Who Reviews Whom, Where, and Why? Evidence from the Peer Review Process of the OECD Development Assistance Committee

Alice Iannantuoni* Simone Dietrich† Bernhard Reinsberg‡

International Studies Quarterly§

*European School of Political and Social Sciences, Université Catholique de Lille, alice.iannantuoni@univ-catholille.fr.
†Department of Political Science and International Relations, University of Geneva, simone.dietrich@unige.ch.
‡School of Social and Political Sciences, University of Glasgow, bernhard.reinsberg@glasgow.ac.uk.
§The authors thank Katharina Fleiner and Tomas Spahini for research assistance and Çağlayan Başer, Ekrem Başer, Nuole Chen, Richard Clark, Christoph Mikulaschek, Daniel Nielson, and participants to the PEIO 2022 annual meeting and the ECPR Joint Sessions Workshop "The Performance of Global Governance Institutions" for helpful comments on a previous draft.

Appendix

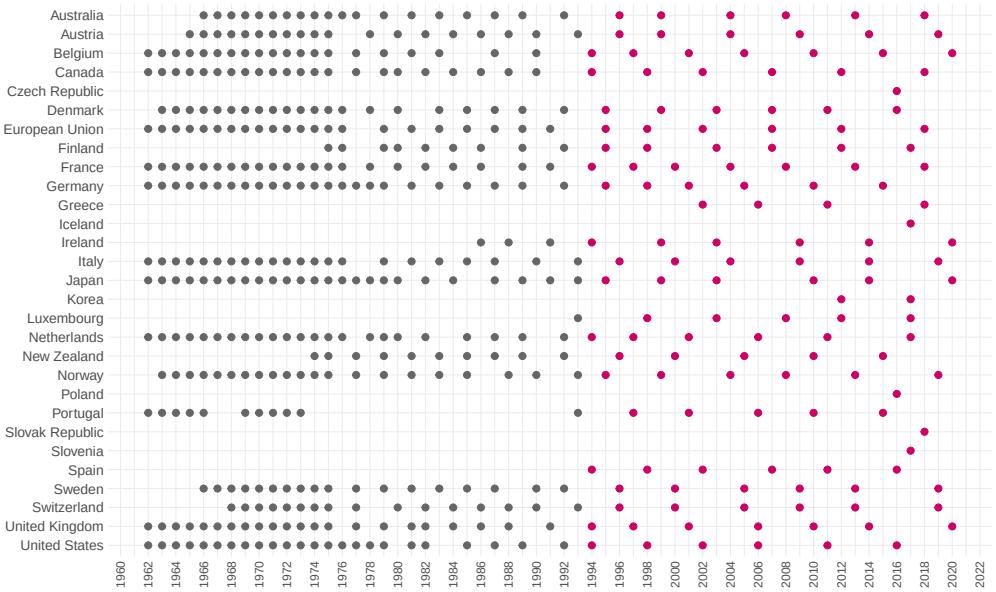
Updated August 2024

A1 OECD DAC Peer Reviews: The Data

The analyses presented in this paper rely on metadata about the OECD DAC donor peer reviews. We collected information about the entire universe of existing peer reviews—517 reviews conducted 1962–2020—through documents available at the OECD library.[1] For all 517 reviews, we know the DAC member who was reviewed, the date of the review, and the peer examiner countries, as we are able to retrieve this information without the need to access the full peer review report. These are the data we use in the analyses of the choice of peer examiners. For 143 of the most recent reviews, we also have access to the full peer review report through the OECD library. This includes all reviews from 1994 to 2020. The full peer review documents contain information about the in-depth assessments of one or more recipients of the reviewed donor's aid, which was collected manually by research assistants. They also allow us to classify reviews by document format, an indication of the different "eras" in which reviews were conducted under different operating guidelines. These are the data we use in the analyses of the choice of recipients for in-depth assessments.

Figure A1 illustrates the universe of peer reviews, highlighting the reviews for which we have access to the full report document, in pink.

Figure A1: Universe of OECD DAC Peer Reviews



[1] Specifically, information about review dates and examiners for older review was retrieved from the "Chronological List of Development Assistance Committee (DAC) Aid Review (Peer Review) Meetings (by year) - 1962 to 1994.pdf" document.

A2 Determinants of Peer Examiner Choice: The Data

Table A1: Definitions and Descriptive Statistics for Data on Peer Examiner Choice

Variable Name	Description	Count	Mean	SD	Min	Max
Is examiner	Binary indicator for whether a given pair of examiners conducted a given review (OECD 2020)	88374	0.006	0.076	0	1
At least one similar size	Binary indicator for whether at least one examiner is of the same size (where size is measured as being a major donor—United States, United Kingdom, Germany, France, Japan)	88374	0.827	0.378	0	1
Both similar size	Binary indicator for whether both examiners are of the same size (where size is measured as above)	88374	0.502	0.5	0	1
At least one similar EU	Binary indicator for whether at least one reviewee–examiner pair has the same status with respect to being a EU member in the year of review	88374	0.701	0.458	0	1
Both similar EU	Binary indicator for whether both examiners have the same status as the reviewee with respect to being a EU member in the year of review	88374	0.245	0.43	0	1
At least one common language	Binary indicator for whether at least one reviewer matches on an indicator of French language (in which the review may be carried out). French-speaking members are Belgium, Canada, the EU institutions, France, Luxembourg, and Switzerland. Note that a match also exists if both are English-speaking countries	88374	0.797	0.402	0	1
Both common language	Binary indicator for whether both reviewers match on the French language indicator with the reviewee	88374	0.397	0.489	0	1
At least one contiguous	Binary indicator for whether at least one reviewee–examiner pair entails states that share a border (Mayer and Zignago 2011)	84079	0.191	0.393	0	1
Both contiguous	Binary indicator for whether both reviewee–examiner pairs are contiguous	84079	0.026	0.158	0	1
At least one same model	Binary indicator for whether at least one reviewee–examiner pair has the same organizational model (DAC 2009)	88374	0.54	0.498	0	1
Both same model	Binary indicator for whether both reviewers have the same organizational model as the reviewee (DAC 2009)	88374	0.096	0.295	0	1
Review twice in a row	Binary indicator for whether both examiners also reviewed the reviewee at its immediately prior review	88374	0.189	0.391	0	1
Both before as pair	Binary indicator for whether both examiners as a pair reviewed the reviewee before at any point in time	88374	0.065	0.246	0	1
At least one before	Binary indicator for whether at least one examiner was allocated this reviewee ever before (OECD 2020)	88374	0.793	0.405	0	1
Both before	Binary indicator for whether both examiners were each (individually) allocated this reviewee ever before (OECD 2020); the theoretically interesting effect can be obtained by adding the coefficient for "one examiner reviewed reviewee" (which includes cases of two examiners)	88374	0.405	0.491	0	1
Average UNGA distance	Average of the pairwise idealpoint distances between each examiner and the reviewee, based on UN General Assembly voting behavior (Bailey, Strezhnev, and Voeten 2017)	72276	0.021	0.575	-2.444	2.216
Dispersion of UNGA distances	Dispersion of UNGA distances, defined as the difference between the larger distance and the smaller distance across both pairs; distances are absolute distances	72276	0.347	0.431	0	2.857
At least one same CPE	Binary indicator for whether at least one reviewee–examiner pair had the same "donor political economy" (or CPE), that is, whether they were both LMEs or non-LMEs (Dietrich 2021)	88374	0.861	0.346	0	1
Both same CPE	Binary indicator for whether both reviewee–examiner pairs had the same CPE (Dietrich 2021)	88374	0.593	0.491	0	1

