.S. GSS `emopoli2` item, at $R = 0.98$. We take up the issue of the importance of taking uncertainty into account when working with the PGE data in a subsequent section.

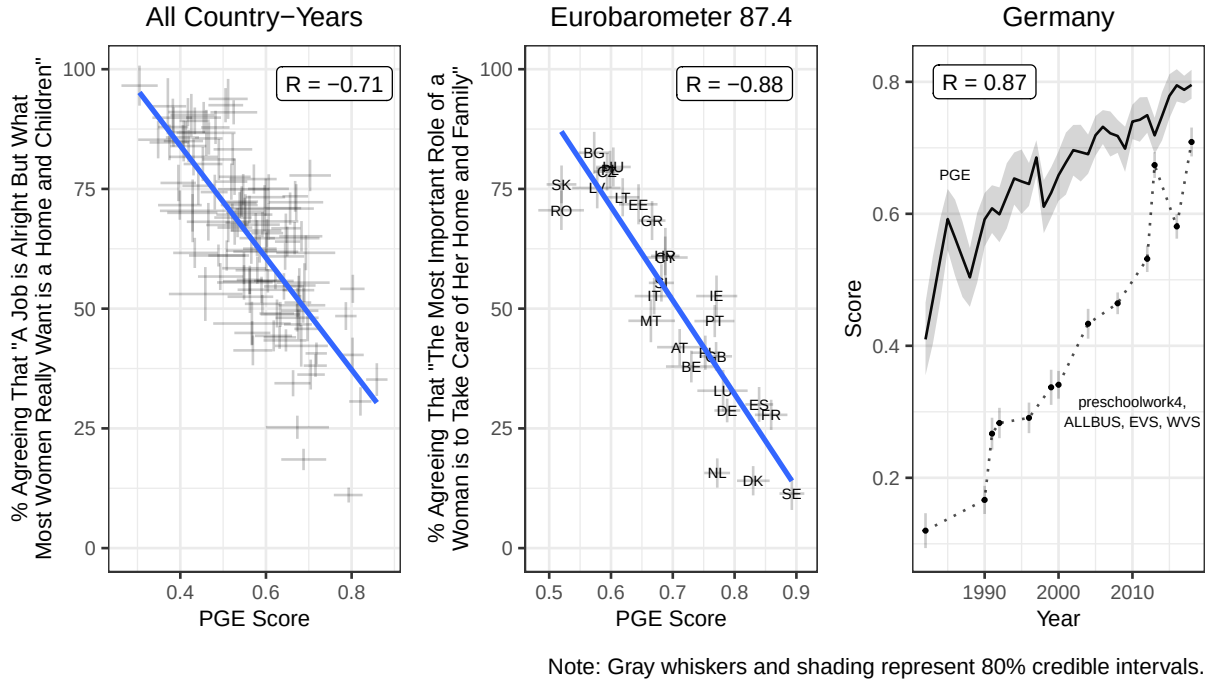

Figure 5: Construct Validation: Correlations Between PGE Scores and Individual 'Balancing' Gender Egalitarianism Survey Items

if they do not—will lead to less gender egalitarian opinions with regard to these latter, public-sphere activities, evidence for this theoretical relationship will provide construct validation for the PGE score. Exemplars of such items across all available country-years ("a job is alright but what most women really want is a home and children" from the WVS and EVS), in cross-section ("the most important role of a woman is to take care of her home and family" from the Eurobarometer 87.4), and in time series ("a pre-school child is likely to suffer if his or her mother works" from the German ALLBUS, WVS, and EVS) all show strong correlations with the PGE scores.

Finally, Figure 6 shows additional tests of construct validation. As attitudes toward gender egalitarianism in the public sphere plausibly both cause and are caused by women's gains in the workplace, strong relationships between the PGE scores and measures of workplace gender equality provide construct validation for our measure. In the left panel of Figure 6, we compare the PGE scores to the ratio of women's to men's labor force participation rates in 76 countries in 2016, drawing on data compiled by the Statistics Division of the UN Department of Economic and Social Affairs (2020). In the right panel, we plot the PGE scores against the percentage of women on the boards of directors of the largest publicly listed companies in 41 countries in 2017 (see OECD

