Letting the Driver Steer: Organizational Structure and Country Context in Delivering Better Aid

Dan Honig[1]

December 23, 2013
Political Economy of International Organizations (PEIO)

Abstract

This paper explores attempts by external aid agencies (e.g. the World Bank, USAID) to improve systems and deliver services in contexts of heterogeneous stability. The key finding is that giving greater autonomy to agents in the field improves the success of international development projects, with increasing returns to autonomy in fragile states and in project domains where it is more difficult to externally observe outcomes (e.g. democratic participation projects as opposed to road construction projects). The analysis uses regression analysis of a novel dataset composed of evaluations of over 14000 projects from nine international development organizations, using self-evaluated project outcomes as a measure of success, the state fragility index as a measure of state fragility, and both expert surveys and a measure constructed from organization-level responses to Paris Declaration monitoring surveys as measures of agent autonomy. In addition to shedding light regarding important facets of the development outcome production process, this work suggests that bureaucratic delivery channels have an important role to play in the variance of outcomes, suggesting these processes ought receive greater attention in the development literature.

[1]I thank the National Science Foundation Graduate Research Fellowship for their support under grant# DGE-1144152. Matt Andrews, Peter Hall, Art Kraay, Sheila Page, Steve Radelet, Tristan Reed, Ryan Sheely, Beth Simmons, Eric Werker, and many others have all provided incredibly helpful comments on earlier versions of these ideas and/or this work. I can be reached at dhonig@fas.harvard.edu.

Introduction

> *Although the uniqueness of the foreign aid agency's task has been recognized and understood, the organizational environment that such a task requires has never been specified…. I ascribe problem results to an organizational, rather than a historical, determinism (Tendler 1975, p. 9, 110).*

This paper argues that international development organizations (IDOs) are significantly hampered by their desire to make an environment legible – tractable to analysis and intervention.[2] In pursuing this desire they lose the benefit of what James Scott calls "mētis", namely "a wide array of practical skills and acquired intelligence in responding to a constantly changing natural and human environment." (Scott 1998, p. 313)

Mētis is a form of what economists call "soft information", or

> Information that cannot be directly verified by anyone other than the agent who produces it. For example, a loan officer who has worked with a small-company president may come to believe that the president is honest and hardworking – in other words, the classic candidate for an unsecured 'character loan.' Unfortunately, these attributes cannot be unambiguously documented in a report that the loan officer can pass on to his superiors. (Stein 2002, p.1892)

Mētis, then, is inherently intractable to the "hardening" that is necessary for its transmission across organizational distance, e.g. up a hierarchy.

This work aims to test if mētis is a beneficial component of the IDO development results production process. It does so by investigating whether the degree of autonomy given to field-level agents, who can then incorporate mētis, is correlated with project success. It uses a novel data set composed of evaluations of 14,000 unique projects in 182 counties over the past 50 years from nine donor agencies.

Theory & Research Design

IDOs & Autonomy

This work assumes that in most cases bureaucracies, and field-level bureaucrats,[3] matter. Outcomes are not merely dictated by political or management-level staff or executive boards. Scholars have argued that bureaucracies, and bureaucrats, have a realm for independent action (Carpenter 2001; Heimer 2013; Huber and Shipan 2002) in which individual bureaucrats influence outcomes via their exercise of discretion (Canales 2010). These works build on the power borne of information asymmetries perhaps first identified in Weber (1922) and developed most explicitly in contract theory (Aghion and Tirole 1997; Calvo and Wellisz 1978; Friebel and Raith 2004; Grossman and Hart 1983; Hart and Holmstrom 1987; Holmstrom 1999; Laffont and Tirole 1988; Tirole 1986). The literature

[2] Pritchett & Woolcock (2004) also briefly employ Scott's notion of legibility in describing IDOs and their push for isomorphism among recipients.
[3] Development's version of 'street level bureaucrats' (Lipsky 2010).

argues that the greater the informational asymmetry[4] and the more difficult monitoring, the more scope will be present for agents to exert effective control, to gain autonomy. Thus how tasks are organized has an impact on outcome in settings as diverse (and difficult to monitor) as isolated forest rangers (Kaufman 2006) and state prison employees (DiIulio 1987).

In public bureaucracies in general – and IDOs are no exception – the combination of many tasks (and in multilaterals and some bilaterals) many principals (Dewatripont, Jewitt, and Tirole 1999; Dixit 1997, 2002; Johns 2007; Tirole 1994; Wilson 1989) exacerbates the information problem.[5] As a consequence agents have more slack (Lyne, Nielson, and Tierney 2006), particularly where long chains of nested principal-agent relations are present (Kofman and Lawarree 1993; Strausz 1997) or where managerial rotation is frequent (Kiser and Kane 2007). These conditions arise frequently in IDOs. (see Azam & Laffont, 2003; Chauvet, Collier, & Fuster, 2006; Cooley, 2005; Murshed, 2009).

While some aid is associated with real and impressive improvements – e.g. life expectancy and schooling attendance/literacy outcomes have risen markedly in the past forty years (UNDP 2010), an increase that cannot be explained by income or economic growth; it is hard to imagine aid having played no role in these gains. That said, the literature largely focuses on the ineffectiveness of significant amounts (perhaps even most) aid, bilateral and multilateral, to achieve its intended goals. A number of scholars (Allegret and Dulbecco 2007; Barnett and Finnemore 2004; Birdsall 2008; Cooley and Ron 2002; Gibson et al. 2005; Ostrom, Gibson, and Shivakumar 2002; Whittle and Kuraishi 2008) see organizational dysfunction as central in explaining failings of foreign aid delivery, a thread upon which this work builds.