Variable Name	Description	Count	Mean	SD	Min	Max
Average partisan distance	Average of the partisan ideology distance between each examiner and the reviewee, based on a quasi-continuous partisan ideology measure (Bjornskov and Potrafke 2013)	66575	0.421	0.295	0	1.5
Dispersion of partisan distances	Dispersion of partisan ideology distances, defined as the difference between the larger distance and the smaller distance across both pairs; distances are in absolute terms	66575	0.288	0.323	0	1.5
Average difference in deliberative democracy	Average of the differences in the V-Dem deliberative democracy index (Coppedge et al. 2016) between each examiner and the reviewee; direction of difference is preserved;	82455	0.006	0.131	-0.832	0.832
Dispersion of differences in deliberative democracy	Dispersion of differences in deliberative democracy, defined as the difference between the higher difference and the lower difference in deliberative democracy across both pairs	82455	0.085	0.123	0	0.832
Arms trade between at least one	Binary indicator for whether at least one reviewee–examiner pair had any arms trade in the year of review (SIPRI 2021)	88374	0.447	0.497	0	1
Arms trade between both	Binary indicator for whether both reviewee–examiner pairs had any arms trade in the year of review (SIPRI 2021)	88374	0.115	0.319	0	1
Alliance between at least one	Binary indicator for whether at least one reviewee–examiner pair had any alliance obligation, which includes bilateral alliances and joint membership in multilateral alliances—drawn from ATOP database (Leeds et al. 2002)	88374	0.406	0.491	0	1
Alliance between both	Binary indicator for whether both reviewee–examiner pairs had alliance obligations, which includes bilateral alliances and joint membership in multilateral alliances (Leeds et al. 2002)	88374	0.145	0.352	0	1
Diplomatic ties between at least one	Binary indicator for whether at least one reviewee–examiner pair had any diplomatic representation (Bayer 2006). Since the data end in 2005, we carried them forward based on the last available year	88374	0.693	0.461	0	1
Diplomatic ties between both	Binary indicator for whether both reviewee–examiner pairs had any diplomatic representation (Bayer 2006). Since the data end in 2005, we carried them forward based on the last available year	88374	0.623	0.485	0	1
Average ODA budget ratio	Geometric mean of the ratio of (logged) ODA budgets between the reviewee and examiners; data on ODA commitments available from DAC1 table (OECD 2019)	81216	1.041	0.235	0.41	2.476
Dispersion of ODA budget ratio	Ratio of the larger ODA budget ratio and the smaller ODA budget ratio across both reviewee–examiner pairs; underlying data available from DAC1 table (OECD 2019)	84922	1.248	0.231	1	2.757
First pair role-swopped	Binary indicator for whether the current reviewee was an examiner at the most recent (past) review in which the first examiner was being reviewed (OECD 2020)	88374	0.656	0.475	0	1
Second pair role-swopped	Binary indicator for whether the current reviewee was an examiner at the most recent (past) review in which the second examiner was being reviewed (OECD 2020)	88374	0.601	0.49	0	1
Number of available examiners	Number of potential examiners available for review, which is determined by DAC membership only (OECD 2020)	88374	19.574	4.752	10	28
Three examiners	Binary indicator for whether the review involved three examiners (OECD 2020)	88374	0.06	0.237	0	1
Review year	Year of the review	88374	1991.423	17.62	1962	2020

A3 Determinants of Peer Examiner Choice: Robustness Tests

Table A2: Determinants of Peer Examiner Choice:
Using Only "At Least One" Pair-Matching Covariates

	(A1)		(A2)		(A3)	
At least one similar size	0.536***	(0.108)			0.616***	(0.123)
At least one similar EU	0.056	(0.066)			0.072	(0.082)
At least one common language	0.131	(0.087)			0.157	(0.109)
At least one contiguous	0.000	(0.073)			-0.136	(0.091)
At least one same model	-0.007	(0.053)			-0.004	(0.068)
Review twice in a row	-0.078	(0.072)			-0.152	(0.092)
Both before as pair	0.239†	(0.129)			0.197	(0.153)
At least one before	0.075	(0.082)			0.049	(0.107)
Average UNGA distance			-0.188†	(0.109)	-0.210†	(0.112)
Dispersion of UNGA distances			-0.041	(0.085)	-0.070	(0.087)
At least one similar CPE			-0.021	(0.102)	0.016	(0.103)
Average partisan distance			0.111	(0.134)	0.125	(0.134)
Dispersion of partisan distances			0.037	(0.112)	0.048	(0.111)
Average difference in democracy			-0.659	(0.684)	-0.728	(0.697)
Dispersion democracy differences			-0.536	(0.431)	-0.471	(0.435)
Observations	84079		57015		57015	
Within R2	4.80E-04		2.30E-04		8.90E-04	

Linear regressions with review fixed effects and standard errors clustered on reviews in parentheses.

