

Online Appendix for “Elusive Collaboration? The Determinants of Lead Donorship in International Development”

Miles D. Williams

Appendix A

The measure of lead donorship provides objective and consistent criteria for quantifying lead donorship, but looking at a set of examples can help make the implications of these criteria more concrete. As a matter of observation, identifying clear cases of lead donorship (narrowly defined in terms of long-term high donor concentration) should be easy enough—so, too, with identifying clear cases of its absence. To illustrate, consider Figure A1, which shows the share of aid committed by the top DAC donor in a given year to two aid recipients: Mexico (left) and Mozambique (right). The x-axes of each panel in the figure denote years from 1996 to 2014. The y-axes show the percent of total aid going to a recipient that the top donor is responsible for. These data come from the OECD’s Creditor Reporting System (CRS) and reflect the total of bilateral Official Development Assistance (ODA) commitments in 2017 constant US dollars across the 24 main development sectors as defined by the OECD. Abbreviated country names are included so it is clear which donor is responsible for committing the top share of aid. Also, lighter columns have been added that indicate the share of aid committed by the donor responsible for the second greatest amount in a given year. This provides a point of reference for how much more the top donor in a given year committed relative to the next most generous donor.

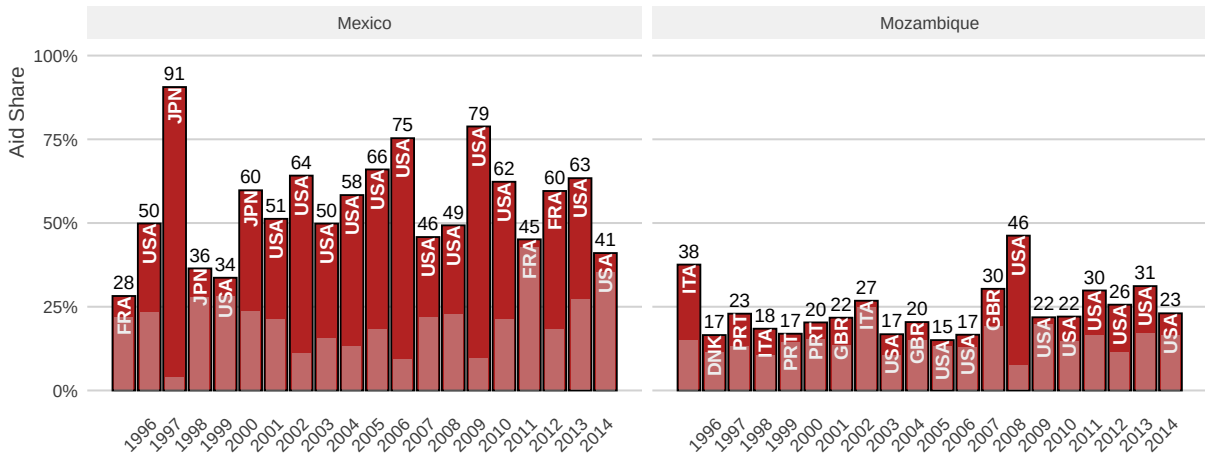


Figure A1: Aid shares of top donors to Mexico and Mozambique.

Among the 29 DAC countries committing foreign aid across 24 development sectors to Mexico, the United States held top donor status for 14 out of 19 years from 1996 to 2014. The US further bore responsibility for not just the plurality, but the majority of aid for eight of those years. This comes as little surprise. Mexico shares a border with the United States and has greater long-term security, diplomatic, and migration policy importance for the US government than for other DAC members.

Conversely, over this same period the share of aid committed by the top DAC donor per year to Mozambique was 23% on average, declining to as low as 15% in 2005, and only climbing as high as 46% in 2008 before falling to 22% the next

year. While the United States committed the largest share of aid for several years, top donor status fluctuated among a group of donors from 1996 to 2007. And, even though the US committed the most aid to Mozambique from 2008 to 2014, in nearly all of those years (2008 being the exception), the second most generous donor did not fall far behind.¹ While Portugal formerly colonized Mozambique—making it a plausible contender for lead donor given the role of colonial legacy in driving aid giving (Chiba and Heinrich 2019)—it only held top-donor status in 1997 and again in 1999 and 2000.