I theorize that the greater the need of any given level of the hierarchy to manage up to those who do not share the context – e.g. an executive board (or national legislators/appropriators) to whom the management of an IDO must account - the more constrained that level will be in enabling down to the field, in treating the task environment as about project success (down) rather than reputation management and legitimacy (up). This dynamic is identified by Tendler (1975) in her description of the large amount of criticism USAID received:

> It has been generally recognized that criticism of the foreign aid program weakened [USAID] and kept it from doing what it wanted to do. Less understood is the fact that the process of living with criticism profoundly affected what the agency *wanted* to do and **what it was capable of doing** (p. 40).[6]

[4] That is, the more the agent knows what the principal does not and cannot without the agent conveying the information to the principal.
[5] This exacerbates both of the classic problems of the principal-agent relationship, that of hidden action (moral hazard) and hidden information (adverse selection/information asymmetries). Multiple principals in bilaterals might include domestic NGOs, legislators, the ministry of foreign affairs/state department, etc.
[6] Boldfacing mine.

Interaction with Recipient Country Governments

Recipients are differentially capable, and differentially legible to IDOs. The development community has broadly accepted that country context matters as regards so-called "fragile states", and has accepted the notion that they may have different needs than other countries; fragile states are mentioned in all of the major recent international aid conventions (Accra 2008, Busan 2011, Paris 2005). The need for discretion and flexibility does not end at the threshold of fragility, however; the 2011 WDR (World Bank 2011) follows this logic, endorsing the importance of adapting the assistance modality to risk along a broader spectrum of recipient government quality and suggesting internal donor reform is needed to achieve these goals (pgs 276-8).

In those developing countries characterized by higher levels of stability the name on the door of a government unit is well-correlated with the activities that take place within, and medium- and long-term plans made have some reasonable chance of proceeding apace, with predictable risks to implementation. In others, none of this is the case.

In the organizational behavior literature it is widely accepted that organizational forms perform differentially by task environment (Brechin 1997; Lawrence and Lorsch 1967; Thompson 1967), and that the more unpredictable the work process or environmental volatility, the higher the optimal level of agent discretion/autonomy (Dobbin and Boychuk 1999; March and Simon 1958; Thompson 1967). In high-fragility countries there is a greater need for the incorporation of mētis, of using the kind of information that is more difficult to make 'seeable' to a distant supervisor. The more predictable (the more naturally legible to a distant principal) the context, the less severe the 'penalty' for failing to incorporate soft information (mētis) into decision-making.

What scant outcome evidence there is seems to point in this direction regarding at least one IDO (the World Bank) by demonstrating that the WB does less well in less predictable contexts.[7] Pohl & Mihaljek (1992) find suggestive evidence that better recipient government economic management rating is associated with higher, and more certain, project returns. Chauvet, Collier, & Duponchel (2010) find that the probability of a World Bank's project success increases as peace lasts (as the country becomes more stable). A number of papers (Chauvet, Collier, and Fuster 2006; Denizer, Kaufmann, and Kraay 2011; Dollar and Levin 2005) find that better-quality recipient government institutions are associated with a greater probability of project success.

This work builds upon that logic by looking at the effect of state fragility *in interaction with* the field-level autonomy/control of IDOs. I expect, consistent with prior work, that situations with greater fragility (less stability) will be more difficult for all IDOs; however they will be differentially increasing in difficulty for IDOs with lower levels of autonomy. Key to the empirical strategy below, then, will be examining outcomes across a

[7] WB outcome data seems to be far and away the most accessible, if the large number of studies is any guide. There are no studies prior to this work, to my knowledge, that look at inter-IDO comparisons in a similar (quantitative) fashion.

variety of levels of IDO flexibility and recipient fragility. The hypothesized relationship is depicted in a stylized manner in figure 1 below.

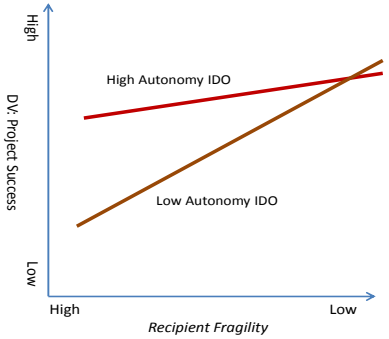

Figure 1: Hypothesized relationship between recipient fragility and project success for IDOs of differing autonomy levels

Task Domains: Interaction with Observability

Not all project types have an equal chance of success, ceteris paribus. Empirical evidence on World Bank projects demonstrates this as well (Chauvet, Collier, and Duponchel 2010; Pohl and Mihaljek 1992), finding (broadly) that transport and urban development projects are the outstanding performers;[8] there is also evidence that the institutions associated with good performance differ by the type of loan instrument and purpose employed (Dollar and Levin 2005).

The sectors in which success seems greatest – transport and urban development – are those seemingly with the smallest potential gap between 'de jure' and 'de facto' performance. Although a road that looks good on the surface can hide deep inequities (and indeed, Olken 2007 suggests that, due to corruption, this is sometimes the case), there is clearly less potential for decoupling of form and function in road construction than in e.g. the establishment of an anti-corruption commission. In such a case it will be easy to monitor that the building has been built and commissioners appointed but much more difficult to determine what is being done inside that building and whether those acts have a deterrent effect on others.

The notion that we might expect differences based on how well the monitorable proxies (existence of a road vs. existence of a building) are necessarily coupled with desired/intended *de facto* outcomes is supported by the contract theory literature on monitorability of tasks referenced above. This work investigates how variation in inherent legibility of tasks has on the returns to incorporating mētis via the autonomy of field-level agents.

[8] Doing better than education, agriculture, energy, and private sector development, among other sectors.

Moving from Theory to Empirics

In sum, IDOs are the kinds of organizations that we should expect, ex-ante, to be procedural (in the Wilsonian sense), with outcomes difficult to monitor. In this environment we should expect that an attempt by principals to control agents by writing tight contracts will not be successful in creating better outcomes; it will, however, succeed in constraining agents such that they will be unable to use mētis and have to rely on hard information. The expectation investigated by this work is that IDOs that allow (or simply have organizational features that augur towards, whether intended or not) higher levels of autonomy for field-level agents will be more successful. These returns to autonomy are expected to be higher in more fragile environments, and when the task domain is harder to monitor.