Significance levels: † <.1 * <.05 ** <.01 *** <.001

Table A3: Determinants of Peer Examiner Choice:
Using Only "Both" Pair-Matching Covariates

	(A1)		(A2)		(A3)	
Both similar size	-0.057	(0.065)			0.066	(0.097)
Both similar EU	-0.034	(0.063)			-0.104	(0.077)
Both common language	-0.048	(0.060)			-0.031	(0.078)
Both contiguous	-0.123	(0.169)			-0.022	(0.237)
Both same model	0.033	(0.092)			0.063	(0.117)
Review twice in a row	-0.084	(0.073)			-0.157†	(0.090)
Both before as pair	0.219	(0.133)			0.188	(0.158)
Both before	0.137*	(0.065)			0.147†	(0.082)
Average UNGA distance			-0.180	(0.110)	-0.194	(0.129)
Dispersion of UNGA distances			-0.047	(0.085)	0.007	(0.089)
Both similar CPE			-0.099	(0.089)	-0.083	(0.091)
Average partisan distance			0.108	(0.134)	0.121	(0.133)
Dispersion of partisan distances			0.035	(0.112)	0.056	(0.112)
Average difference in democracy			-0.663	(0.684)	-0.671	(0.701)
Dispersion democracy differences			-0.535	(0.431)	-0.467	(0.444)
Observations	84079		57015		57015	
Within R2	1.80E-04		2.50E-04		4.60E-04	

Linear regressions with review fixed effects and standard errors clustered on reviews in parentheses.

Significance levels: † <.1 * <.05 ** <.01 *** <.001

Table A4: Determinants of Peer Examiner Choice: Additional Donor Politics Variables

	(A1)		(A2)		(A3)	
Arms trade between at least one	0.094	(0.061)	0.055	(0.083)	0.081	(0.081)
Arms trade between both	0.179†	(0.104)	0.133	(0.119)	0.111	(0.120)
Alliance between at least one	0.002	(0.079)	-0.050	(0.105)	-0.051	(0.105)
Alliance between both	0.052	(0.090)	0.128	(0.117)	0.161	(0.117)
Diplomatic ties between at least one	-0.516	(0.384)	-0.468	(0.404)	-0.452	(0.403)
Diplomatic ties between both	-0.103	(0.113)	-0.079	(0.122)	-0.015	(0.122)
Average UNGA distance			-0.150	(0.115)	-0.089	(0.133)
Dispersion of UNGA distances			-0.044	(0.086)	-0.077	(0.091)
Average power difference			0.019	(0.107)	0.057	(0.107)
Dispersion in power difference			-0.109	(0.093)	-0.096	(0.096)
Average partisan distance			0.104	(0.134)	0.116	(0.133)
Dispersion of partisan distances			0.036	(0.112)	0.052	(0.112)
Average difference in democracy			-0.578	(0.742)	-0.642	(0.757)
Dispersion democracy differences			-0.543	(0.460)	-0.366	(0.471)
At least one similar size					0.643***	(0.130)
Both similar size					-0.063	(0.101)
At least one similar EU					0.110	(0.086)
Both similar EU					-0.128	(0.082)
At least one common language					0.174	(0.111)
Both common language					-0.070	(0.081)
At least one contiguous					-0.140	(0.093)
Both contiguous					0.009	(0.248)
At least one same model					-0.019	(0.071)
Both same model					0.058	(0.122)
Review twice in a row					-0.164†	(0.093)
Both before as pair					0.118	(0.160)
At least one before					-0.008	(0.110)
Both before					0.131	(0.084)
Observations	84079		57015		57015	
Within R2	1.50E-04		3.60E-04		1.13E-03	

Linear regressions with review fixed effects and standard errors clustered on reviews in parentheses.
Significance levels: † <.1 * <.05 ** <.01 *** <.001

Table A5: Determinants of Peer Examiner Choice: Alternative Measure of Donor Size

	(A4)		(A5)		(A6)	
At least one similar size (GDP-5)	0.523***	(0.086)			0.572***	(0.100)
Both similar size (GDP-5)	-0.061	(0.064)			-0.026	(0.089)
At least one similar EU	0.088	(0.068)			0.123	(0.086)
Both similar EU	-0.052	(0.066)			-0.119	(0.082)
At least one common language	0.149†	(0.089)			0.171	(0.111)
Both common language	-0.073	(0.061)			-0.071	(0.080)
At least one contiguous	0.006	(0.076)			-0.132	(0.094)
Both contiguous	-0.121	(0.177)			0.030	(0.246)
At least one same model	-0.037	(0.056)			-0.047	(0.072)
Both same model	0.031	(0.096)			0.060	(0.122)
Review twice in a row	-0.091	(0.074)			-0.154	(0.094)
Both before as pair	0.175	(0.134)			0.137	(0.159)
At least one before	0.026	(0.084)			-0.006	(0.110)
Both before	0.118†	(0.065)			0.128	(0.083)
Average UNGA distance			-0.179	(0.110)	-0.140	(0.120)
Dispersion of UNGA distances			-0.048	(0.085)	-0.037	(0.090)
Average power difference			0.008	(0.107)	0.051	(0.109)
Dispersion in power difference			-0.101	(0.093)	-0.089	(0.096)
Average partisan distance			0.108	(0.134)	0.117	(0.134)
Dispersion of partisan distances			0.035	(0.112)	0.059	(0.112)
Average difference in democracy			-0.663	(0.684)	-0.810	(0.713)
Dispersion democracy differences			-0.535	(0.431)	-0.327	(0.450)
Observations	84079		57015		57015	
Within R2	6.70E-04		2.50E-04		1.10E-03	

Linear regression with review fixed effects and standard errors clustered on reviews in parentheses.
Significance levels: † <.1 * <.05 ** <.01 *** <.001

Table A6: Determinants of Peer Examiner Choice: Alternative Measure of Donor Size

	(A7)		(A8)		(A9)	
Average ODA budget ratio	-1.050***	(0.198)			-1.567***	(0.327)
Dispersion of ODA budget ratio	0.106	(0.109)			0.363†	(0.191)
At least one similar EU	0.116†	(0.068)			0.161†	(0.086)
Both similar EU	-0.005	(0.068)			-0.086	(0.084)
At least one common language	0.193*	(0.090)			0.265*	(0.109)
Both common language	-0.032	(0.064)			-0.083	(0.083)
At least one contiguous	-0.051	(0.078)			-0.177†	(0.096)
Both contiguous	-0.241	(0.169)			-0.135	(0.228)
At least one same model	0.016	(0.057)			-0.022	(0.073)
Both same model	0.048	(0.101)			0.064	(0.126)
Review twice in a row	-0.107	(0.078)			-0.141	(0.098)
Both before as pair	0.174	(0.137)			0.152	(0.161)
At least one before	-0.000	(0.091)			-0.014	(0.117)
Both before	0.022	(0.073)			0.055	(0.090)
Average UNGA distance			-0.179	(0.110)	0.011	(0.140)
Dispersion of UNGA distances			-0.048	(0.085)	-0.114	(0.096)
Average power difference			0.008	(0.107)	0.045	(0.109)
Dispersion in power difference			-0.101	(0.093)	-0.092	(0.096)
Average partisan distance			0.108	(0.134)	0.126	(0.137)
Dispersion of partisan distances			0.035	(0.112)	0.082	(0.114)
Average difference in democracy			-0.663	(0.684)	-0.644	(0.834)
Dispersion democracy differences			-0.535	(0.431)	0.195	(0.726)
Observations	77726		57015		54463	
Within R2	5.50E-04		2.50E-04		9.00E-04	

