Honig When Reporting Undermines Performance Online Appendix: Data Collection Methods and Additional Empirics

Dan Honig

March 10, 2018

This document is the online appendix referred to in Honig, D. (2019). When Reporting Undermines Performance: The Costs of Politically Constrained Organizational Autonomy in Foreign Aid Implementation. *International Organization*, Forthcoming Winter 2019.

Data Collection

There is no existing cross-IDO database of project outcome data. The project success data therefore had to be collected from each IDO in the sample individually. I pursued project success data from every OECD bilateral aid agency in the top 10 in terms of the volume of official development assistance aid delivered directly in 2010, the last available data when this research commenced. This includes agencies in the US, Germany, the UK, France, Japan, Canada, Norway, Australia, Sweden, and Denmark. I also pursued data from all of the biggest multilateral aid agencies (the European Commission, UN Development Programme, World Bank, African and Asian Development Banks, and Global Fund), as were other agencies with which I had links (e.g. Irish Aid, International Fund for Agricultural Development, Food and Agriculture Organization, and International Monetary Fund).

There were two basic reasons to exclude IDOs from the sample. First, many IDOs do not in fact assign an overall, holistic success rating to projects ex-post. Second, for some IDOs I could not get access to outcome data that does exist (e.g. the African Development Bank). The IDOs included in this analysis are a convenience sample, raising concerns regarding broader generalizability. To the extent that the willingness to make data public, or the agency's decision to give projects an overall success rating, are plausibly correlated to an agency's autonomy this is a threat to generalizability that must be considered in examining these quantitative results in isolation (that is, without incorporating the case study findings). Table I.12 suggests there is cause for concern, as none of the bottom ten IDOs in autonomy are included in this analysis. It seems plausible that the least autonomous agencies, those with the least stable relationships with their political authorizing environments, are less likely to collect and/or make public information that might cast some of their projects in a less than stellar light.

The most straightforward result of this under-sampling would be to reduce the power of the quantitative tests; it is harder to imagine how this might lead to spurious findings. Spurious findings would result if the "true" shape of the relationship between autonomy and project success were parabolic. This seems most likely in the sense that the most autonomous agencies might engage in "too much" autonomy; however there is a good sampling of the "top" of the IDO distribution as regards autonomy. While seeming unlikely theoretically, if those with modest degrees of autonomy e.g. fared even worse than those with the lowest degrees of autonomy as environmental predictability role, this parabolic relationship (with the extremes of autonomy both faring better than the middle) would be missed due to the lack of data availability for the least autonomous IDOs.

The nine IDOs included are the the World Bank (WB), Asian Development Bank (AsDB), the UK's Department for International Development (DFID), the European Commission (EC), the Global Fund for AIDS, TB, and Malaria (GFATM), the German Agency for Technical Cooperation (GiZ), the German Development Bank (KfW), the International Fund for Agricultural Development (IFAD), and the Japanese International Cooperation Agency (JICA). Of the nine IDOs included only the World Bank's information is publicly accessible. The Asian Development Bank and DFID released data following formal public information requests. The European Commission and KfW released data under confidentiality agreements that limited their disclosure and further use. KfW later waived its confidentiality provision, allowing its data to be included in the publicly posted data that accompanies this paper; the EC declined to do so.

GiZ, IFAD, and JICA all maintain publicly accessible archives of individual project evaluation documents. In converting these individual project documents into a usable database I contracted research assistants using the online job contracting platform Odesk. RAs speaking the appropriate language (English, German, or Japanese as appropriate) extracted the relevant data – project names, performance scores, start and completion dates, budgets, etc. – from source documents, with me selectively double-checking their work (in the case of foreign language documents, with the help of Google translate). After compilation of each IDO's data I sent to each excel spreadsheets containing each agency's data were sent back to the originating agency for comment and/or correction. GiZ was kind enough to respond with a handful of minor corrections, which were incorporated. JICA had no substantive comment on the data itself, but wished it to be made clear that these data were generated by me rather than by JICA, which bears no responsibility for

errors or omissions. I hereby note that is the case, with all JICA data unofficial and unverified. IFAD never responded to multiple queries.

Archival Work on Project Success

As mentioned in the main text, I engaged in archival work to examine the documents underlying project documents. The World Bank uniquely allows access to archived primary project documents.[1] These documents include correspondence between project staff and between World Bank staff and national governments, back-to-office reports and (often handwritten) notes by those monitoring projects, detailed financial and performance indicators, and detailed evaluation reports. For approximately a dozen projects I reviewed archival documents at length, focusing on cases in which similar projects (such as the first and second phases of a particular project in a particular country) received quite different ratings and one might therefore be particularly doubtful about the reliability of ratings. In reviewing the archival documents (which in every case occurred many months after identifying the projects to be reviewed), I intentionally proceeded without knowledge of which projects were more or less successful and attempted to generate my own rating from the primary documentation. I cannot say that my rating on a six-point scale always matched the World Bank's score precisely. Indeed, this would be troubling if true, since evaluators also engage in conversations with project personnel, recipient government officials, and project beneficiaries, transcripts of which are not included in the archived files. However, there were no cases in which my self-generated rating differed by more than one point from the World Bank's official rating on a six point scale. In short, in this small sample success and failure do seem to be different and do map onto real features of the projects.

Construction of the Paris Declaration Monitoring Survey-Derived Measure of IDO Autonomy

As mentioned in the main text, the primary IDO autonomy measure employed in this work is a composite of two scales, one focusing on authorizing environment insecurity and the other on IDO propensity to devolve control over project implementation. The authorizing environment insecurity measure is constructed from two indicators. These indicators are, first, the degree to which aid is untied; that is, the extent to which it is not required that funds be spent on goods and services produced by the donor country. A high level of tied aid is a sign of an IDO's need to build political consensus for aid by serving domestic political constituencies and thus reflects more insecure footing in the IDO's political authorizing environment. The second indicator is the predictability of aid. The Paris

[1] Access to these documents, which require an extended vetting and declassification process, is via the World Bank Group Archives. These documents can be accessed by making requests under the World Bank's Access to Information policy.