Description of Data

As noted above, this analysis uses a novel dataset composed of over 14,000 unique projects in 182 counties across nine donor agencies.[9] 75% of the projects have been completed since 1991 and 50% have been completed since 2002. [10]

The key outcome measure in the regression analysis below is overall project success, a holistic rating of the overall success of a project undertaken by independent evaluators (either external evaluation contractors or independent evaluation units) or by project staff in project completion reports.[11]

Organizational autonomy is proxied in two ways: by a scale drawn from the Paris Declaration monitoring indicators and by direct field surveys of aid experts.

To build the autonomy scale I take five measures which are indicative of either 1) IDOs' comfort level with/propensity to devolve control regarding project implementation to recipient countries or 2) the degree of autonomy the agency itself has relative to its political authorizing environment. [12] The first group are indicators of the extent to which an organization values control (and is thus a proxy for field-level autonomy of staff), the second indicators of the autonomy of the agency itself, which I theorize above is highly

[9]Those agencies are the UK's Department for International Development (DFID), the Japan International Cooperation Agency (JICA), KfW (the German Development Bank), GiZ (The German Society for International Cooperation) the International Fund for Agricultural Development (IFAD), the Asian Development Bank (AsDB), the World Bank (WB), the Global Fund for AIDS, Tuberculosis, and Malaria (GFATM), and the European Commission (EC). WB data used in this analysis are publicly available; JICA, GiZ, and IFAD data were assembled from individual project completion reports; AsDB, KfW, EC, GFATM, and DFID data were shared by the institutions themselves. I thank all of the agencies that provided data for this project, as well as all of those who were kind enough to respond to queries for their assistance.
[10] Or to be completed; a small number (less than 5%) of observations are for projects which have not yet been completed.
[11] To the extent I have coded and explored differences in this regard (which I plan to do more fully in the coming months) there seems to be no statistically significant relationship between the type of evaluation and evaluated project outcomes.
[12] These are indicators for monitoring the implementation of the 2005 Paris Declaration, which the official OECD website (http://www.oecd.org/dac/effectiveness/parisdeclarationandaccraagendaforaction.htm) describes as "a practical, action-oriented roadmap to improve the quality of aid and its impact on development. It puts in place a series of specific implementation measures and establishes a monitoring system to assess progress and ensure that donors and recipients hold each other accountable for their commitments."

correlated with the autonomy agencies give to field-level staff. The scale used here is a time-invariant measure formed from the average of the three (2005, 2007, and 2010) waves of the Paris Declaration survey.

The first set of indicators 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 (that is, separate units inside recipient countries which use donor standards and thus give donors more control/separation of funds or procurement). The second set are the degree to which aid is untied (i.e. not required that funds be spent on goods produced and contractors from the donor government; a high level is a sign of an IDO's need to build political consensus for aid giving via serving domestic political constituencies and thus a more insecure footing in the IDO's political authorizing environment) and the predictability of aid (the extent to which estimates of aid volume ex-ante are proved accurate ex-post; variations are a sign of IDO funding insecurity and/or political meddling by actors in the political authorizing environment, e.g. members of congress or executive boards). The two sub-scales are highly correlated and principal components analysis yields a single component with relatively equal weighting between measures. The results presented below are robust to the dropping of either subscale as well as dropping any single measure.[13]

In addition to this scale, I conducted direct field surveys of aid experts – individuals with substantial development experience, or with jobs (e.g. working for a recipient government's ministry of finance aid management unit) which brought them into contact with a wide swath of donors, with a concentration on those nationals and internationals with expertise in Liberia and South Africa (as these are case study countries for related qualitative work). Respondents rated a number of development agencies (including, but not limited to, those in the sample) on a scale of 1-7 in response to the following question:

> "To what degree do you believe the in-country field office/bureaus 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.[14] The Asian Development Bank is not included in the survey as (unsurprisingly, given the focus on Sub-Saharan Africa) there were no respondents who responded to the autonomy question regarding the AsDB. The correlation between the scale and the survey means for each donor is .71.

[13] The autonomy scale is a simple average of the five measures except in the case of multilaterals, for which tied aid is not calculated; in these cases the scale is an average of the remaining four measures.

[14] This is the remaining N after removing surveys which were not substantively responsive, or gave indications of nonsense answers; the 3 largest reasons for exclusion wererating 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 rating the survey's anchoring vignettes (these anchors are not yet incorporated into the survey evaluation) such that the most autonomous text was evaluated as the same or less autonomous than the least autonomous text.

State fragility is measured via the Polity IV/Integrated Network for Societal Conflict Research (INSCR)'s State Fragility Index (Center for Systemic Peace 2012). OECD Development Assistance Committee (DAC) sector codes are standard classifications that are in most assigned by the IDOs themselves either in their databases/project report or their reports on aid flows to DAC.[15]

All specifications below use IDO fixed effects, thus accounting for any (constant) systematic differences between IDO evaluation criteria or measurement standards. This specification is also why the autonomy scale itself does not appear as an independent variable in the models below (only appearing as part of an interaction term); as it is fixed by IDO, it is collinear with IDO fixed effects.

Results

Table 1 below presents basic summary statistics for the few variables that form the core of the analysis.

Table 1: Summary Statistics for core variables					
	N	Mean	SD	Min	Max
Overall Project Success (6 point scale)	14048	4.23	1.2	1	6
Overall Project Success (z scores)	14048	$1.36*10^{-8}$	.9997	-3.53	2.01
State Fragility Index	8938	12.5	5.05	0	25
Project Size (USD Millions)	9712	36.115	116.57	$1.36*10^{-7}$	6309[16]
Autonomy scale	14424	.655	.059	.564	.790

Project success is an ordinal variable ranging from 1-6; I also employ a z-transformed (by IDO) version of this variable, thus allowing for the fitting of OLS (rather than ordered logit) models.