Linear regression with review fixed effects and standard errors clustered on reviews in parentheses.
Significance levels: † <.1 * <.05 ** <.01 *** <.001

Table A7: Determinants of Peer Examiner Choice: Random Effects and Additional Controls

	(A10)		(A11)		(A12)	
At least one similar size	0.373***	(0.075)			0.480***	(0.107)
Both similar size	-0.174**	(0.064)			-0.207*	(0.087)
At least one similar EU	0.067	(0.064)			0.095	(0.081)
Both similar EU	-0.054	(0.064)			-0.125	(0.080)
At least one common language	0.119†	(0.070)			0.124	(0.089)
Both common language	-0.079	(0.059)			-0.085	(0.075)
At least one contiguous	0.009	(0.067)			-0.091	(0.082)
Both contiguous	-0.113	(0.175)			0.062	(0.245)
At least one same model	-0.007	(0.054)			-0.001	(0.069)
Both same model	0.035	(0.095)			0.064	(0.120)
Review twice in a row	-0.110	(0.074)			-0.114	(0.091)
Both before as pair	0.186	(0.134)			0.159	(0.158)
At least one before	0.029	(0.065)			-0.014	(0.087)
Both before	0.099	(0.061)			0.130†	(0.075)
Available examiners	0.002	(0.063)	-0.125	(0.093)	-0.160	(0.130)
Three examiners	1.596	(1.079)	0.323	(1.034)	-0.127	(1.449)
Average UNGA distance			-0.065	(0.041)	0.057	(0.048)
Dispersion of UNGA distances			-0.022	(0.079)	-0.080	(0.089)
Average power difference			0.021	(0.097)	0.058	(0.099)
Dispersion in power difference			-0.110	(0.088)	-0.101	(0.091)
Average partisan distance			0.096	(0.098)	0.100	(0.097)
Dispersion of partisan distances			0.040	(0.108)	0.054	(0.108)
Average difference in democracy			-0.278	(0.272)	-0.357	(0.278)
Dispersion democracy differences			-0.696*	(0.313)	-0.442	(0.317)
Observations	84079		57015		57015	
Within R2	5.50E-04		2.20E-04		9.20E-04	

Linear regression with random effects, year effects, and standard errors clustered on reviews in parentheses.
Significance levels: † <.1 * <.05 ** <.01 *** <.001

Table A8: Determinants of Peer Examiner Choice: Review Interdependence

	(A13)		(A14)		(A15)	
At least one similar size	0.571***	(0.130)			0.647***	(0.152)
Both similar size	-0.122†	(0.069)			-0.090	(0.099)
At least one similar EU	0.072	(0.069)			0.106	(0.085)
Both similar EU	-0.051	(0.067)			-0.121	(0.075)
At least one common language	0.155†	(0.084)			0.179†	(0.105)
Both common language	-0.070	(0.051)			-0.063	(0.070)
At least one contiguous	-0.003	(0.080)			-0.132	(0.098)
Both contiguous	-0.138	(0.168)			0.004	(0.264)
At least one same model	-0.015	(0.065)			-0.016	(0.087)
Both same model	0.031	(0.079)			0.065	(0.103)
Review twice in a row	-0.115	(0.080)			-0.156†	(0.086)
Both before as pair	0.157	(0.129)			0.121	(0.155)
At least one before	0.034	(0.078)			-0.006	(0.103)
Both before	0.122†	(0.067)			0.136†	(0.075)
Average UNGA distance			-0.179†	(0.103)	-0.111	(0.121)
Dispersion of UNGA distances			-0.048	(0.081)	-0.085	(0.085)
Average power difference			0.008	(0.099)	0.043	(0.101)
Dispersion in power difference			-0.101	(0.100)	-0.087	(0.104)
Average partisan distance			0.108	(0.141)	0.118	(0.141)
Dispersion of partisan distances			0.035	(0.119)	0.056	(0.116)
Average difference in democracy			-0.663	(0.462)	-0.729	(0.494)
Dispersion democracy differences			-0.535	(0.331)	-0.339	(0.353)
Observations	84079		57015		57015	
Within R2	6.10E-04		2.50E-04		1.02E-03	

Linear regression with review fixed effects and standard errors clustered on years in parentheses.
Significance levels: † <.1 * <.05 ** <.01 *** <.001

Table A9: Determinants of Peer Examiner Choice: Added Review Characteristics

	(A16)		(A17)		(A18)	
Highest unemployment	-0.002	(0.003)	-0.007	(0.005)	-0.005	(0.005)
Lowest growth	-0.013	(0.019)	-0.027	(0.024)	-0.028	(0.024)
Highest corruption	-0.272	(0.266)	-0.176	(0.391)	-0.204	(0.400)
At least one similar size	0.404**	(0.139)			0.506**	(0.161)
Both similar size	-0.097	(0.071)			-0.043	(0.111)
At least one similar EU	0.189*	(0.075)			0.218*	(0.095)
Both similar EU	0.015	(0.072)			-0.033	(0.095)
At least one common language	0.151†	(0.090)			0.173	(0.110)
Both common language	0.042	(0.068)			0.067	(0.093)
At least one contiguous	-0.086	(0.079)			-0.180†	(0.100)
Both contiguous	-0.047	(0.205)			-0.126	(0.250)
At least one same model	0.073	(0.060)			0.009	(0.081)
Both same model	-0.048	(0.102)			-0.046	(0.127)
Review twice in a row	-0.034	(0.089)			-0.157	(0.114)
Both before as pair	0.161	(0.151)			0.120	(0.178)
At least one before	0.107	(0.078)			0.124	(0.105)
Both before	0.149*	(0.073)			0.118	(0.093)
Average UNGA distance			-0.265*	(0.129)	-0.270†	(0.151)
Dispersion of UNGA distances			-0.091	(0.095)	-0.132	(0.104)
Average power difference			-0.010	(0.107)	0.007	(0.108)
Dispersion in power difference			-0.073	(0.093)	-0.027	(0.095)
Average partisan distance			0.126	(0.150)	0.147	(0.147)
Dispersion of partisan distances			0.134	(0.124)	0.129	(0.124)
Average difference in democracy			-0.581	(0.898)	-0.459	(0.932)
Dispersion democracy differences			-0.430	(0.643)	-0.419	(0.675)
Observations	60036		40289		40289	
Within-R2	7.20E-04		3.50E-04		1.10E-03	