Declaration asked donors for the first time to report formal projections of disbursements for future years; the monitoring surveys compare the last (that is, most recent) ex-ante projection of aid spending in a given year to the actual volume of aid disbursed.[2] Previous scholarship suggests that deviations from estimated sums are linked to IDO funding insecurity and political interference in IDO funding levels and direction.[3]

The propensity to devolve control measure is constructed from three indicators examining an IDO's project implementation behaviors. There is no available measure of IDO behavior with regards to their own agents; there are, however, systematic measures of IDO behavior as regards recipient country governments, and the frequency with which IDOs let go of principal control in favor of implementation led by these governments. Many of the same factors that I theorize drive IDOs' inappropriate retention of principal control vis-à-vis their agents – e.g. a worry about reputational risk and a desire to ensure short-term delivery is successful at the expense of long-term development goals – should also reduce an IDOs' propensity to hand over substantive control to developing country governments. I use IDO control tendencies toward recipient governments as a proxy for IDO control tendencies towards their own agents. The specific measures employed are the use of recipient-country public financial management (PFM) systems; the use of recipient-country procurement systems; and the avoidance of parallel implementation units.[4]

Additional Robustness

Additional Summary Statistics: Project Success

Tables II.1 and II.2 provide additional information regarding the key dependent variable, overall project success. Project success is,"inflated" to a six point scale from whatever the likert-type base scale is for each donor. This has no implication for the econometrics so long as IDO fixed effects are employed, but makes interpretation of the results more intuitive.

---

[2] This is a slight simplification; the indicator also penalizes over-disbursement, in fact calculating something like the absolute value of the deviation from prediction. In addition, disbursements are as-reported by partner government, adding inaccuracy borne of partner government data systems. See Ibid., 73–74 for more detail.

[3] Celasun and Walliser 2008; Desai and Kharas 2010.

[4] Parallel implementation units are separate operating units established at donors' insistence. These units use donor standards and thus give donors more control than would the routing of funds fully through recipient country government systems.

IDO	count	mean	sd	min	max
AsianDB	999	4.007508	1.036263	1.5	6
DFID	1795	4.610808	.9684172	1.2	6
EC	586	4.067406	.9810926	1.5	6
GFATM	538	4.750929	1.229771	1.5	6
GiZ	108	4.407407	.9175041	2	6
IFAD	31	4.16129	.7347006	2	5
JICA	672	4.984375	1.188046	1.5	6
KfW	1052	4.223384	1.02328	1	6
WB	3531	4.09544	1.18068	1	6
Total	9312	4.303898	1.138767	1	6

Table I.1: Summary Statistics of Project Success by Donor (6-point scale)

Perhaps an even more intuitive way of thinking about these data is to think of them as z-scores, given that – once IDO fixed effects are taken – each project is essentially being compared to the distribution of a given IDO's other projects. In the robustness checks in the paper and below I also drop the IDO fixed effect from regression models, instead using the z-score as the dependent variable.

IDO	count	mean	sd	min	max
AsianDB	999	.1100482	.9887898	-2.282585	2.011261
DFID	1795	-.0085493	.9993619	-3.528346	1.425033
EC	586	-.0084486	1.004518	-2.637156	1.970289
GFATM	538	.0025107	1.002199	-2.646827	1.020438
GiZ	108	-.0707946	1.050646	-2.827548	1.752903
IFAD	31	-.0282393	1.010707	-3.001466	1.12555
JICA	672	-.0083597	.9893771	-2.910065	.8374288
KfW	1052	.052686	.9197903	-2.8447	1.649623
WB	3531	-.0162114	.9279303	-2.449007	1.480637
Total	9312	.008058	.9637126	-3.528346	2.011261

Table I.2: Summary Statistics of Project Success by Donor (Z-scores)

Direct Effect of IDO Autonomy on Outcomes in Primary Analysis

The models in Table 2 do not incorporate a base term for IDO autonomy as it is collinear to IDO fixed effects. Table I.3 replicates Table 2 incorporating IDO autonomy and dropping IDO fixed effects. Results are substantively identical.

	(1) Project Success (Z-score)	(2) Project Success (Z-score)	(3) Project Success (Z-score)	(4) Project Success (Z-score)
IDO Autonomy	-1.559**	-2.033***	-1.725**	-2.016***
	(0.651)	(0.589)	(0.752)	(0.714)
Environmental Unpredictability	-0.127***	-0.146***	-0.123***	-0.132***
	(0.0297)	(0.0299)	(0.0359)	(0.0380)
Env Unpred*IDO Auton	0.148***	0.174***	0.151***	0.161***
	(0.0462)	(0.0446)	(0.0545)	(0.0562)
Project Size (USD Millions)		0.000562***		0.000353**
		(0.000182)		(0.000141)
Recipient Fixed Effects	N	N	Y	Y
R^2	0.025	0.024	0.077	0.080
Observations	9312	7247	9312	7247

Standard errors in parentheses, clustered by recipient country

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table I.3: Adding Base Term for IDO Autonomy to Table 2. Running regressions without IDO fixed effects but with the "base" autonomy scale leaves the key results on the interaction term substantively unchanged.

Validity of IDO Autonomy Measure

- Principal Components Analysis

The main text explained the construction of the primary IDO autonomy measure, and my decision to use a simple average of the five component measures drawn from the Paris Declaration Monitoring Surveys rather than a principal components approach. A principal components analysis of these five measures – aid predictability, untied aid, use of parallel implementation units, use of country public financial management systems, and use of country procurement systems - yields a first principal component with an eigenvalue of 3.09, thus explaining 62% of the variance in the five measures. This first principal component has quite even loading across the five constituent measures. The second component has eigenvalue of 1.08, just barely above the traditional cutoff of 1. Figure I.1 presents the scree plot.