The coverage of the State Fragility Index – one of the key covariates – is limited to 2005-2011 at present, which limits the analysis to the 62% (8938) of projects that fall in that time period. This is also the range over which the Paris Declaration monitoring surveys from which the autonomy scale is drawn were conducted.

Main Results

Tables 2 and 3 below (using ordered logit on the 6 point scale and OLS on the z-scores of outcome, respectively) demonstrate the core findings from these results using the autonomy scale. As expected, there is a robust and statistically significant negative relationship between level of state fragility and evaluations of project success; projects in more fragile (more stable) environments are less (more) successful. This relationship is mitigated by increasing autonomy to field-level personnel; organizations with greater

[15] In a small number (fewer than 3%) of cases codes are assigned by me/research assistants based on the detailed contents of project reports.

[16] Investigating the dozen or so suspicious-looking outliers in project size on a project-by-project basis (to ensure no reporting errors by IDOs etc.) is one of my tasks for the coming months.

(lower) levels of autonomy have less (more) pronounced relationships between state fragility and project success; the results of their projects are less (more) negatively affected by an increase in fragility. These relationships are robust to the inclusion of project size as a control variable (under the logic that agencies might place differential attention on projects of different sizes).

Table 2: Relationship Between Autonomy and State Fragility (Ordered Logit)

DV: Overall Project Success (6 pt scale)	1	2	3	4
State Fragility Index (SFI)	-0.051	-0.042	-0.271	-0.257
	(12.80)***	(8.86)***	(7.71)***	(6.34)***
Autonomy (scale)*SFI			0.328	0.310
			(6.31)***	(5.33)***
Project Size (USD Millions)		0.001		0.001
		(3.68)***		(3.69)***
Constant	1.600	1.844	1.426	1.548
	(26.89)***	(20.61)***	(21.75)***	(14.14)***
IDO Fixed Effects	Y	Y	Y	Y
N	8,702	6,911	8,702	6,911
Pseudo-R^2	.02	.02	.02	.02

* p<0.1; ** p<0.05; *** p<0.01

Table 3: Relationship Between Autonomy and State Fragility (Multilevel OLS)

DV: Overall outcome(Z-score)	1	2	3	4
State Fragility Index (SFI)	-0.027	-0.024	-0.143	-0.147
	(12.78)***	(9.03)***	(7.36)***	(6.55)***
Autonomy (scale)*SFI			0.174	0.179
			(6.01)***	(5.54)***
Project Size (USD Millions)		0.000		0.000
		(4.06)***		(4.07)***
Constant	0.335	0.286	0.322	0.303
	(11.67)***	(7.02)***	(11.24)***	(7.52)***
IDO Fixed Effects	Y	Y	Y	Y
R^2-Within	0.02	0.01	0.02	0.02
R^2-Between	0.22	0.04	0.01	0.06
N	8,702	6,911	8,702	6,911

* p<0.1; ** p<0.05; *** p<0.01

The basic message of these results is presented graphically in figure 2 below, which is drawn from model 4 of table 3 above. A hypothetical IDO with an autonomy level comparable to that of the EC (.56) relative to DFID (.79) sees on average a much greater difference in performance in a country of relatively high fragility (e.g. a country with a state fragility index comparable to that of Liberia) as compared to a country of relatively low

fragility (e.g. a country with a state fragility index comparable to that of South Africa).[17]

The substantive implication of these changes is not insubstantial, with a gap of about ¾ of a standard deviation between the performance of the two hypothetical IDOs. This is the equivalent of about 1 point on the 6-point outcome scale.

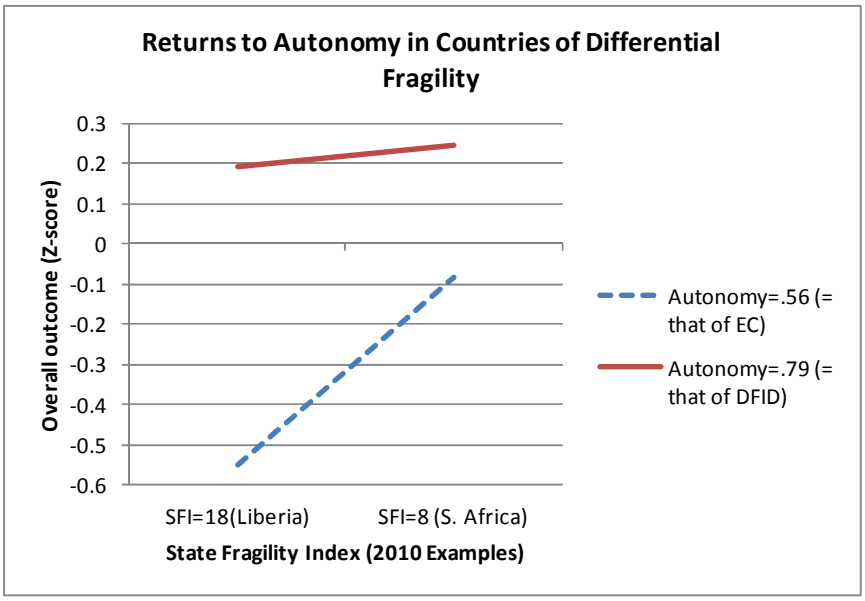

Figure 2: Fitted plot drawn from Table 3, Model 4

Table 4 below incorporates recipient country fixed effects; results are substantively unchanged, suggesting that as the fragility level of countries change the relationship between state fragility and outcome continues to be mediated by autonomy. It seems not to be the case, then, that time-invariant features of the countries (or the set of countries in which various IDOs have more/fewer programs due to idiosyncratic features, e.g. colonial ties) are driving the findings presented here.

[17] This is not a comparison between the EC and DFID nor between South Africa and Liberia; figure 2 is a fitted plot drawn from the central trends of the overall model, using values equivalent to real-world examples to create context.