Linear regression with review fixed effects and standard errors clustered on reviews in parentheses.
Significance levels: † <.1 * <.05 ** <.01 *** <.001

Table A10: Determinants of Peer Examiner Choice: Additional Potential Determinants

	(A19)		(A20)		(A21)	
First pair role-swopped	0.230***	(0.057)	0.207***	(0.062)	0.209**	(0.079)
Second pair role-swopped	0.293***	(0.055)	0.263***	(0.061)	0.320***	(0.076)
At least one similar size			0.548***	(0.111)	0.626***	(0.130)
Both similar size			-0.121†	(0.066)	-0.098	(0.102)
At least one similar EU			0.066	(0.068)	0.091	(0.086)
Both similar EU			-0.044	(0.066)	-0.118	(0.081)
At least one common language			0.165†	(0.089)	0.190†	(0.111)
Both common language			-0.069	(0.061)	-0.070	(0.080)
At least one contiguous			-0.001	(0.076)	-0.112	(0.093)
Both contiguous			-0.134	(0.177)	0.015	(0.245)
At least one same model			-0.019	(0.055)	-0.029	(0.071)
Both same model			0.027	(0.096)	0.051	(0.122)
Review twice in a row			-0.106	(0.075)	-0.152	(0.093)
Both before as pair			0.148	(0.135)	0.112	(0.160)
At least one before			-0.027	(0.087)	-0.060	(0.113)
Both before			0.018	(0.070)	0.036	(0.087)
Average UNGA distance					-0.076	(0.130)
Dispersion of UNGA distances					-0.063	(0.091)
At least one same CPE					0.035	(0.108)
Both same CPE					-0.078	(0.096)
Average partisan distance					0.109	(0.133)
Dispersion of partisan distances					0.064	(0.111)
Average difference in democracy					-0.624	(0.709)
Dispersion democracy differences					-0.332	(0.449)
Observations	84079		84079		57015	
Within R2	4.00E-04		8.80E-04		1.34E-03	

Linear regression with review-fixed effects and standard errors clustered on reviews in parentheses.
Significance levels: † <.1 * <.05 ** <.01 *** <.001

A4 Determinants of Recipient Choice: The Data

A4.1 Data on Official Development Assistance

The data on foreign aid flows used in this paper come from the OECD DAC Creditor Reporting System (CRS). The version of these data used in this draft was updated on the OECD.Stat website on April 28 and downloaded by us on May 6, 2021.[2] We aggregate the CRS's project-level information into yearly figures by donor, recipient, and donor-recipient pairs as necessary. In our aggregation, we only consider flows marked as "ODA Grants" or "ODA Loans"—thus excluding flows marked as "Equity Investment," "Private Development Finance," and "Other Official Flows (non Export Credit)" from our analyses. The resulting data includes information on 47 years, from 1973 to 2019; 95 donors; and 203 recipients (note that 21 of the recipients are "regional" and two are "unspecified").

We use commitment data in million US dollars, 2019 constant amounts (variable `USD_Commitment_Defl`). Out of the 3,986,771 project-level observations from 1990 to 2019, 265 have a negative commitment amount (0.0066%); 802,917 have a commitment amount equal to zero (20%); and 1,013,029 are listed as not available (25%). As we aggregate amounts from the project level to construct donor-year, recipient-year, and donor-recipient-year totals, we sum across all projects—including those with zero or negative amounts—while the projects with no commitment amount information are dropped.

[2] `https://stats.oecd.org/DownloadFiles.aspx?DatasetCode=CRS1`

A4.2 Data on Recipient Characteristics

A4.2.1 Quality of Government (QoG) Data

Many of the recipient characteristics included in our models are data from the Quality of Government (QoG) Basic Dataset 2021.[3] Specifically, we use the tenfold classification of colonial origin from Wahman, Teorell and Hadenius (Hadenius and Teorell, 2007; Teorell and Wahman, 2018; Wahman et al., 2013)—variable `ht_colonial` in the QoG Basic Dataset 2021—to construct our indicator of colonial history, which equal to one for recipients who were primarily a former colony of the donor, and zero otherwise.

We measure democracy with the Revised Combined Polity Score (`p_polity2`) from the Polity project (Marshall and Gurr, 2020); and government effectiveness (`wbgi_gee`), rule of law (`wbgi_rle`), and political stability and absence of violence/terrorism (`wbgi_pve`) with estimates from the World Bank's Governance Indicators. We also use measures of life expectancy at birth (`wdi_lifexp`) and GDP per capita in constant 2010 US dollars (`wdi_gdpcapcon2010`) from the Bank's World Development Indicators.

Note that, when including these covariates, we lose observations from the 22 recipients who are not countries in the QoG data.[4] This is a problem for Kosovo, which was visited once by Austria in 2020; and for the West Bank and Gaza Strip, which was visited by Italy in 2000. Since Italy also visited Ethiopia in 2000, that review remains in the sample. However, Austria only visited Kosovo in 2020—this means that lines from the Austria 2020 review remain in the sample for our models, but no recipient is selected with a value of one for the dependent variable once the Austria-2020-Kosovo observation is dropped.

A4.2.2 UN Security Council Data

We include an indicator for whether a recipient country holds one of the rotating positions in the UN Security Council in the year of a review. These data are from Dreher, Sturm, and Vreeland (2009) and were updated by Dreher, Lang, Rosendorff, and Vreeland (2018).[5]

Note that the following recipients do not appear in these data: Anguilla, Aruba, Bermuda, British Virgin Islands, Cayman Islands, China (People's Republic of), Cook Islands, French Polynesia, Gibraltar, Hong Kong (China), Kosovo, Macau (China), Mayotte, Montenegro, Montserrat, Netherlands Antilles, New Caledonia, Niue, Northern Mariana Islands, Saint Helena, Serbia, South Sudan, Timor-Leste, Tokelau, Turks and Caicos Islands, Tuvalu, Wallis and Futuna, West Bank and Gaza Strip.

This means that, when including this indicator, we lose relevant observations for recipients who were

[3] Retrieved on May 3, 2021, from `http://www.qog.pol.gu.sedoi:10.18157/qogbasjan21`.