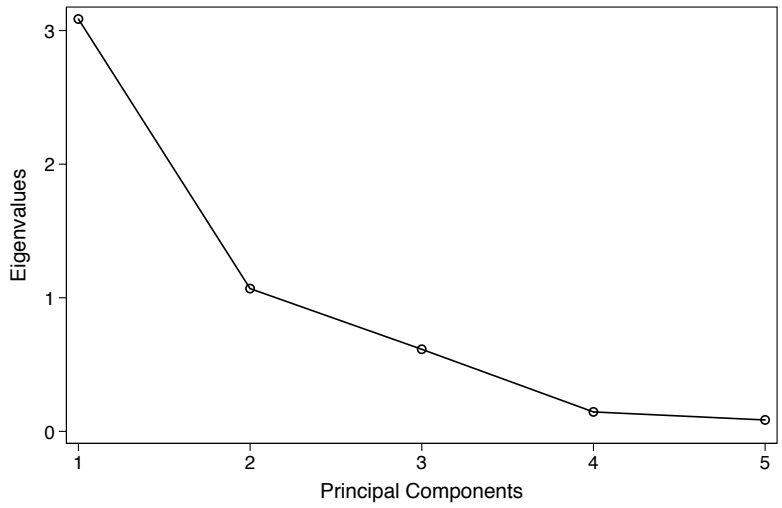

Figure I.1. Scree Plot of Principal Component Eigenvalues from IDO Autonomy Scale Measures

The second principal component, then, is quite marginal to begin with. Table I.4 examines the loading of the variables onto the first three principal components. The loading makes clear that the second component is picking up devolution propensity (with all three of the measures that form part of that subscale positive) where it does not overlap with authorizing environment insecurity (with both the measures that form that subscale negative). Thus combining the two principal components will lead to an overemphasis on devolution propensity relative to authorizing environment.

Variable	Component 1	Component 2	Component 3
Use of PIUs	.2796	.6668	.6767
Aid Predictability	.5254	-.1996	.1177
Use of Country PFM	.5339	.1003	-.2800
Use of Country Procurement	.4980	.1905	-.5030
Untied Aid	.3358	-.6849	.4436

Table I.4: Loading of IDO Autonomy measures onto Principal Components

Using only the first principal component struck me as quite similar, but much less intuitive, then simply averaging the five measures. A cluster analysis (via Stata's clv command) suggests what is implied by both the principal components analysis and intuition, that a single cluster with all five measures – that is, a single scale – is most appropriate here. As such I construct a simple average; the Cronbach's alpha of this scale (.825) suggests to me that this simple averaging is reasonable.

Nonetheless, I do retain the first principal component in the data to allow a robustness check; Table I.5 displays the results, which show the same effect as does the measure of autonomy employed in the primary results, e.g. Table 2 (in fact, t-statistics are higher using this principal components approach than with the primary measure).

DV: Project Success	(1) 6 pt scale	(2) Z-score	(3) 6 pt scale	(4) Z-score
Environmental Unpredictability	-0.0587***	-0.0456***	-0.0466***	-0.0393***
	(0.00716)	(0.00626)	(0.0134)	(0.0111)
Env Unpred*IDO Autonomy (Principal Component)	0.0199***	0.0142***	0.0171***	0.0125**
	(0.00517)	(0.00514)	(0.00650)	(0.00612)
IDO Autonomy (Principal Component)		-0.146**		-0.129
		(0.0717)		(0.0833)
Constant	4.341***	0.530***	3.758***	0.510**
	(0.126)	(0.0771)	(0.207)	(0.209)
IDO Fixed Effects	Y	N	Y	N
Recipient Fixed Effects	N	N	Y	Y
R^2	0.098	0.024	0.146	0.076
Observations	9312	9312	9312	9312

Standard errors in parentheses, clustered by recipient country

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table I.5: Results Using First Principal Component Instead of Simple Average for Paris Declaration Monitoring Survey-derived Scale

- Using my Alternate Field Survey Measure

One might be concerned that the Paris Declaration monitoring survey-derived IDO autonomy measure is not actually mapping autonomy. As noted in the main text, I conducted a small survey of aid experts in the field regarding IDO autonomy. A typical role for one of the survey respondents would be a senior position in the aid management unit of a recipient government's ministry of finance. Respondents rated a number of development agencies (including but not limited to those in the sample) on a scale of 1 to 7 in response to the following question:

> To what degree do you believe the in-country field office/bureau of the agencies listed below (presented in random order) are enabled to make decisions with a significant impact on the direction, nature, or quality of development projects? Please only respond for those agencies you have had exposure to either via working with the agencies or discussions with colleagues.

The survey N is 28, with varying coverage for different donors. This is a small but well-informed sample; methodological studies suggest small numbers of high-quality respondents will prove more accurate than significantly larger samples that

lack expertise.[5] The N of 28 is the remaining N after removing surveys which were not substantively responsive or gave indications of nonsensical answers; the two largest reasons for exclusion were (a) rating the Asian Development Bank despite stating that all relevant development-related work experience was in an African country (where the Asian Development Bank does not function) or (b) rating the survey's anchoring vignettes such that the most autonomous text was evaluated as being just as autonomous or less autonomous than the least autonomous text. The survey N is limited by the small number of individuals in any given country who can make expert inter-donor comparisons (this generally excludes employees of development agencies, who can only speak intelligently regarding their own organization).

The correlation between this survey measure and the autonomy scale drawn from the Paris Declaration surveys is .73. Table I.6 substitutes the survey measure for that of the Paris Declaration-derived measure, otherwise paralleling the analysis of Table 2; the results are similar, which should increase confidence in the primary analysis.