The United States arguably qualifies as a lead donor in the first case (Mexico), while no obvious lead donor exists in the second (Mozambique). Of course, not all cases can be categorized so decisively. The question of where to draw the line is unavoidably subjective and arbitrary, which poses a challenge to setting down criteria for lead donorship. However, Steinwand (2015) provides some criteria that make it easy to rigorously and consistently quantify lead donorship status. As noted in the main text, these criteria are:

1. The Herfindahl Index (HI) for a given recipient in a given year is greater than the median HI for the data sample;
2. The difference between the total amount of aid given by the top donor relative to the next largest donor to a recipient in a year is greater than the sample median of this difference;
3. The share of aid to a recipient in a given year from the top donor is greater than the sample median of aid shares.

Lead donorship also has a temporal dimension, which Steinwand (2015) captures using the following additional criteria:

4. A donor must have top donor status for at least 5 out of 9 consecutive years;
5. A donor must not drop from top donor status for more than 2 consecutive years in that 9 year time span.

In devising a measure of lead donorship for this study, I adopt some, but not all, of these decision rules. In particular, I rely on the first three while I opt for a slightly more flexible approach to deal with the longitudinal dimension of lead donorship. Instead of the year-based criteria, I use a simple lagged rolling average of lead donorship:

For a given year, I calculate the proportion of times a donor meets the criteria specified in points 1-3 in that year and the two previous.

While certain features of this measurement strategy are obviously arbitrary, they do a fair job of identifying cases where, on its face, lead donorship is present. Consider again the aid recipients Mexico and Mozambique. Figure A2 reports the aid shares of the top donors to each recipient per year from 1998 to 2014—1996 and 1997 now have to be dropped. In addition, it reports the lead donorship status of the top aid donors by the column shades. The darker the shade, the stronger the measure of lead donorship.

As it turns out, the measurement strategy does a fair job of identifying lead donors status. In the case of Mexico, the leadership of the United States is strongest precisely when it should be based on its giving. Further, when its giving wanes, this is reflected in the US having a weaker measure of lead donorship or even being supplanted by other donors in a few years. Japan, for example has a modest degree of lead donorship in 1998 and 2000 (1/3), as does France in 2012 (1/3). Even so, as we would intuitively expect, the United States has non-zero lead donorship status for 12 of the 17 years from 1998 to 2014 and has full lead donorship status from 2003 to 2006 and again in 2010.

Conversely, no donor rises to even a modest level of lead donorship status in Mozambique for the entire period from 1998 to 2014. Even from 2008 to 2014 when the US consistently holds the position of top donor, its level of giving is insufficient to earn it the lead donorship title. This is as would be expected given the historical widespread attention Mozambique receives from the donor community. In such a fragmented donor environment, lead donorship remains elusive.

¹The sizeable and rapid jump in US support in 2008 coincided with the onset of a devastating flood in Mozambique in early 2008 that killed several and displaced tens of thousands. Total DAC giving to Mozambique increased to 1.78 billion from 1.5 billion 2017 constant US dollars this same year. It is interesting to note that the US was responsible for much of this increase. However, in the following years, in addition to aid totals slowly declining, the US clearly began to bear less responsibility for the total amount given.

