Datasets and replication instructions

Michal Parizek and Matthew Stephen, 2020. The Increasing Representativeness of International Organizations' Secretariats: Evidence from the United Nations System, 1997-2015. *International Studies Quarterly*

Datasets used

The analysis is based on one main dataset and three additional small datasets which are used for some of the graphs and supplementary analyses. The use of the four datasets for the construction of individual tables and figures in the article and appendix is summarized in the table below.

Main dataset: representativeness_international_secretariats_main_data.txt

Additional datasets: representativeness_international_secretariats_igo_visibility_comparison.txt, representativeness_international_secretariats_additional_data.txt, representativeness_international_secretariats_imf_wto.txt

Replication instructions

For the most part, replication can be run line by line or at once, for the whole file. The figures and tables are not presented sequentially in the order in which they appear in the article and appendix, but rather in the order that follows the structure of the data, as they are aggregated and created in the script.

One exception to the general procedure where the script can be run line by line or at once is the analyses for Table A11 and Table A12, at the end of the replication file. These work with alternative measures of IO visibility. In Table A11, this measure is based on Google, instead of Factiva. In Table A12, the three measures of broader political visibility, or significance, are size of budget, the absolute growth of secretariat size, and significance of operation activity of IOs.

To replicate models in Table A11 and A12 correctly, the entire dataset needs to be reloaded and the correct visibility variable needs to be set, all at the very beginning of the script. Because of the way data are aggregated later on, the specific visibility variable must be changed there, at the beginning of data processing. This may inconvenient, because for Table A12, the entire data processing script needs to be re-run three time, once for each of the three alternative visibility measures (and, of course, always only the two specific models need to be created, the others not over-run). However, given that data are also aggregated by visibility scores, it is not possible to include more than one visibility measure in the data preparation process. Also note that visibility measures differ slightly in missing values so, again, for each of the analyses in Table A12, and models 4-6 in Table A11, the entire data processing script, from main data load, needs to be re-run. Do remember that for all other figures and tables than Table A12 and models 4-6 in Table A11, Factiva-based measure of visibility, the default, needs to be used.

The replication script contains a sizeable amount of additional specific instructions and explanatory notes for individual lines of code.

Table: Use of datasets in replication of figures and tables

		Main dataset used	Any additional dataset
Figure 1	The development of the share of OECD citizens on the UN system, IMF, and WTO staff distribution	_main_data	_imf_wto_y
Table 1	OLS and panel regression ('within-between' random effects) results	_main_data	
Figure 2	The diminishing effects of economic power (GNI) and the rising effect of population size on staffing over time (extracted from Model 5 in Table 1)	_main_data	
Table A1	Interviews with senior diplomatic staff, Geneva 2017 and 2018		
Table A2	Overview of IOs (bodies) in the sample	_main_data	
Figure A1	The development of UN bodies staff size, 1997-2015	_main_data	
Table A3	Country representation and changes	_main_data	
Figure A2	Countries' absolute and per million inhabitants representation on staff (average for the period 1997-2015)	_main_data	
Figure A3	Staff and weighted staff		_additional_data
Table A4	WHO and UN Secretariat staff seniority		_additional_data
Table A5	Contract types (aggregate across all UN system bodies)		
Figure A4	Change in % share of permanent contracts, related to economic wealth and power		_additional_data
Figure A5	GNI and budgetary contributions		_additional_data
Table A6	GNI and budgetary contributions		_additional_data
Figure A6	Google and Factiva		_igo_visibility_comparison
Table A7	Factiva, Google, and GDELT as visibility scores sources		_igo_visibility_comparison
Table A8	Pooled OLS models, replicating results from Models 5-7 from Table 1	_main_data	
Table A9	Percent (%) share on staff as the outcome variable, replicating Models 5-7 from Table 1	_main_data	
Table A10	Cross sectional OLS models, replicating results from Table 1	_main_data	
Table A11	Google search measure of visibility and exclusion of extreme observations	_main_data	

Figure A7	Alternative conceptualizations of IO 'visibility': budget size, staff expansion, and operational activity	_main_data	
Table A12	Alternative explanations: budget size, staff expansion, and operational activity	_main_data	