[4] Anguilla, Aruba, Bermuda, British Virgin Islands, Cayman Islands, Cook Islands, French Polynesia, Gibraltar, Hong Kong (China), Kosovo, Macau (China), Mayotte, Montserrat, Netherlands Antilles, New Caledonia, Niue, Northern Mariana Islands, Saint Helena, Tokelau, Turks and Caicos Islands, Wallis and Futuna, West Bank and Gaza Strip.

[5] The data were last updated on August 4, 2020, and retrieved by us on May 10, 2021, from `https://www.uni-heidelberg.de/fakultaeten/wiso/awi/professuren/intwipol/datasets_en.html`.

visited in the cases of Kosovo and West Bank and Gaza Strip (already discussed); but also Montenegro and Timor-Leste.

A4.2.3 UN General Assembly Voting Data

We include ideal point distances between the reviewed donors and their aid recipients in UN General Assembly voting. These data are from Voeten, Strezhnev, and Bailey (2009) and were updated in July 2023.[6]

Similar to the UNSC data, recipients who are not UN members do not appear in these data: Anguilla, Aruba, Bermuda, British Virgin Islands, Cayman Islands, Chinese Taipei, Cook Islands, French Polynesia, Gibraltar, Hong Kong (China), Kosovo, Macau (China), Mayotte, Montserrat, Netherlands Antilles, New Caledonia, Niue, Northern Mariana Islands, Saint Helena, Serbia, Tokelau, Turks and Caicos Islands, Wallis and Futuna, West Bank and Gaza Strip.

A4.2.4 World Bank's World Development Indicators Data

We include information on recipients' tourism and statistical capacity from the World Bank's World Development Indicators (WDIs). To measure a recipient's level of tourism, we use the number of tourists who travel to a country other than that in which they have their usual residence, but outside their usual environment, for a period not exceeding 12 months and whose main purpose in visiting is other than an activity remunerated from within the country visited ("International inbound tourists (overnight visitors)" indicator). These data are originally from the *Yearbook of Tourism Statistics* of the World Tourism Organization, and they were last updated on June 30, 2021.

To measure statistical capacity in the recipient country, we use the overall average Statistical Capacity score from the World Bank's *Bulletin Board on Statistical Capacity*, last updated on July 30, 2021. This indicator is a composite score assessing the capacity of a country's statistical system. It is based on a diagnostic framework assessing the following areas: methodology; data sources; and periodicity and timeliness. Countries are scored against 25 criteria in these areas, using publicly available information and/or country input. The overall Statistical Capacity score is then calculated as a simple average of all three area scores on a scale of 0–100.

A4.3 Data on Recipient Assessments

We retrieved information on recipient-specific assessments of the donor's development assistance program from the peer review reports. As depicted in Figure A1, we have access to the full report document for 143

[6]The data were retrieved by us on November 2, 2023, from `https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/LEJUQZ`.

peer reviews, 1994–2020; and 133 of them mention one or more recipient-specific assessments. The following figures provide additional descriptive information about these data.

Figure A2 lists all recipients subject to at least one field visit in the sample of 133 reviews (1994–2020) for which we have visit information. Contrary to what is prescribed in the peer review methodology, it is not uncommon for the same recipient to receive visits for reviews in back-to-back years; and in six instances, the same recipient was visited for more than one review in the same year, although this has not happened since 2003.

Figure A2: Recipient Visit Count by Year. Universe of 211 recipient-specific assessments conducted across our sample of 143 reviews, 1994–2020. Each data point represents one recipient-specific assessment, with dot size and label highlighting recipients who were visited for more than one review in the same year.