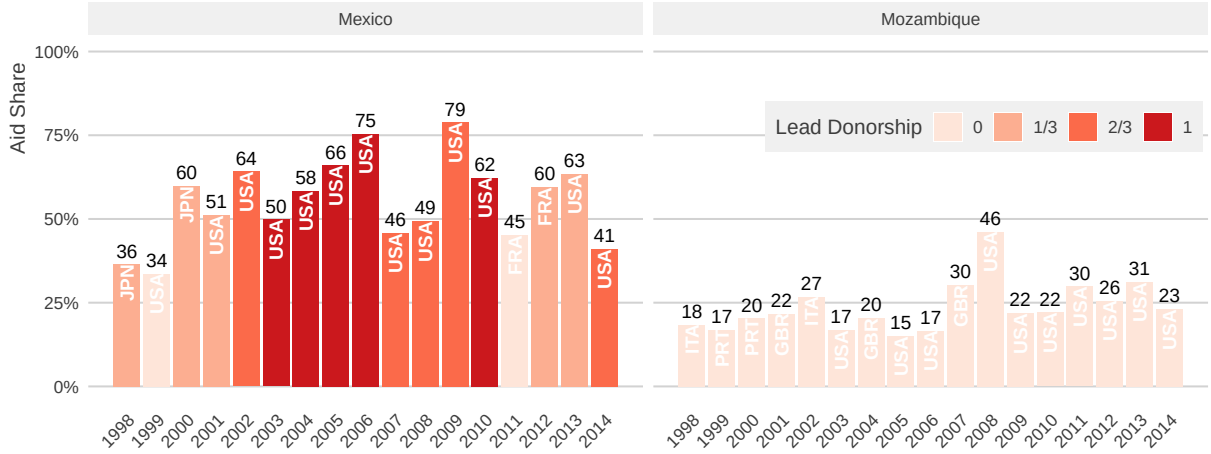


Figure A2: Lead donorship of top donors to Mexico and Mozambique.

Appendix B

This section provides a summary of the method used to construct the measures of donor-recipient ties and recipient need. Define a vector of $K > 1$ covariates as $(X_{i1}, \dots, X_{ik}, \dots, X_{iK}) = X_i \in X \subseteq \mathbb{R}^K$ and an objective function $f : \mathbb{R}^K \rightarrow \mathbb{R}$. Subscript i indexes N units of observation. The goal with the objective function f is to identify a vector of weights $(\omega_1, \dots, \omega_k, \dots, \omega_K) = \omega \in \Omega \subseteq \mathbb{R}^K$ that optimize f given X_i . Denote $\hat{\omega}$ as the solution set of weights:

$$\hat{\omega} = \underset{\omega \in \Omega}{\operatorname{argmax}} f(\omega; X_i), \quad (\text{A1})$$

and define the vector Z_i as the optimal linear combination of the variables:

$$Z_i = X_i^\top \hat{\omega}. \quad (\text{A2})$$

Several functional forms might do for f , but one of the simpler is the sum of the squared covariances or SSC:

$$f(\omega; X_i) = \sum_{k=1}^K \operatorname{cov}(X_{ik}; Z_i)^2. \quad (\text{A3})$$

Squaring ensures that greater values indicate a stronger positive correlation between a variable k and the linear combination Z_i while also allowing f to be continuously differentiable, which is a nice property for optimization.

In words, this approach finds a linear combination of K covariates such that said linear combination maximally covaries with each individual variable k . The end product is a variable Z that maximally covaries with the set of components used in its construction.

Because scale may be an issue when combining covariates, in practice it is advisable to standardize each of the covariates in the set X_i . The most commonly applied method is to mean center and scale covariates such that they have a standard deviation of 1. After the linear combination Z is generated, it again is desirable to standardize the final output. Thus, Z_i is simply a z-score representing variation in the composite measure of interest.

Because much has been written about other existing approaches of dimensionality reduction (or linear projection), a comprehensive summary of each is beyond the scope of this paper. However, two popular alternatives are worth mentioning: principal components analysis (PCA) and multivariate factor analysis (MFA). These approaches are applied across a range of scientific fields to produce a smaller set of covariates from a much larger set for use in regression analysis. Why develop and use SSC when these other approaches are available?

The answer to this question has both a theoretical and a practical element. With respect to the first, it is necessary to briefly discuss the intuitions behind PCA and MFA. The goal of the former is to find a multi-dimensional ellipsoid with axes that represent “principal components.” Using the notation from above, the first such component is identified for a set of covariates X_i by generating a set of weights where:

$$\hat{\omega} = \operatorname{argmax}_{\omega \in \Omega} \sum_{i=1}^N (X_i^\top \omega)^2. \quad (\text{A4})$$

Second-, third-, and k -order principal components are identified by first subtracting out the first and then repeating the procedure again in successive fashion.