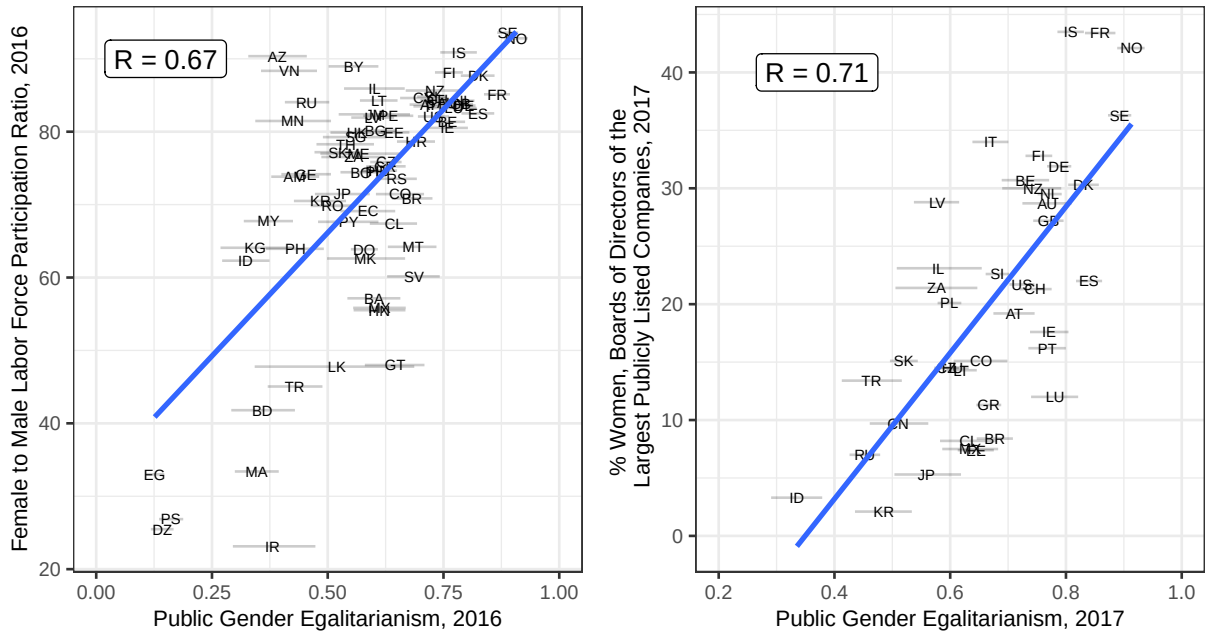

Figure 6: Construct Validation: Correlations Between PGE Scores and Indicators of Workplace Gender Equality

2020). Both correlations are strong. Together, this evidence of construct validation and convergent validation attests to the validity of the PGE scores as measures of public opinion towards gender equality in the public sphere.

Using the Public Gender Egalitarianism Dataset

Version 1.0 of the PGE dataset includes PGE scores for 123 countries for as many years as possible from 1972 to the present, a total of 2651 country-years. It can be accessed in two ways: via a user-friendly web application on the PGE website, which plots scores for as many as four countries for easy comparison of levels and trends, and via the Harvard Dataverse, where the entire dataset is available for download for use in statistical analysis. We will revise and update the dataset as new survey data on public gender egalitarianism becomes available.

One aspect of latent-variable estimates of public opinion like the PGE dataset that is easy for researchers to overlook is the uncertainty in the estimates. But neglecting to incorporate this uncertainty by using only the mean estimate for each country-year—that is, for example, the mean of the PGE scores—in an analysis can lead one to mistakenly conclude that the analysis supports

the hypothesis (see **?**) as well as to mistakenly conclude that it does *not* support the hypothesis (see Crabtree and Fariss 2015). Therefore, taking the uncertainty in the PGE scores into account is crucial to reaching well-grounded conclusions.

The download includes pre-formatted data to facilitate incorporating the uncertainty in the PGE scores. In R, the functions of the `purrr` package (Henry and Wickham 2019, also included in the widely-used `tidyverse` package (Wickham 2017)) make it entirely straightforward to incorporate the uncertainty of the PGE estimates. In Stata, the `mi estimate:` command prefix originally developed for analyzing multiply imputed data can be used to automate the process of building uncertainty into nearly any analysis. Step-by-step instructions on how to use these tools, complete with examples, are included in the data download.

References

Adcock, Rober and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95(3):529–546.

Caughey, Devin, Tom O'Grady and Christopher Warshaw. 2019. "Policy Ideology in European Mass Publics, 1981–2016." *American Political Science Review* pp. 1–20.

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

Crabtree, Charles D. and Christopher J. Fariss. 2015. "Uncovering Patterns Among Latent Variables: Human Rights and De Facto Judicial Independence." *Research & Politics* 2(3):1–9.

Henry, Lionel and Hadley Wickham. 2019. "purrr: Functional Programming Tools." Available at the Comprehensive R Archive Network (CRAN). `https://CRAN.R-project.org/package=purrr`.

Kolczynska, Marta, Paul-Christian Bürkner, Lauren Kennedy and Aki Vehtari. 2020. "Trust in State Institutions in Europe, 1989-2019." SocArXiv.

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

OECD. 2020. "Female Share of Seats on Boards of the Largest Publicly Traded Companies." `https://stats.oecd.org/index.aspx?queryid=54753`. Accessed 3 December 2020.

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*. "Measuring Income Inequality Across Countries and Over Time: The Standardized World Income Inequality Database." *Social Science Quarterly* .

Solt, Frederick. 2020*c*. "Modeling Dynamic Comparative Public Opinion.".
URL: *osf.io/preprints/socarxiv/d5n9p*

United Nations Department of Economic and Social Affairs, Statistics Division. 2020. "Minimum Set of Gender Indicators." `https://genderstats.un.org/#/downloads`. Accessed 3 December 2020.

Wickham, Hadley. 2017. "tidyverse: Easily Install and Load the 'Tidyverse'." R package version 1.2.1. `https://CRAN.R-project.org/package=tidyverse`.