	(1)	(2)	(3)	(4)
DV:	6pt scale	Z-score	6pt scale	Z-score
Environmental Unpredictability	-0.102***	-0.0852***	-0.0760**	-0.0724***
	(0.0255)	(0.0228)	(0.0326)	(0.0269)
Env Unpred*IDO Autonomy (Survey)	0.0170***	0.0146**	0.0123*	0.0119*
	(0.00641)	(0.00582)	(0.00713)	(0.00622)
Autonomy (Survey)		-0.142*		-0.132
		(0.0762)		(0.0817)
addlinespace Constant	4.941***	0.892***	4.489***	0.974***
	(0.102)	(0.287)	(0.189)	(0.354)
IDO Fixed Effects	Y	N	Y	N
Recipient Fixed Effects	N	N	Y	Y
R^2	0.094	0.021	0.142	0.072
Observations	8313	8313	8313	8313

Standard errors in parentheses, clustered by recipient country

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table I.6: Robustness to Use of Survey Measure.

Outcome Variance

One might be worried that results are driven by quirks in the variance of outcomes. Table I.7 examines this concern in a simple nonparametric manner, by dividing environmental predictability and autonomy scores at their respective

[5] Leuffen, Shikano, and Walter 2012.

means and then examining the variance in project success z-score by autonomy and environmental predictability quadrant, and finds no substantively large differences. By calculation (see Table I.2), the Z-score outcome measure has a mean near 0 and standard deviation 1 for each IDO. Table I.7 allows us to examine if the variance in this measure differs systematically along the autonomy and environmental predictability axes, thus potentially distorting the interpretation of OLS results. The question, then, is whether any of the quadrants deviate substantially enough from 1 to cause concern. Both low autonomy and high autonomy IDOs do better in contexts of lower environmental unpredictability; the gap between low- and high-SFI contexts is larger for low-propensity to IDOs (approximately .39 SD) than for high-propensity to IDOs (.17 SD).[6]

	Low Autonomy IDOs	High Autonomy IDOs
Low environmental unpredictability	.163 (.863)	.123 (.969)
High environmental unpredictability	-.226 (.998)	-.047 (.983)

Table I.7: Analysis by IDO Autonomy and Environmental Unpredictability Quadrant.

Features of the modeling (e.g. Overfitting concerns, 2nd-level N distortions, etc.)

One might also worry, particularly given the small number of IDOs in this multilevel model, results are driven by features of the modeling. To address this concern, I first calculated the simple mean of project success (expressed as a z-score) for each IDO for projects above and below the mean of environmental unpredictability, yielding eighteen observations (two per IDO). I then calculated the gap between each IDO's high unpredictability and low unpredictability project mean success (thus leaving one observation per IDO). I then used this gap as the dependent variable in a regression with only a single explanatory variable, IDO autonomy. This result is Table I.8 below. There remains a clear relationship between IDO autonomy and the impact of environmental unpredictability on performance, significant at the 90% confidence interval (and just short of

[6] Given the large N, the analysis can of course confirm that that these variances are not equal; the question is whether they are substantively different enough to potentially bias results. I would argue the answer to this is in the negative.

significance at the 95% level, with a t-statistic of 2.32 but only nine observations). The R-squared is also .43, suggesting IDO autonomy explains a great deal of the difference in this gap in comparing IDOs. [7]

	(1) Gap Between High and Low Unpredictability Success by IDO
Autonomy (PD Scale)	1.319^{+}
	(0.569)
Constant	-1.177*
	(0.375)
R^2	0.434
Observations	9

Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < .01$

Table I.8: Difference in Difference (in Z-scores) of Gap Between High and Low Unpredictability Project Success (split at mean of Environmental Unpredictability)

Table I.9 below further examines the underlying relationship between IDO autonomy and project success at the IDO-by-IDO level, summarizing the relationship between environmental unpredictability and overall project success for each donor in isolation; that is, using only data from one donor at a time and implementing nine different regressions.[8] In each case, the model is of the form

$$\text{Project Success}_{i,j} = \beta_1 * \text{Environmental Unpredictability (State Fragility Index)}_j + \varepsilon_i.$$

[7] The coefficient is positive even though the gap between high unpredictability and low unpredictability projects is smaller as IDO autonomy rises because the dependent variable (the gap itself) is always negative; every IDO has less success in high unpredictability environments than low unpredictability environments. For more autonomous IDOs this is a smaller negative number than for less autonomous IDOs.

[8] This is intuitively similar to a rank-based regression.

IDO	Autonomy Score	Correlation between Env Unpred & Success (Z-score) for this donor with only this donor's data in regression
EC	.559	-0.0249**
		(0.0103)
Global Fund	.594	-0.0471***
		(0.0112)
World Bank	.608	-0.0365***
		(0.0043)
GiZ	.666	-0.0525***
		(0.0175)
KfW	.666	-0.0331***
		(0.0101)
JICA	.667	-0.0221
		(0.0133)
Asian DB	.669	-0.0671***
		(0.0217)
IFAD	.681	-0.0183
		(0.0362)
DFID	.799	-0.0019
		(0.0084)

Standard errors in parentheses, clustered by recipient country
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table I.9: IDO-by-IDO Regressions. This table allows a direct examination of the 2nd level N that drives results. IDOs with lower levels of autonomy see a greater negative correlation between environmental unpredictability and project success.

As expected, greater environmental predictability has a more negative and statistically significant relationship with overall project success for less autonomous donors. This confirms—using an approach that does not rely on the parameterization of the interaction term—that higher levels of autonomy mitigate the inverse relationship between the environmental unpredictability measure (the State Fragility Index) and overall project success. A figure presenting the data underlying table I.9 is presented as Figure I.2 below. [9]

[9] Credit to Chris Kilby, who as a discussant at NEUDC 2014 first generated this graph (that is, the graph is generated by me, but is inspired by a similar graph generated by Kilby).