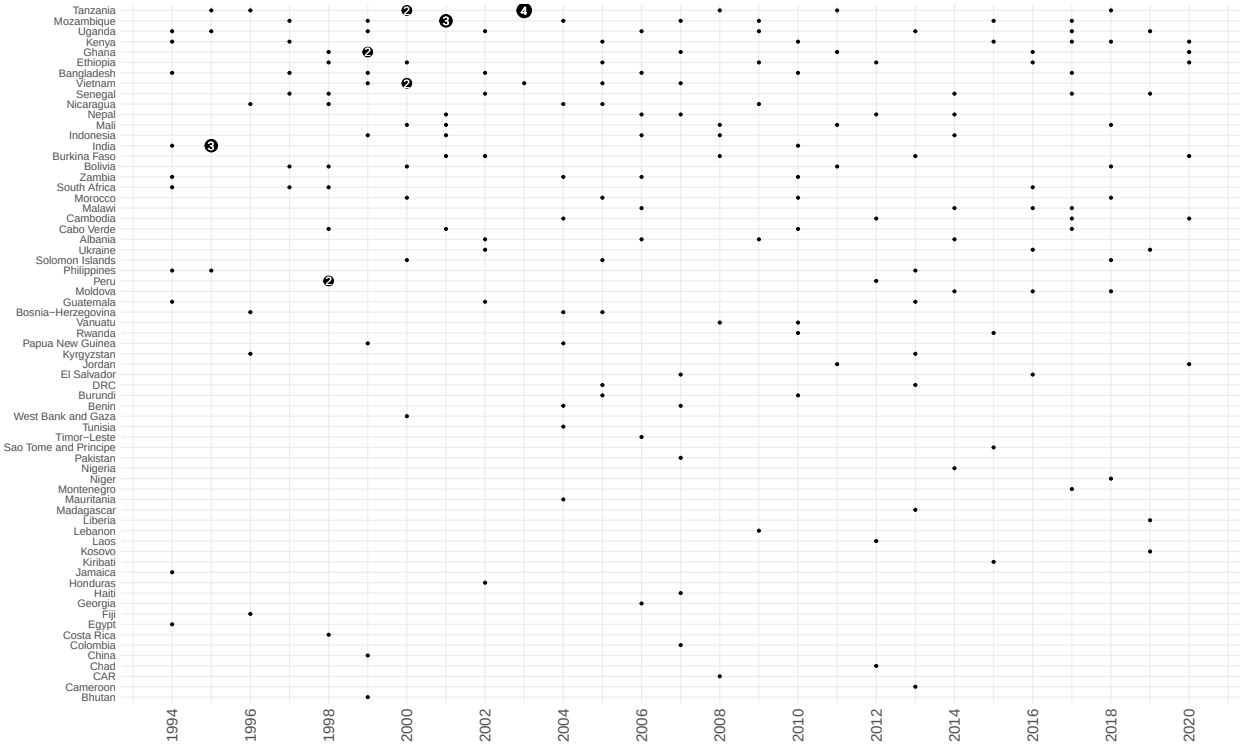


Figure A3: Number of Recipient-Specific Assessments per Donor

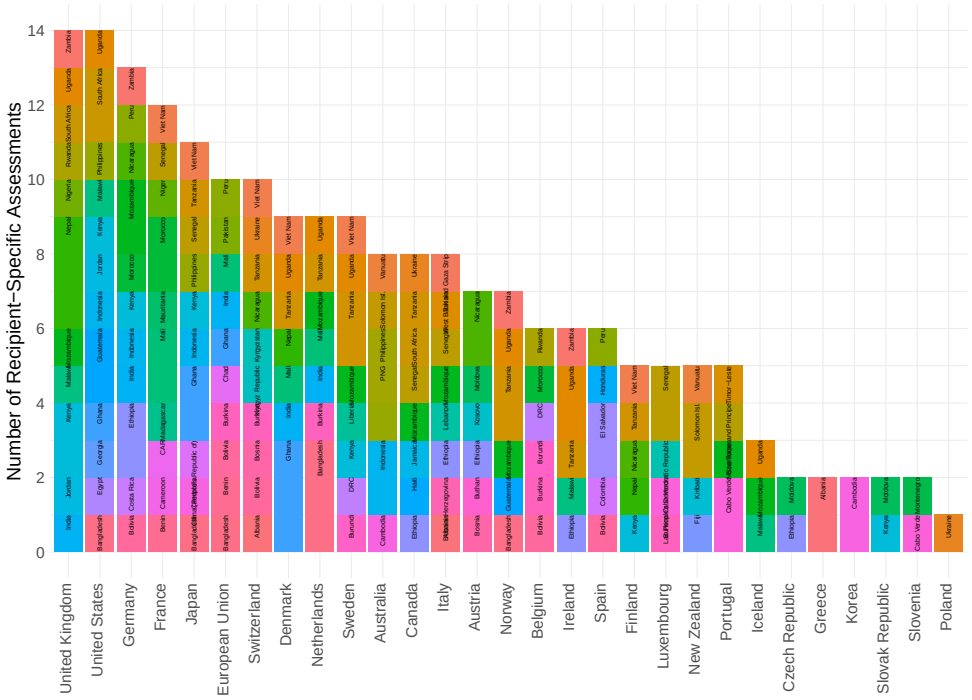


Figure A4: Number of Recipient-Specific Assessments per Recipient

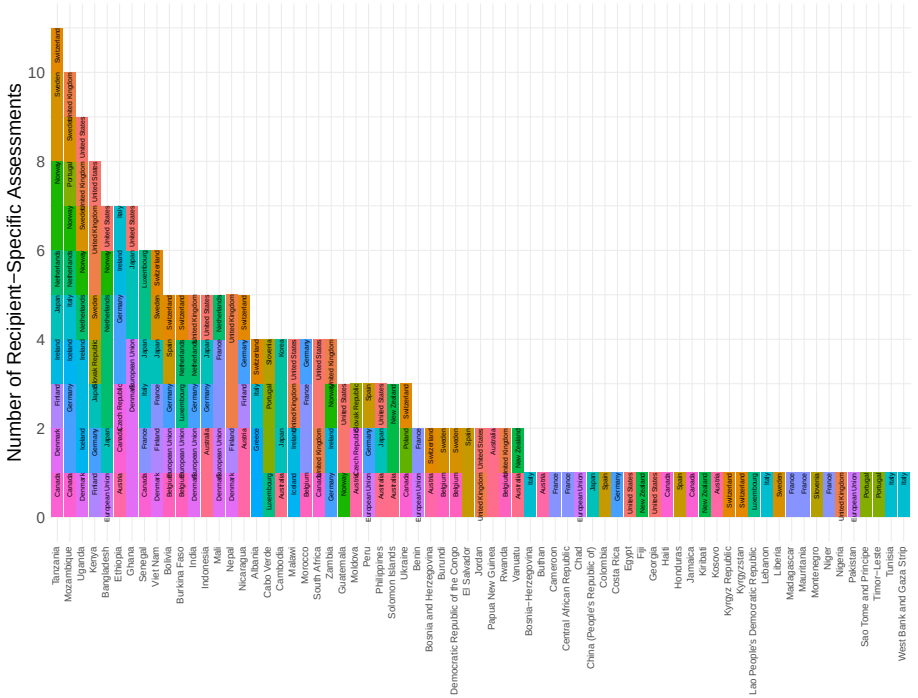


Figure A5: Number of Recipient-Specific Assessments per Review, by Type

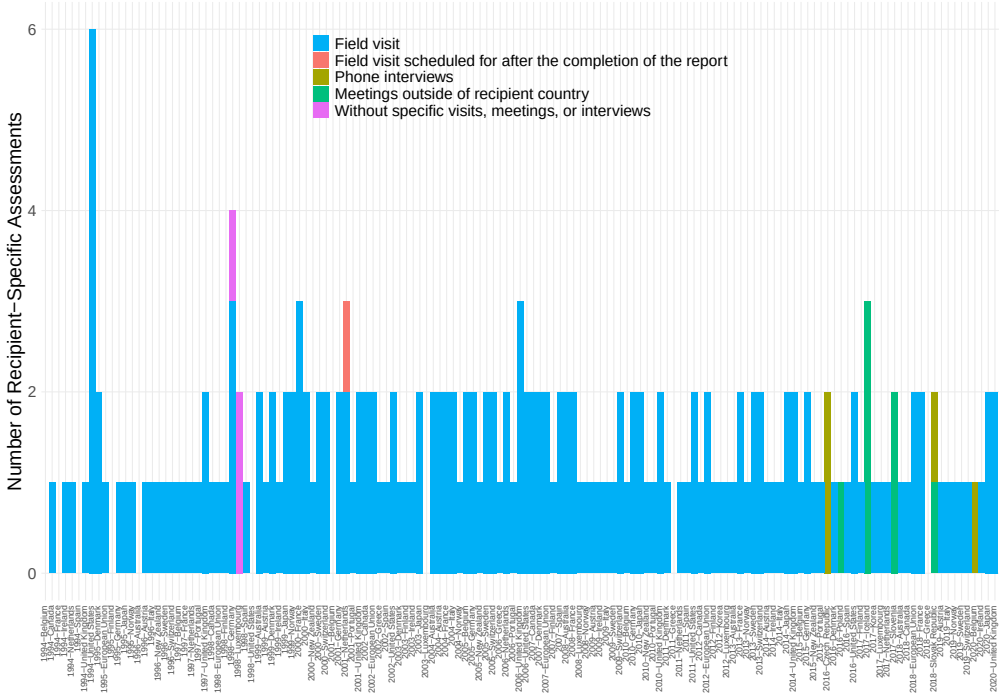
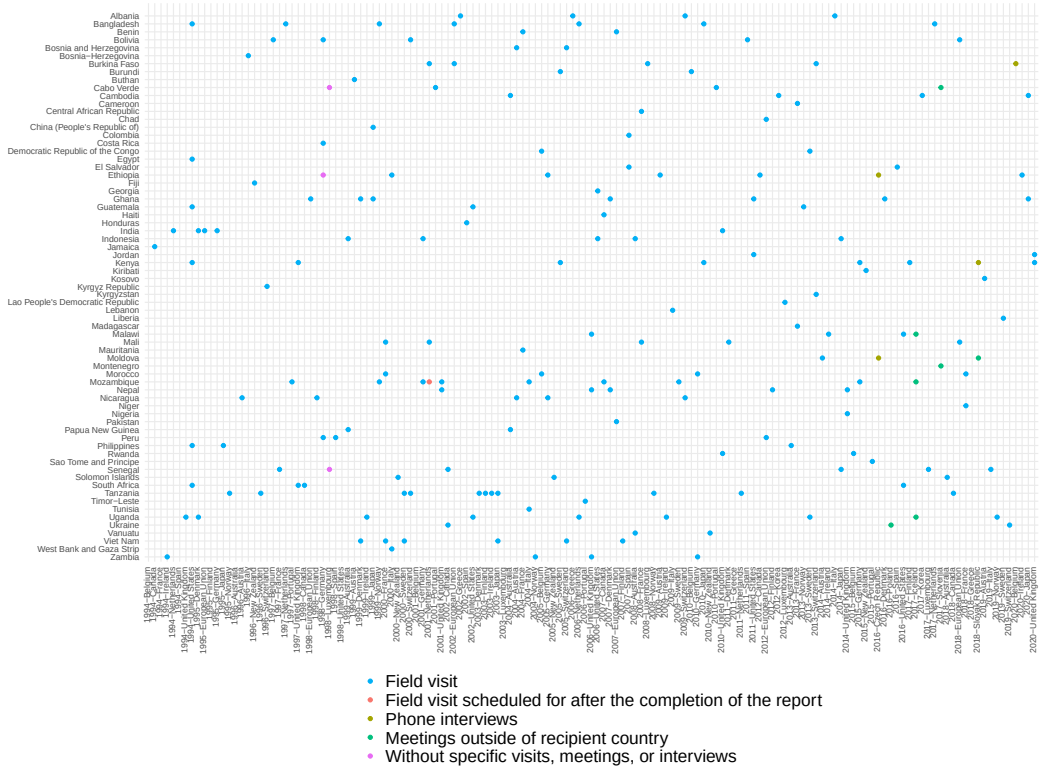


Figure A6: Universe of Recipient-Specific Assessments



A5 Heterogeneity Across Member States

Table A11: Member State Characteristics Across Different Peer Review Mechanisms

DAC Peer Review

	Members		Non-members		Difference	Significance
	Mean	SD	Mean	SD	(in means)	(p<0.05)
UNGA distance to G7	-0.472	0.227	-1.789	0.670	1.317	*
VDem Polyarchy Index	0.834	0.078	0.507	0.209	0.327	*
VDem Gender Equality Index	0.921	0.038	0.745	0.143	0.176	*
VDem Corruption Index	0.134	0.123	0.568	0.245	-0.434	*
Logged GDP per capita	10.687	9.949	8.858	9.264	1.830	*
KOF Political Globalization Index	89.175	8.103	66.533	16.287	22.642	*
Economic Freedom Index	72.017	7.048	59.367	8.676	12.651	*

Anti-Bribery Convention

	Members		Non-members		Difference	Significance
	Mean	SD	Mean	SD	(in means)	(p<0.05)
UNGA distance to G7	-0.692	0.652	-1.897	0.554	1.205	*
VDem Polyarchy Index	0.779	0.153	0.481	0.197	0.298	*
VDem Gender Equality Index	0.882	0.108	0.734	0.139	0.148	*
VDem Corruption Index	0.211	0.185	0.602	0.232	-0.391	*
Logged GDP per capita	10.376	10.010	8.704	9.270	1.671	*
KOF Political Globalization Index	86.257	9.844	64.375	15.798	21.883	*