Table 4: Relationship Between Autonomy and State Fragility
with Recipient Country Fixed-Effects (OLS)

DV: Overall outcome(Z-score)	1	2	3	4
State Fragility Index (SFI)	-0.010	-0.012	-0.108	-0.114
	(1.42)	(1.41)	(4.62)***	(4.44)***
Autonomy (scale)*SFI			0.148	0.152
			(4.40)***	(4.08)***
Project Size (USD Millions)		0.000		0.000
		(2.91)***		(2.91)***
Constant	-0.175	0.245	-0.243	0.180
	(0.19)	(0.26)	(0.26)	(0.19)
IDO Fixed Effects	Y	Y	Y	Y
Recipient Country Fixed Effects	Y	Y	Y	Y
R^2-Within	0.07	0.07	0.08	0.08
R^2-Between	0.00	0.14	0.00	0.12
N	8,702	6,911	8,702	6,911

* p<0.1; ** p<0.05; *** p<0.01

Table 5 below runs models 3&4 (including the Autonomy/SFI interaction) of Tables 2-4 above, substituting autonomy survey responses for the autonomy scale. All results are robust to this substitution, which should provide confidence in the validity of the autonomy scale measure.

Table 5: Substituting Autonomy Survey for Scale

	1:Ordered Logit, DV 6 pt scale	2:Ordered Logit, DV 6 pt scale	3:OLS, DV z-score	4:OLS, DV z-score	5:OLS, DV z-score	6:OLS, DV z-score
State Fragility Index (SFI)	-0.142	-0.107	-0.088	-0.080	-0.058	-0.051
	(5.04)***	(3.19)***	(5.73)***	(4.26)***	(3.17)***	(2.35)**
Autonomy(survey)*SFI	0.023	0.016	0.016	0.014	0.012	0.010
	(3.38)***	(2.08)**	(4.11)***	(3.18)***	(2.95)***	(2.03)**
Project Size (USD Mil)		0.001		0.000		0.000
		(2.72)***		(3.04)***		(1.23)
Constant	1.506	1.562	0.285	0.246	0.249	0.284
	(22.40)***	(14.48)***	(9.66)***	(6.40)***	(0.26)	(0.30)
IDO Fixed Effects	Y	Y	Y	Y	Y	Y
Recipient Fixed Effects					Y	Y
R^2-Within/Pseudo-R^2	0.01	0.02	0.02	0.01	0.07	0.07
R^2-Between			0.02	0.15	0.02	0.13
N	7,766	5,975	7,766	5,975	7,766	5,975

* p<0.1; ** p<0.05; *** p<0.01

Table 6 below summarizes the relationship between state fragility and project success for each donor in isolation – that is, using only data from one donor at a time – further confirming the role of higher levels of autonomy in mitigating the inverse relationship between the State Fragility Index (SFI) and project success in a way that does not rely on the parameterization of the interaction term.

Table 6: Relationship Between SFI and Success for Sample IDOs in Isolation

IDO	*Autonomy Scale Score from Paris Declaration Survey Measures (Range: 0-1)*	*Autonomy Survey Mean from Field Surveys (Range: 1-7)*	*Relationship between SFI and Success for this donor with only this donor's data included in regression (Z-score)*
EC	.564	3	-.025***
Global Fund	.603	5	-.040***
WB	.622	3.65	-.036***
Asian DB[18]	.651		-.064***
JICA	.661	4.67	-.017
GiZ[19]	.674	4.33	-.05**
KfW	.674	4	-.033***
IFAD	.721	6	-.017
DFID	.790	4.71	-.002

Comparing Autonomy Scale to Other Plausible Measures

It seems natural to investigate if other scales – other measures of aid, or organizational, quality – yield similar results. Is autonomy unique in its relationship with success, or is it just that autonomy is one of a larger package of good donor practices which are related to relative success in more fragile contexts? Table 7 below gives summary statistics on two alternate scales – the Commitment to Development Index (CDI) and the Quality of Official Development Assistance (QuODA).[20] In both cases I also look at the subscales which seem most relevant – in the case of CDI, the aid component; in the case of QuODA, the Maximizing Efficiency and Fostering Institutions subscales. It is worth noting that there is a degree of overlap between these measures and the autonomy scale (which is re-presented below for ease of reference). The CDI aid index also penalizes tied aid (a component of the autonomy scale); untied aid is also a component of QuODA's Maximizing

[18] As noted above, no survey respondents answered the autonomy question with regards to the Asian Development Bank; thus there is no AsDB score reported.
[19] The Paris Declaration monitoring survey does not differentiate between institutions from a single country; thus GiZ and KfW (both arms of the German government) have the same autonomy score.
[20] The CDI is an annual product of the Center for Global Development; QuODA is an occasional (last wave was 2010) product of the Brookings Institution in collaboration with the Center for Global Development. The CDI has a number of components (Aid, Investment, Migration, Environment, Security, and Technology) which assess the commitment of nations (multilateral organizations such as the WB are not included) to assisting the developing world. QuODA has four components – Maximizing Efficiency, Transparency & Learning, Reducing Burden, and Fostering Institutions. All components in both the CDI and QuODA involve a variety of sub-measures.

Efficiency measure. QuODA's Fostering Institutions component draws from the Paris Declaration monitoring surveys as well, incorporating avoidance of project implementation units and use of recipient country systems.[21]

Table 7: Summary Statistics for Alternate Scales					
	N	Mean	SD	Min	Max
Autonomy scale	14424	.655	.059	.564	.790
Commitment to Development Index (CDI) 2012 Overall	4867	5.27	.71	3.4	5.7
Commitment to Development Index (CDI) 2012 Aid	4867	4.76	1.79	1.6	6.8
Quality of Development Assistance (QuODA) 2010 Overall	14297	.529	.147	.0425	.713
Quality of Development Assistance (QuODA) 2010 Maximizing Efficiency	13703	.161	.276	-.89	.51
Quality of Development Assistance (QuODA) 2010 Fostering Institutions	14297	.392	.276	-.1	.93

Table 8 below shows none of the other measures have anywhere near the strength of association of the autonomy scale. A better QuODA overall score and a better score on QuODA's maximizing efficiency subscale yields results in the opposite direction as that of autonomy, with higher scores associated with a stronger relationship between state fragility and evaluated project success. QuODA's fostering institutions measure moves in the same direction (unsurprising inasmuch as 2 of the 8 indicators are included in the autonomy measure), but with a very small point estimate. CDI's aid measure (which also has a small amount of overlap with the autonomy measure) has similar results as QuODA's.