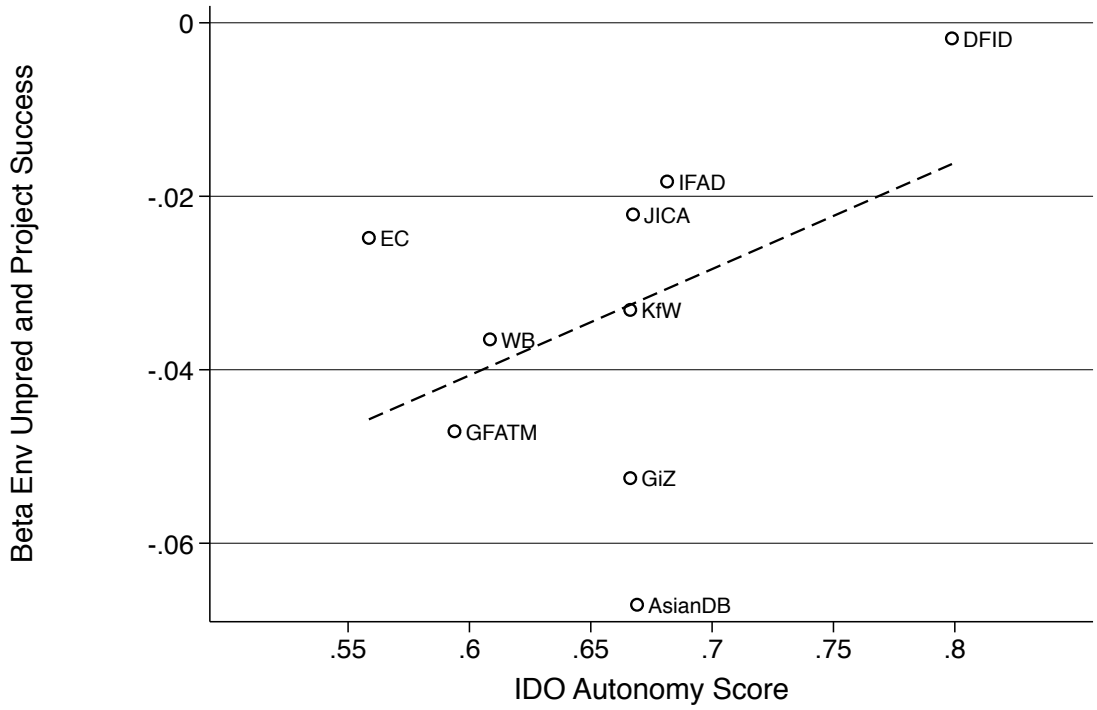

Figure I.2: Graph of IDO-by-IDO slopes.

While Figure 2 in the main text chooses the highest and lowest observed values of IDO Autonomy in demonstrating effects, Figure I.3 shows that the 25th and 75th percentile observations of IDO Autonomy are still differentiable from one another, drawing from the same model as figure 2 in the main text.

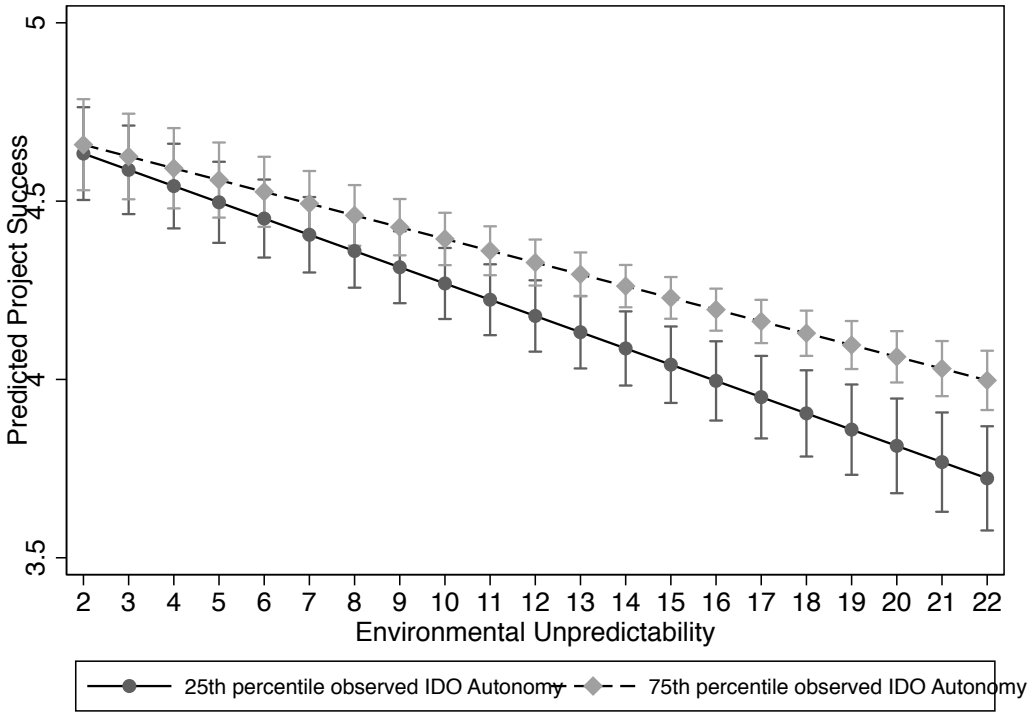

Figure I.3: Differentiating the 25th and 75th percentile of IDO Autonomy from one another

Clustering of Standard Errors

The primary analysis in this paper has clustered standard errors at the recipient country level to adjust for the possibility that project success may be correlated within a given recipient country. The results presented in the primary analysis are robust to alternative clustering strategies; that is, to clustering standard errors at the IDO level or, where practicable, to double clustering at both the IDO and the recipient country level.

As noted in the main text, it is also possible that project success is correlated within IDOs. While the small number of clusters when clustering at the IDO level may negatively affect the coverage properties of clustered standard errors (one of the motivations for clustering at the recipient level in the primary analysis), Table I.10 shows that results with standard errors clustered on the IDO are strongly consistent with the main text. The most conservative clustering strategy would be to double-cluster at both the IDO and recipient level. However, doing so precludes inclusion of fixed effects; that is, the limited degrees of freedom (given the 2nd-level N of nine) makes the inclusion of either donor or recipient fixed effects and double-clustering

simultaneously impossible.[10] Table I.11 implements double-clustering in the only case where it is viable to do so, paralleling model 2 in Table I.11. Using the Z-score of project success as the dependent variable (and thus no IDO fixed effect), Table I.11 suggests that the primary results are also robust to simultaneous clustering at the recipient and IDO level, to the extent that estimating such a model is possible.