Keeping with the first principal component (since this is what we would use to generate the composite measures of interest), this method, in words, finds the variance maximizing set of weights. This approach obviously differs from that of SSC, which finds weights such that the linear combination of covariates maximally covaries with each component variable. This difference implies different kinds of information from the data that the resulting linear combinations capture. SSC provides a measure that is maximized to be correlated with its components, while PCA provides a measure that has maximum variance given the component measures. The former objective is better aligned with the goal of measurement creation.

MFA, unlike PCA, takes into account covariances in the data in the construction of “factors.” MFA models observed covariates in X_i as linear functions of latent factors F_i where the number of factors is less than K . The idea is that groups of variables may move in common directions because they are the product of a smaller set of unobserved variables. For instance, variables like income, education, and employment status could be thought of as determined by a latent concept called socioeconomic status.

In some ways MFA is closer in spirit to SSC than is PCA, but it still optimizes on a different objective, namely the extent to which variables are correlated *with each other*. This means that using the first factor from MFA (which we might try to do in constructing measures of interest for this study) may not capture all the variation in a concept that is of interest. If two variables are part and parcel of the same overall concept, but are not themselves strongly correlated, one of these may receive a much lower factor loading and thus contribute much less to the variation in the estimated factor.

To assist in validating SSC and justifying its use as opposed to PCA or MFA, a simple predictive analysis is performed to compare versions of the main measures of interest (recipient need or RN and donor interest or DI) as calculated using SSC, PCA, and MFA.

To test the performance of recipient need and donor-recipient ties as predictors of aid allocation, I use the same dataset as that used for the analysis shown in Table 1 in the main text. To compare the predictive performance of ties and need as created via the different approaches (SSC, PCA, and MFA), I first randomized observations in the dataset into training and test sets. Because the dataset is a panel, and thus dyads over time are dependent, in constructing the training and test sets I randomized on the basis of dyads rather than individual units of observation. The training set was created by randomly selecting 70 percent of the dyads in the full panel, and the test set consisted of the remaining 30 percent.² Prior to dividing the data, I demeaned the outcome and the measures of interest at the donor-year level (so that variation reflects within donor and year differences in the variables). This approach allows me to avoid using donor and year factors in model training since some donors or years may not appear in both training and test sets.³

I trained four different predictive models using the training dataset: (1) a linear model estimated via OLS, (2) a generalized additive model (GAM), (3) a random forest regression, and (4) a neural network. The idea is to ensure that the performance of one set of measures is not unduly driven by parametric assumptions. The linear model was specified as

$$\text{Aid}_{irt} = \beta_0 + \beta_1 \text{ties}_{irt} + \beta_2 \text{need}_{rt} + \epsilon_{irt}, \quad (\text{A5})$$

and the specification for the generalized model was

$$\text{Aid}_{irt} = \beta_0 + s_1(\text{ties}_{irt}) + s_2(\text{need}_{rt}) + \epsilon_{irt}, \quad (\text{A6})$$

where $s(\cdot)$ denotes nonparametric smoothers. These fit local regression lines at certain levels of the predictor variables rather than imposing constant slopes as in a standard linear regression model.

²A seed of 111 was used.

³Predictions would be undefined for such cases.

Table B1: Prediction MSE

Model	Test Predictions			Train Predictions		
	SSC	PCA	MFA	SSC	PCA	MFA
Linear	1.537	1.56 (8.63)***	1.555 (4.1)***	1.516	1.536 (11.49)***	1.54 (8.02)***
GAM	1.517	1.539 (5.44)***	1.536 (3.45)***	1.481	1.522 (15.3)***	1.508 (7.35)***
Random Forests	1.564	1.646 (6.09)***	1.692 (8.11)***	0.395	0.383 (-6.07)***	0.308 (-35.95)***
Neural Net	1.525	1.549 (7.54)***	1.537 (2.42)*	1.502	1.523 (9.75)***	1.529 (7.82)***

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Sig. tests are from pair-wise t-tests (t-stat in parentheses) of the squared error in predictions using SSC derived measures relative to one of the alternatives. Entries not in parentheses are the prediction mean squared error.