[21] It's not entirely clear what this last measure incorporates, but – given that the stated source is the Paris Declaration monitoring reports – it seems reasonable to assume this combines the procurement and public financial management measures I use in the autonomy scale.

Table 8: Relationship Between Alternative Scales and State Fragility (OLS)

DV: Overall outcome(Z-score)	Autonomy (scale)	CDI (Overall)	CDI (Aid)	QuODA (Overall)	QuODA (Maximizing Efficiency)	QuODA (Fostering Institutions)
State Fragility Index (SFI)	-0.147	-0.047	-0.048	-0.003	-0.023	-0.038
	(6.55)***	(1.59)	(4.23)***	(0.32)	(8.37)***	(6.83)***
Scale in Column Title*SFI	0.179	0.006	0.006	-0.045	-0.028	0.030
	(5.54)***	(1.17)	(3.26)***	(2.58)**	(3.71)***	(3.02)***
Project Size (USD Mil)	0.000	0.000	0.000	0.000	0.001	0.000
	(4.07)***	(1.93)*	(1.91)*	(4.11)***	(4.36)***	(4.14)***
Constant	0.303	0.181	0.193	0.290	0.310	0.278
	(8.16)***	(3.53)***	(3.79)***	(7.71)***	(7.88)***	(7.41)***
IDO Fixed Effects	Y	Y	Y	Y	Y	Y
R^2	0.02	0.01	0.01	0.01	0.02	0.01
N	6,911	3,501	3,501	6,815	6,261	6,815

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Incorporating Observability of Outcomes

Given that the theory underlying this analysis rests upon the ability of mētis to be incorporated into the direction of development projects, it seems logical to investigate whether results differ based on the extent to which outputs and outcomes can be observed from afar, and thus the extent to which theory would suggest mētis/soft information is needed as part of the process of producing good development results.

Tables 9 and 10 below attempt to do this, examining these relationships in a variety of OECD-DAC sectors. In sectors (STD control including HIV/AIDS, road infrastructure construction, electrical transmission/distribution, building educational facilities, and elections monitoring) where outcomes are relatively easily observed (either because they involve physical construction or because there are other clearly observed outcomes – ballots tampered with, patients on treatment/drug delivered, etc.), there is no statistically significant relationship between the interaction of autonomy and state fragility and project success (though regarding road infrastructure the interaction comes close to the 90% significance level). In some of the sectors (democratic participation and civil society as well as transport management, which is separated from actual construction in the transportation sector) where it is more difficult to observe outcomes, greater levels of autonomy continue to lessen the inverse relationship between state fragility and project success.

Table 9: Relationship Between Autonomy and State Fragility (Ordered Logit) by Sector (Outcomes Easily Observed; sector by CRS code)

DV: Overall outcome(6 pt scale)	STD Control (Including HIV/AIDS)	Road Infrastructure	Electrical Transmission/ Distribution	Building Educational Facilities	Elections Monitoring
State Fragility Index (SFI)	-0.311	-0.919	-0.284	-0.589	2.554
	(1.11)	(1.66)*	(0.26)	(0.80)	(1.94)*
Autonomy*SFI	0.363	1.255	-0.069	0.895	-4.192
	(0.82)	(1.47)	(0.04)	(0.77)	(2.29)**
Size of project (USD mil)	0.000	0.001	0.002	0.001	0.094
	(0.04)	(0.27)	(0.90)	(0.68)	(1.99)**
IDO Fixed Effects	Y	Y	Y	Y	Y
N	253	300	141	38	19
Pseudo-R^2	.04	.04	.06	.09	.37

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 10: Relationship Between Autonomy and State Fragility (Ordered Logit) by Sector (Outcomes Difficult to Observe; sector by CRS code)

DV: Overall outcome(6 pt scale)	Public Sector Policy and Administration	Democratic Participation and Civil Society	Transport Management
State Fragility Index (SFI)	-0.227	-0.821	-4.145
	(0.96)	(1.60)	(3.82)***
Autonomy*SFI	0.274	1.40	6.09
	(0.81)	(1.66)*	(3.89)***
Size of project (USD millions)	0.003	0.21	0.031
	(1.23)*	(1.19)	(2.37)**
IDO Fixed Effects	Y	Y	Y
N	191	51	35
Pseudo-R^2	.02	.04	.20

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

All results presented above are robust to the use of both OLS (using z-transformed project success as the DV) and ordered logit models. The robustness of these results to (donor-specific) z-transformation of outcomes strongly suggests the parameterization of the dependent variable is not driving the results. These results are also robust to transforming the outcome into a simple binary variable (and thus fitting logit models)

using either 4 or 5 (of 6) as the lowest level at which a project is considered a success. These results are also robust to constructing the autonomy measure from only the first (2005) wave of the Paris Declaration survey.[22]

Discussion/Conclusion

There seems to be reasonably strong evidence that greater levels of autonomy are associated with a reduction in the difference between performance in countries of differing levels of state fragility. This effect seems to be more pronounced in sectors where it is more difficult to externally observe and externally verify (from e.g. a distant headquarters), and thus contract on, outcomes.

This suggests that IDOs with higher levels of autonomy are better able to gather and incorporate the kind of soft information so critical to the success of development interventions. It is also possible that organizations with greater levels of field level autonomy do a better job of recruiting, retaining, and promoting (externally unobservably) higher quality personnel, as individuals gravitate towards institutions that facilitate and reward initiative-taking and deep engagement in the environments in which they work.