	(1)	(2)	(3)	(4)
DV:	6pt scale	Z-score	6pt scale	Z-score
Environmental Unpredictability	-0.170***	-0.127***	-0.149***	-0.123***
	(0.0272)	(0.0235)	(0.0310)	(0.0192)
Env Unpred*IDO Autonomy	0.205***	0.148***	0.187***	0.151***
	(0.0398)	(0.0372)	(0.0440)	(0.0277)
IDO Autonomy		-1.559**		-1.725***
		(0.671)		(0.354)
Constant	4.423***	1.383**	3.807***	1.564***
	(0.0367)	(0.418)	(0.381)	(0.283)
IDO Fixed Effects	Y	N	Y	N
Recipient Fixed Effects	N	N	Y	Y
R^2	0.099	0.025	0.147	0.077
Observations	9312	9312	9312	9312

Standard errors in parentheses, clustered by IDO

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table I.10: Main Results with Standard Errors Clustered by IDO

	Project Success (Z-score)
Environmental Unpredictability	-0.127***
	(0.0245)
IDO Autonomy	-1.559**
	(0.644)
Env Unpred*IDO Autonomy	0.148***
	(0.0388)
Constant	1.383***
	(0.386)
R^2	0.025
Observations	9312

Standard errors in parentheses, double clustered by IDO and recipient country via cgmreg

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table I.11: Main Results with Double Clustering, outcome as z-score, base term for autonomy, and no fixed effects for recipient country, IDO, or sector

Additional Robustness Tests

[10] Estimation is via Cameron, Gelbach, and Miller 2006 and their *cgmreg* routine. I mean to say that *cgmreg* cannot estimate standard errors – the routine fails – when fixed effects are included.

In addition to the robustness checks discussed here, the results above are robust to:

- Using ordered logit models on six point project outcome scales (rather than OLS)
- Using z-scores as outcomes (rather than the six-point scale where employed)
- Compressing success and failure to a binary outcome and employing logit models
- Restricting SFI to common support; that is, only the range of SFI realized in all donors' data (2-22, rather than 0-25 in the main analysis)
- Dropping the latter two waves of the Paris Declaration survey in generating the IDO autonomy measure (to allay concerns that donors responded to measurement by changing their practices)
- Dropping either subscale of the state fragility index (legitimacy or effectiveness)
- Using any of the four domains of the state fragility index (security, political, economic, or social)

IDO Autonomy Measure in Full

Table I.12 provides a full list of all organizations for which IDO autonomy scores were generated and those scores, for full transparency regarding the measure.

IDO	Autonomy Score	Rank
Ireland	0.878579795	1
UK (DFID)	0.798823953	2
Norway	0.796352506	3
Netherlands	0.773272038	4
Sweden	0.719851851	5
IMF	0.714166641	6
Finland	0.689640522	7
IFAD	0.681465507	8
Denmark	0.678942561	9
Canada	0.677956104	10
AsianDB	0.669080436	11
Japan (JICA)	0.667425275	12
Germany (GiZ/KfW)	0.666281819	13
France	0.628306508	14
WB	0.608462632	15
Switzerland	0.605228841	16
GFATM	0.593850553	17

New Zealand	0.593333304	18
EC	0.558577597	19
Austria	0.535915732	20
Spain	0.533711374	21
Belgium	0.501379311	22
Luxembourg	0.492137939	23
African DB	0.488045961	24
Australia	0.480275869	25
Portugal	0.476367801	26
Italy	0.459770113	27
IADB	0.392873555	28
Korea	0.375316083	29
United States	0.36240229	30
GAVI Alliance	0.330833346	31
Turkey	0.285268188	32
United Nations	0.234992817	33

Table I.12: Full List of IDO Autonomy Scores. Note that the Paris Declaration Monitoring Surveys are, for bilateral donors, at the country level; thus KfW and GiZ share Germany's score. For the other bilateral donors in the sample I have added the IDO name to the country where appropriate.

Case Study Interviewees

Interviewee Data and Numbering Schemes

The following table indicates all individuals that provided information (mostly by interview, but in a handful of instances by email correspondence) that informs the broader qualitative data in *Navigation by Judgment* (Oxford University Press, 2018), not only those interviewees upon whose comments this article draws. A few notes on these tables – first, the number in the left-hand column does not correspond to the number in the interview citations (e.g. interview 63, 6/25/13). The cited numbers are randomized to maintain the promised confidentiality to interviewees. The dates of the interviews are omitted from the tables below, as including dates would make it much easier to infer the identity of a given speaker. For the same reason, individuals who contributed in more than one domain – e.g. speaking to South African interventions in both municipal governance and health – are given a new randomly generated number for use in each section, as to do otherwise would make it exceedingly easy to identify these speakers. As such, though 147 interviews are listed below, there is a degree of overlap; it is small, however. There are more than 135 unique interviews on which these cases collectively draw.