Economic Freedom Index	70.042	7.708	58.285	8.254	11.757	*

Universal Periodic Review

	Members		Non-members		Difference	Significance
	Mean	SD	Mean	SD	(in means)	(p<0.05)
UNGA distance to G7	-1.470	0.782	-1.621	0.789	0.151	
VDem Polyarchy Index	0.588	0.239	0.548	0.223	0.040	
VDem Gender Equality Index	0.778	0.153	0.771	0.145	0.006	
VDem Corruption Index	0.476	0.279	0.508	0.281	-0.032	
Logged GDP per capita	9.637	9.919	9.390	9.788	0.247	
KOF Political Globalization Index	77.617	13.032	67.277	18.028	10.340	*
Economic Freedom Index	62.070	9.442	61.125	9.712	0.945	

Appendix Bibliography

Bayer R. 2006. "Diplomatic Exchange Data Set, v2006-1." Accessed June 15, 2021. `http://correlatesofwar.org`

Bjørnskov C. Potrafke N., 2013. "The Size and Scope of Government in the US States: Does Party Ideology Matter?" *International Tax and Public Finance* 20: 687–714.

Coppedge M. Gerring J., Lindberg S.I., Skaaning S.E., Teorell J., Altman D., Bernhard M., et al. 2016. V-Dem Dataset v6.2.

Dreher A., Lang V., Rosendorff P., Vreeland J. R.. 2018. "Buying Votes and International Organizations: The Dirty Work-Hypothesis." CEPR Discussion Paper No. 13290.

Leeds B., Ritter J., Mitchell S., Long A. 2002. "Alliance Treaty Obligations and Provisions, 1815–1944." *International Interactions* 28 (3): 237–60.

Mayer T., Zignago S. 2011. "Notes on CEPII's Distances Measures: The GeoDist Database." Working Papers 2011–25, CEPII Research Center.

SIPRI. 2021. SIPRI Arms Transfers Database (TIV of Arms Exports from All). Accessed June 15, 2021. `http://www.sipri.org/databases/armstransfers/sources-and-methods/`