The data presented above suggest that organizational features 1) vary across IDOs and 2) have systematic associations with the relative success of projects. This suggests that there may be returns to focusing on comparative organizational features in an effort to improve the performance of development projects, particularly in contexts (such as more fragile states) where development interventions have relatively large potential impact. This is perhaps the 'low hanging fruit' of international development, inasmuch as it focuses on features that – unlike corruption or political will – are entirely within the control of organizations themselves.

Quantitative analysis can potentially play an important role in this effort in concert with qualitative work and careful organizational assessment. In addition to decision-making regarding program design, supervision, and revision, human resource systems (including how individuals are evaluated in the context of career/promotion concerns) and staff rotation practices are potentially interesting areas of examination in this regard.

[22] Relevant were one to be concerned that IDOs might have responded to the Paris Declaration monitoring regime over time and altered their practices with regards to what was measured without changing their underlying way of working/the autonomy given to field level staff.

Bibliography

"Accra Agenda for Action." 2008. In Accra, Ghana: Available online at Http://www.oecd.org/dataoecd/11/41/34428351.pdf.

Aghion, Philippe, and J Tirole. 1997. "Formal and Real Authority in Organizations." *Journal of political economy* 105(1): 1–29.

Allegret, Jean-Pierre, and Philippe Dulbecco. 2007. "The Institutional Failures of International Monetary Fund Conditionality." *The Review of International Organizations* 2(4): 309–27.

Ascher, W. 1983. "New Development Approaches and the Adaptability of International Agencies: The Case of the World Bank." *International Organization* 37(3).

Azam, Jean-Paul, and Jean-Jacques Laffont. 2003. "Contracting for Aid." *Journal of Development Economics* 70(1): 25–58.

Babb, Sarah. 2003. "The IMF in Sociological Perspective : A Tale of Organizational Slippage." *Studies in Comparative International Development* 38(2): 3–27.

Barnett, Michael N., and Martha Finnemore. 2004. *Rules For The World: International Organizations In Global Politics*. Cornell University Press.

Bird, Graham. 1998. "The Effectiveness of Conditionality and the Political Economy of Policy Reform: Is It Simply a Matter of Political Will?" *The Journal of Policy Reform* 2(1): 89–113.

Birdsall, Nancy. 2008. "No Seven Deadly Sins: Reflections on Donor Failings." In *Reinventing Foreign Aid*, ed. William Easterly. MIT Press, 515–51.

Brechin, Steven. 1997. *Planting Trees in the Developing World: a Sociology of International Organizations*. Johns Hopkins University Press.

"Busan Partnership For Effective Development Co-Operation." 2011. In *4th High Level Forum on Aid Effectivness*, Busan, South Korea.

Calvo, G.A., and Stanislaw Wellisz. 1978. "Supervision, Loss of Control, and the Optimum Size of the Firm." *The Journal of Political Economy* 86(5): 943–52.

Canales, Rodrigo. 2010. "Weaving Straw into Gold: Rule Bending, Localism, and Managing Inconsistencies in Organizational Rules." 1–51.

Carpenter, Daniel P. 2001. *The Forging of Bureaucratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies, 1862-1928*. Princeton University Press.

Center for Systemic Peace. 2012. “State Fragility Index.”

Chauvet, L, P Collier, and M Duponchel. 2010. “What Explains Aid Project Success In Post-Conflict Situations?” *World Bank Policy Research Working Paper*.

Chauvet, L, P Collier, and A Fuster. 2006. “Supervision and Project Performance: a Principal-agent Approach.”

Cooley, Alexander. 2005. *Logics of Hierarchy: The Organization of Empires, States, And Military Occupations*. Cornell University Press.

Cooley, Alexander, and James Ron. 2002. “The NGO Scramble.” *International Security* 27(1): 5–39.

Denizer, Cevdet, Daniel Kaufmann, and Aart Kraay. 2011. “Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance.” *World Bank Policy Research Working Paper 5645* (May).

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole. 1999. “The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies.” *Review of Economic Studies* 66(1): 199–217.

DiIulio, John. 1987. *Governing Prisons*. Free Press.

Dixit, A. 1997. “Power of Incentives in Private Versus Public Organizations.” *The American Economic Review* 87(2).

———. 2002. “Incentives and Organizations in the Public Sector: An Interpretative Review.” *Journal of Human Resources* 37(4): 696–727.

Dobbin, F., and T. Boychuk. 1999. “National Employment Systems and Job Autonomy: Why Job Autonomy Is High in the Nordic Countries and Low in the United States, Canada, and Australia.” *Organization Studies* 20(2): 257–91.

Dollar, D, and V Levin. 2005. “Sowing and Reaping: Institutional Quality and Project Outcomes in Developing Countries.” *World Bank Policy Research Working Paper*.

Dreher, Axel, Jan-Egbert Sturm, and James Raymond Vreeland. 2009. “Global Horse Trading: IMF Loans for Votes in the United Nations Security Council.” *European Economic Review* 53(7): 742–57.

Friebel, Guido, and Michael Raith. 2004. “Abuse of Authority and Hierarchical Communication.” *The RAND Journal of Economics* 35(2): 224.

Gibson, Clark, Krister Andersson, Elinor Ostrom, and Sujai Shivakumar. 2005. *The Samaritan's Dilemma: The Political Economy of Development Aid*. Oxford University Press.

Gould, E. 2006. "Delegating IMF Conditionality: Understanding Variations in Control and Conformity." In *Delegation and Agency in International Organizations*, eds. Darren G. Hawkins, David A. Lake, Daniel L. Nielson, and Michael J. Tierney. , 281–311.

Grossman, Sanford J., and O.D. Hart. 1983. "An Analysis of the Principal-agent Problem." *Econometrica: Journal of the Econometric Society* 51(1): 7–45.

Hart, Oliver, and Bengt Holmstrom. 1987. "The Theory of Contracts." In *Advances in Economic Theory*, ed. T.F. Bewley. , 71–155.