The random forest and neural network models are denoted each as

$$\text{Aid}_{irt} = \text{RF}(\text{ties}_{irt}, \text{need}_{rt}), \quad (\text{A7})$$

and

$$\text{Aid}_{irt} = \text{NN}(\text{ties}_{irt}, \text{need}_{rt}). \quad (\text{A8})$$

The random forest model was trained using the default tuning parameters for the `ranger` R package, and the neural net was trained using one hidden layer with two neurons and a linear output.

Random forests provide a non-parametric approach to predicting outcomes given a set of factors by bagging the predictions of multiple regression trees trained on bootstrapped samples of the training dataset and random selections of covariates. Neural networks are a deep learning method that recursively fit layers of underlying linear and/or nonlinear equations relating a set of predictors to the outcome via a series of so-called “hidden” layers of neurons.⁴

Table B1 reports the mean squared error (MSE) for both out-of-sample predictions (under the “Test Predictions” header) and within-sample-predictions (under the “Train Predictions” header) using each of the above modeling strategies and the different approaches to measurement. The out-of-sample predictions were performed on the 30 percent of dyads randomly excluded from the training dataset. The within-sample predictions were performed on the 70 percent of dyads randomly included in the training dataset. Across all cases, DI and RN as derived via SSC out-performs the PCA and MFA versions, and these differences are statistically significant. For within-sample predictions, on the other hand, MFA and PCA do significantly worse in all cases except random forests.

Appendix C

Table C1 reports summary statistics for the main variables included in the dyadic foreign aid analysis. Table C2 reports the summary statistics for the variables included in the recipient-level analysis of lead donorship.

⁴More about implementing neural networks in R can be found at: <https://rpubs.com/julianhatwell/annr>.

Table C1: Summary Statistics for Dyadic Aid Analysis

	Mean	Median	SD	Min.	Max.	N
Alliance	0.119	0.000	0.323	0.000	1.000	68,856
Civil War	0.177	0.000	0.381	0.000	1.000	68,856
Colony	0.030	0.000	0.170	0.000	1.000	68,856
Democracy	6.017	6.000	3.446	0.000	12.000	68,856
ODA (asinh)	1.083	0.022	1.704	0.000	10.468	68,856
Distaster (asinh)	2.042	0.881	2.451	0.000	13.007	68,856
Distance (asinh)	9.466	9.593	0.606	5.458	10.578	68,856
Trade (asinh)	15.657	16.356	4.398	0.000	26.297	68,856
Income (log)	8.526	8.647	0.984	5.241	11.617	68,856
Population (log)	1.978	2.163	1.919	-2.664	7.237	68,856
Recipient Need	0.000	0.027	1.000	-2.081	3.315	68,856
Donor-Recipient Ties	0.000	0.036	1.000	-3.570	3.459	68,856

Table C2: Summary Statistics for Lead Donorship Analysis

	Mean	Median	SD	Min.	Max.	N
Lead Donorship	0.262	0.000	0.340	0.000	1.000	2,398
Lead Donorship (Steinwand 2015 version)	0.046	0.000	0.208	0.000	1.000	1,890
Recipient Need	-0.001	0.025	1.000	-2.081	3.315	2,398
Donor-Recipient Ties Concentration	-3.206	-3.224	0.086	-3.324	-2.752	2,398

References

- Chiba, Daina, and Tobias Heinrich. 2019. "Colonial Legacy and Foreign Aid: Decomposing the Colonial Bias." *International Interactions* 45(3): 474–99.
- Steinwand, Martin C. 2015. "Compete or Coordinate? Aid Fragmentation and Lead Donorship." *International Organization* 69(2): 443–72.