SOUTH AFRICA

Table I.13: South Africa Municipal Governance Capacity Building Interviews					
#	Surname	First Name	Position	USAID	DFID
1	Anonymous	Anonymous	COGTA Senior Official		
2	Bester	Angela	Former DFID staff, then Deloitte; also former DG, Public Service Commission		
3	Brooks	Frikkie	Head of KZN provincial planning department		
4	Chipkin	Ivor	Executive Director, PARI		
5	Chrystal	Blake	Supervisory Program Officer, USAID SA		
6	Dei	Colleen	Former USAID SA Mission Chief		
7	Fortuin	Joe	Director of Aid, COGTA		
8	Francis	Virginia	USAID Health Team, former RTI SA staff		
9	Glasser	Matt	Former USAID advisor in SA on municipal financing		
10	Hackner	Allan	USAID SA Financial Sector Manager (former COR on Municipal)		
11	Harding	Joel	DFID Governance Advisor		
12	Heymans	Chris	Former CMTP chief architect, now WB		
13	Hofmeyr	Beatie	Head of Education and Training Unit, LGSP implementing sub-contractor		
14	Horn	Steve	former ISLGS CoP		
15	Kolker	Joel	Former USAID staff, municipal program		
16	Konig	Ferdie	CMTP ISF in Phalaborwa, Mpumalanga		
17	Layte	Michelle	former RTI LGSP CoP (Vaz's successor)		
18	Madurai	David	Chief Director, Norms, Standards, Policy and Research, COGTA; former Chief Director, Delopment Planning & Local Economic Development		
19	Mangokwena	Andries	Advisor in Thulamela under CMTP		
20	Mathivha	Makonde	Municipal Manager, Thulamela, Limpopo		
21	Matomela	Bongani	Former Deputy Project Director, LGSP		
22	Naidoo	Subethri	Former Governance Advisor, DFID; former Local Government sector manager, USAID; former Deloitte program manager on CMTP		
23	Olver	Chippy	Former Deputy Director General DPLG		
24	Powell	Derek	Former Deputy DG, DPLG		
25	Rambulana	Wilson	former LGSP Revenue Enhancement Advisor (trainer)		

26	Sadan	Mastoera	Programme Manager, PSPPD, Office of the SA Presidency	[redacted]		
27	Savage	David	Former WB staff, now SA Treasury head of Cities Support program	[redacted]		
28	Snook	Steve	former USAID Democracy and Governance deputy team leader	[redacted]		
29	Tazewell	Littleton	Deputy Mission Director, USAID South Africa Regional Program	[redacted]		
30	Thomas	Richard	Former DFID South Africa Governance Advisor on CMTP		[redacted]	
31	Timm	Jeremy	Former CMTP now Treasury muni gov support		[redacted]	
32	Toli	Robin	Chief Director, International Development Coordination, SA Treasury	[redacted]		
33	TV	Pillay	Head of Municipal Finance, SA Treasury	[redacted]		
34	Vaz	Peter	former RTI LGSP CoP	[redacted]		
35	Yako	Pam	Former municipal manager, Amathole District; former DG, Environmental Affairs, Water Affairs	[redacted]		

Table I.14: South Africa Health Interviews						
#	Surname	First Name	Position	USAID	DFID	CDC
1	Agenbag	Rentia	Government and Civil Society Support Manager, SANAC	[redacted]		
2	Anonymous	Anonymous	CDC & USAID PEPFAR Implementer	[redacted]		
3	Anonymous	Anonymous	Senior DC-based PEPFAR official	[redacted]		
4	Anonymous	Anonymous	Senior CDC Official in Another Southern African Country	[redacted]		
5	Anonymous	Anonymous	USAID and CDC PEPFAR Implementer	[redacted]		
6	Barker	Pierre	Senior VP, Institute for Health Care Improvement		[redacted]	
7	Barron	Peter	Public health specialist & advisor to DDG Pillay	[redacted]		
8	Coovadia	Jerry	Director, MaTCH	[redacted]		

9	Coovadia	Ashraf	Head of pediatric HIV,Rahima Moosa Mother and Child Hospital, Johannesburg			
10	Dei	Colleen	Former USAID SA Mission Chief			
11	Desmond	Chris	Chief research specialist, Human Sciences Research Council			
12	Fryatt	Bob	Former DFID Health Advisor, SA			
13	Giddy	Janet	Former HIV program coordinator, McCord Hospital, Durban			
14	Goga	Ameena	Senior Specialist Scientist, MRC			
15	Gorna	Robin	Former Senior Regional Health and AIDS adviser for DFID Southern Africa			
16	Grant	Ken	HLSP Programme Director, SARRAH			
17	Harding	Joel	DFID Governance Advisor			
18	Holst	Helga	CEO, McCord Hospital, Durban			
19	Kok	Michelle	Advisor to Precious Robinson, NDOH			
20	Kumar	Smita	USAID PMTCT Lead			
21	Lesole	Lerato	PMTCT Specialist, CDC SA; previous NDOH			
22	Mahasela	Lusanda	Deputy, Research & M&E, Johns Hopkins Health and Education in South Africa			
23	Mazibuko	Ntombi	RTC PMTCT Project Manager; former EGPAF & NDOH; former NDOH			
24	Ngubane	Gugu	former HLSP A-Plan Project Manager and Technical Adviser on PMTCT			
25	Nkulu	Hilary	former DFID SA Programme Manager			
26	Pattinson	Robert	Director, MRC Maternal and Infant Health Care Strategies Unit, University of Pretoria			
27	Pillay	Yogan	NDOH Deputy Director General			

28	Robinson	Precious	NDOH Deputy Director in charge of PMTCT			
29	Sanne	Ian	CEO, Right To Care			
30	Schneider	Helen	Director, School of Public Health, University of the Western Cape; former SANAC, MRC			
31	Slingers	Nevilene	Donor Coordination Manager, SANAC			
32	Taback	Rayna	Senior Public Health Advisor, CDC South Africa			
33	Tazewell	Littleton	Deputy Mission Director, USAID South Africa Regional Program			
34	Toledo	Carlos	Chief, HIV Prevention Branch, CDC South Africa			
35	Toli	Robin	Chief Director, International Development Coordination, SA Treasury			
36	Venter	Francois	Deputy Executive Director, Wits Reproductive Health Institute			
37	Vranken	Peter	CDC Senior Technical Advisor, PEPFAR			
38	Wilson	John	HLSP Programme Manager - MSP, RRHF, SARRAH			