Heimer, Carol. 2013. "Resilience in the Middle: Contributions of Regulated Organizations to Regulatory Success." *The ANNALS of the American Academy of Political and Social Science* 649(1): 139–56.

Holmstrom, B. 1999. "Managerial Incentive Problems: A Dynamic Perspective." *Review of Economic Studies* 66: 169–82.

Huber, John D., and Charles R. Shipan. 2002. *Deliberate Discretion: The Institutional Foundations of Bureaucratic Autonomy*. Cambridge University Press.

Johns, Leslie. 2007. "A Servant of Two Masters: Communication and the Selection of International Bureaucrats." *International Organization* 61(02): 245–75.

Kaufman, Herbert. 2006. *The Forest Ranger: A Study in Administrative Behavior*. Resources for the Future.

Kiser, Edgar, and Danielle Kane. 2007. "The Perils of Privatization How the Characteristics of Principals Affected Tax Farming in the Roman Republic and Empire." *Social Science History* 2(Summer 2007).

Kofman, Fred, and Jacques Lawarree. 1993. "Collusion in Hierarchical Agency." *Econometrica* 61(3): 629–56.

Kuziemko, Ilyana, and Eric Werker. 2006. "How Much Is a Seat on the Security Council Worth? Foreign Aid and Bribery at the United Nations." *Journal of Political Economy* 114(5): 905–30.

Laffont, JJ, and Jean Tirole. 1988. "The Dynamics of Incentive Contracts." *Econometrica: Journal of the Econometric Society* 56(5): 1153–75.

Lawrence, Paul R., and Jay William Lorsch. 1967. *Organization and Environment: Managing Differentiation and Integration*. Harvard Business School Press.

Lipsky, Michael. 2010. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*. Russell Sage Foundation.

Lyne, M, Daniel L. Nielson, and Michael J. Tierney. 2006. "A Problem of Principals: Common Agency and Social Lending at the Multilateral Development Banks." In *Delegation and Agency in International Organizations*, eds. Darren G. Hawkins, David A. Lake, Daniel L. Nielson, and Michael J. Tierney. , 41–76.

March, James G., and Herbert Alexander Simon. 1958. *Organizations*. Wiley.

Martin, LL. 2006. "Distribution, Information, and Delegation to International Organizations: The Case of IMF Conditionality." In *Delegation and agency in International Organizations*, eds. Darren G. Hawkins, David A. Lake, Daniel L. Nielson, and Michael J. Tierney. , 140–64.

Momani, Bessma. 2006. "IMF Staff: Missing Link in Fund Reform Proposals." *The Review of International Organizations* 2(1): 39–57.

Murshed, S. Mansoob. 2009. "On the Non-Contractual Nature of Donor-Recipient Interaction in Development Assistance." *Review of Development Economics* 13(3): 416–28.

Ness, GD, and Steven Brechin. 1988. "Bridging the Gap: International Organizations as Organizations." *International Organization* 30(March 1986): 39–57.

Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115(2): 200–249.

Ostrom, Elinor, Clark Gibson, and Sujai Shivakumar. 2002. *Aid, Incentives , and Sustainability: An Institutional Analysis of Development Cooperation*.

Pohl, Gerhard, and Dubravko Mihaljek. 1992. "Project Evaluation and Uncertainty in Practice: A Statistical Analysis of Rate-of-return Divergences of 1,015 World Bank Projects." *The World Bank Economic Review* 6(2): 255–77.

Pritchett, Lant, and Michael Woolcock. 2004. "Solutions When the Solution Is the Problem: Arraying the Disarray in Development." *World Development* 32(2): 191–212.

Sarfaty, GA. 2009. "Why Culture Matters in International Institutions: The Marginality of Human Rights at the World Bank." *American Journal of International Law* 103(4): 647–83.

Scott, James C. 1998. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Yale University Press.

Southard, Frank. 1980. “The Evolution of the International Monetary Fund.” *North Carolina Journal of International Law and Commercial Regulation*.

Stein, JC. 2002. “Information Production and Capital Allocation: Decentralized Versus Hierarchical Firms.” *The Journal of Finance* LVII(5).

Strausz, Roland. 1997. “Delegation of Monitoring in a Principal-Agent Relationship.” *The Review of Economic Studies* 64(3): 337.

Tendler, Judith. 1975. *Inside Foreign Aid*. Johns Hopkins University Press.

“The Paris Declaration on Aid Effectiveness.” 2005. In Paris, France: Avaliable Online at http://www.oecd.org/dataoecd/11/41/34428351.pdf.

Thompson, James D. 1967. *Organizations in Action: Social Science Bases of Administrative Theory*. Transaction Publishers.

Tirole, Jean. 1986. “Hierarchies and Bureaucracies: On the Role of Collusion in Organizations.” *Journal of Law, Economics, & Organization* 2(2).

———. 1994. “The Internal Organization of Government.” *Oxford Economic Papers* 46(1): 1–29.

UNDP. 2010. *Human Development Report 2010 The Real Wealth of Nations: Pathways to Human Development*. United Nations Development Programme.

Vaubel, Roland. 1996. “Bureaucracy at the IMF and the World Bank: A Comparison of the Evidence.” *The World Economy* 19(2): 195–210.

Weber, Max. 1922. “Bureaucracy.” In *Wirtshaft und Gesselschaft (Economy and Society) Volume II*, eds. Guenther Roth and Claus Wittich. University of California Press, 956–1005.

Whittle, Dennis, and Mari Kuraishi. 2008. “Competing with Central Planning: Marketplaces for International Aid.” In *Reinventing Foreign Aid*, ed. William Easterly. MIT Press, 461–84.

Willett, Thomas D. 2001. “Upping the Ante for Political Economy Analysis of the International Financial Institutions.” *The World Economy* 24(3): 317–32.

Wilson, James Q. 1989. *Bureaucracy: What Government Agencies Do And Why They Do It*. Basic Books.

World Bank. 2011. *World Development Report 2011: Conflict, Security, and Development*. World Bank.