LIBERIA

Table I.15: Liberia Health Sector Interviews					
#	Surname	First Name	Position	USAID	DFID
1	Anonymous	Anonymous	Former Liberia NACP Advisor		
2	Anonymous	Anonymous	MoHSW senior personnel		
3	Anonymous	Anonymous	Senior official, USAID Liberia		
4	Anonymous	Anonymous	USAID Liberia international staff		
5	Augustin	Randolph	Lead Health Officer, USAID		
6	Benson	Angela	FARA Coordinator, MoHSW		
7	Bility	Kalipha	Former Program Coordinator, NACP; in 2013 Deputy Minister, Ministry of Agriculture		
8	Bruce	Lwopu	Head of Blood Safety, MoHSW; former deputy head, NACP		

9	Curran	Desmond	DFID Representative in Liberia 2007-2009		
10	Dahn	Eunice	Chief Medical Officer, MoHSW		
11	Davis	Natty B.	Chairman and CEO, NiC; former Minister without Portfolio and National Coordinator, LRDC		
12	Dolopeh	Dr.Eugene	Former Program Manager, NACP		
13	Duncan	Julie	Commissioner, NAC; former Assistant Minister for Preventive Services, MoHSW		
14	Dworku	Tanu	Former USAID Health Officer, Former NACP Coordinator		
15	Dzokoto	Agnes	Senior Technical Officer, AWARE (responsible for Liberia)		
16	Flomo	Matthew	Deputy Minister for Administration, MoHSW		
17	Freeman	Josephine	Former PMTCT Coordinator, NACP		
18	Gabelle	Chris	Former lead Liberia Governance Advisor, DFID		
19	Gaddis	Beth	Health Officer, USAID		
20	Gwenigale	Walter	Minister of Health, MoHSW		
21	Hughes	Jacob	Head of Liberia Health PF Management firm; Former PwC		
22	Hymowitz	Dan	Advisor to the Monserrado County Ebola Response, African Governance Initiative		
23	Jones	Janyaj	M&E Deputy, NACP		
24	Karzon	Toagee	Controller, MoHSW		
25	Lippevald	Theo	RBHS/JSI Deputy CoP		
26	Logan	David	Global Fund Coordinator, MoHSW; former deputy coordinator, NACP		
27	Macaulay	Rose	RBHS/JSI CoP		
28	Manuel	Marcus	Former DFID Deputy Director for West Africa		
29	Mapleh	Louise	PBF Coordinator		
30	Martin	Bill	Former Senior Adviser to the Minister, MoHSW; now PF Manager		
31	McDermott	Chris	Former health lead, USAID		
32	Nartey	Alex	Former lead of PwC team to MoHSW		
33	Niyuhire	Floride	RBHS Advisor on PBF to MoHSW		
34	Nyoweh	Moses	STI Officer, NACP		

35	Sanvee	Dr.Lilly	Head implementer, Catholic Hospital, AWARE		
36	Scheening	Sarah	Senior Policy and Implementation Advisor, USAID Global Health Bureau		
37	Sieh	Sonpon	Program Coordinator (head), NACP; former M&E on HIV, NACP		
38	Sirleaf	Momolu	Head of Aid Coordination, MoHSW		
39	Subah	Pewu	Head of Project Implementation Unit, MoHSW		
40	Tamattey	Felix	Senior Partner leading PwC Engagement, MoHSW		

Table I.16: Liberia Capacity Building Interviews

#	Surname	First Name	Position	USAID	DFID
1	Anonymous	Anonymous	Senior CSA Official		
2	Anonymous	Anonymous	Senior DC-based USAID Official		
2	Allen	William	Former Director General, Liberia Civil Service		
3	Atuanya	Jenkins	Former Deputy Director General, CSA; now assistant minister, ministry of Lands Mines & Energy		
4	Baki	Shadi	Head of Biometrics, CSA		
5	Belleh	Willie	Partner Subah Belleh Associates; local partner for CISCAB		
6	Callender	Elizabeth	Deputy Head, OTI Liberia		
7	Cooper	Vicky	Former WB consultant on Civil Service Pay Reform; current Chief of Party, GEMS		
8	Cooper	Lloyd	Grants Manager, BRDG		
9	Curran	Desmond	DFID Representative in Liberia 2007-2009		
10	Davis	Natty B.	Chairman and CEO, NiC; former Minister without Portfolio and National Coordinator, LRDC		
11	Drosaye	Alfred	CSA PAO		
12	Fahnbulleh	Louise	former OTI staff, Liberia		
13	Fn'Piere	Pat	Consultant, BRDG; OTI Field Advisor		
14	Gabelle	Chris	Former DFID Governance Advisor in Liberia		

15	Gattorn	John	Former Africa Program Manager, OTI		
16	Glentworth	Garth	OBE; former senior Governance Advisor, DFID		
17	Hare	Sam	Former Deputy Minister, Ministry of Youth and Sports		
18	Hunter	Rosslyn	M&E team, BRDG		
19	Johnson	Mimi	HR team, BRDG		
20	Kialain	David	Former principal deputy, GRC		
21	Lauer	Barb	Former CoP for BRDG, DAI		
22	Liberty	T. Edward	Director General, LISGIS		
23	Logan	James	Former Deputy Minister, Ministry of Agriculture		
24	Mayshak	Nellie	Former head, ASI CISCAB team 2007-2009		
25	Muhula	Raymond	Public Sector Specialist, World Bank		
26	Neymah	Oblayon	Former Reform Directorate CSA; current head of LIPA		
27	O'Neill	Dominic	Head of DFID Sierra Leone 2008-2011		
28	Panton	Richard	Deputy Director General, Training, LIPA		
29	Patel	Jalpa	Former coordinator, ASI CISCAB project, 2009-2010		
30	Sigrist	Ken	Former head, ASI CISCAB team 2009-2010		
31	Tarpeh	Dominic	Former CISCAB consultant, now with GRC		
32	Thompson	James	Subah Belleh staff; former member of CISCAB core team		
33	Wilson	Peter	Program Development Officer, BRDG		
34	Wilson	Mark	Grants Manager, BRDG		