

Aiming at the Wrong Targets: The Difficulty of Improving Domestic Institutions with International Aid

Benjamin P. Buch, Mark T. Buntaine, Bradley C. Parks

Abstract:

We explain why the record of aid agencies in building and reforming public sector institutions in developing countries has been broadly unsuccessful, despite extraordinary amounts of time, money, effort, and a commitment to achieve targets. We argue that requirements to specify and monitor observable indicators of success have created strong incentives for aid-dependent countries to signal performance to their foreign sponsors by achieving targets. However, in the absence of requirements about the types of targets that should be pursued, countries that rely heavily upon external sources of financial support select easy targets that have limited value for strengthening public sector institutions. In particular, aid-dependent countries are more likely to select targets that measure how public sector institutions are organized, rather than targets that measure what policy outcomes are achieved through strengthened public sector institutions. We demonstrate that this argument has both explanatory and predictive power for World Bank environment and natural resource management projects.

Acknowledgement: The authors would like to thank Scott Laws, Steven Linett, Rebecca Schectman, Joseph Dobbels, Kelsey Sakumoto, James Marino, and Rob Marty for their excellent research assistance. The authors are also grateful to Matt Andrews, Stephen Krasner, Emma Feutl Kent, and participants in workshops at APSA, PEIO, ISA, Stanford, and UC Santa Barbara for their helpful comments on earlier drafts of this paper.

Keywords: Aid, Environmental Management, Institutions, International Development, Performance Targets, Results Measurement, World Bank

The views expressed in AidData Working Papers are those of the authors and should not be attributed to AidData, funders of AidData's work, or to the institutions the authors represent.

Aiming at the Wrong Targets: The Difficulty of Improving Domestic Institutions with International Aid

Benjamin P. Buch

Correspondence to:
bpbuch@stanford.edu 


Benjamin Buch is a Political Science PhD student at Stanford University. His research focuses on norms and institutional behavior in the national security sphere, with a particular focus on nuclear security.

Mark T. Buntaine

Correspondence to:
buntaine@bren.ucsb.edu 

Mark Buntaine is an Assistant Professor of International Relations and Environmental Policy at the University of California, Santa Barbara and holds a PhD in Environmental Politics & Policy/Research Methods from Duke University. His research investigates the political economy of environmental policymaking in developing countries, with a particular emphasis on the targeting and impact of foreign aid.

Bradley C. Parks

Correspondence to:
bcpark@wm.edu 

Brad Parks is Co-Executive Director of AidData and Research Faculty at the College of William and Mary's Institute for the Theory and Practice of International Relations. He previously administered MCC's annual country selection process and managed anti-corruption and judicial reform programs in Indonesia and the Philippines. He has authored books and articles on aid allocation, aid effectiveness, and development theory and practice.

AidData

Correspondence to:
info@aiddata.org 
www.aiddata.org 

AidData – a joint venture of the College of William and Mary, Development Gateway and Brigham Young University – is a research and innovation lab that seeks to make development finance more transparent, accountable, and effective. Users can track over \$40 trillion in funding for development including remittances, foreign direct investment, aid, and most recently US private foundation flows all on a publicly accessible data portal on AidData.org. AidData's work is made possible through funding from and partnerships with USAID, the World Bank, the Asian Development Bank, the African Development Bank, the Islamic Development Bank, the Open Aid Partnership, DFATD, the Hewlett Foundation, the Gates Foundation, Humanity United, and 20+ finance and planning ministries in Asia, Africa, and Latin America.

Contents

Abstract.....	1
1. Introduction	4
2. Institutional Targets in Development Assistance	6
3. The Selection of Institutional Targets is a Strategic Choice.....	7
3.1 Rewards for Achieving and Maintaining Institutional Form and Function Targets	9
4. Measuring the Choice of Targets for Institutional Development	12
5. Model and Results: The Political Economy of Institutional Targets	14
6. Validating Models with Out-of-Sample Prediction	18
6.1 Testing the “Difficulty of Form and Function Target Achievement” Assumption	22
7. Conclusion	27
References.....	31

1. Introduction

Donor agencies spend an extraordinary amount of time, money, and effort trying to improve the performance of public sector institutions in developing countries. Common wisdom holds that “institutions matter” and that development rarely occurs in the absence of government agencies that have the capacity and incentives to discharge the state’s basic functions (Acemoglu and Robinson 2012; Besley and Persson 2010; Callaghy 1988; Haggard 1990; Herbst 1990; Rodrik 2000). Strong government agencies provide law and order, enforce the rules of economic exchange, raise revenue, and deliver essential public services (Acemoglu and Robinson 2012; Bates 2009; Besley and Persson 2010; North et al. 2007; Rodrik 2000). In the environment sector, the management of pollution and the rational use of natural resources depend on government agencies that can plan effectively, monitor and enforce regulations, and respond to citizen concerns (Esty and Porter 2005; Keohane and Levy 1996; VanDeveer and Dabelko 2001). However, there is little evidence that external development assistance systematically contributes to improved institutional performance in developing countries (Birdsall 2007; Booth 2011; Andrews 2013).

This lack of success is puzzling since both donor agencies and recipient countries have faced growing pressure in recent decades to measure and deliver observable outcomes as part of the “results agenda” in development assistance (Birdsall 2008; Clemens et al. 2007; Ebrahim 2013; Khagram et al. 2009; Natsios 2010; Ramalingam 2013). The World Bank, for example, requires the identification of observable targets and a plan to monitor them before any project is approved. It also conditions future funding on the achievement of these targets (World Bank 2010b).

We offer an explanation for why intensifying pressure for “measurable results” from aid agencies and donor countries has not led to substantial improvements in institutional performance among countries that receive development assistance for this purpose. We argue that, among countries that are more dependent on external sources of funding, performance targets have introduced strong and perverse incentives to signal success to donor agencies without addressing the underlying sources of institutional dysfunction as intended.

In the absence of requirements about what types of targets should be pursued, countries that are dependent on aid select easy targets that have limited value for strengthening public sector institutions. In particular, they are more likely to select easy targets that measure how public sector institutions are organized, rather than more difficult targets that measure the policy outcomes achieved through strengthened institutions. By contrast, achieving easy targets that measure how public sector institutions are organized yields few external rewards for countries that are not heavily dependent on aid, making

these countries more likely to select more difficult targets that measure how public sector institutions solve problems and achieve policy targets. These different payoffs for countries with varying levels of aid dependence result in a counterproductive outcome: those countries most in need of institutional development are the least likely to select targets that promote measurable improvements in institutional function.

To test this argument, we examine institutional development targets that were selected to measure success in World Bank environment and natural resource management projects. This sector provides an ideal empirical setting to test our theoretical argument because environmental aid programs focus heavily on improving public sector institutions (Keohane and Levy 1996; VanDeveer and Dabelko 2001), and donors have stronger preferences than recipients for improved environmental management, which should heighten incentives for aid-dependent countries to signal their performance to external actors (Keohane and Levy 1996; Hicks et al. 2008; Marcoux and Urpelainen 2012; Aklin and Urpelainen 2014; McLean forthcoming).

We created two original datasets by extracting targets related to institutional development from pre-project appraisal reports and post-project evaluations, which we use for an in-sample explanatory model and an out-of-sample predictive model about the choice of targets. After compiling the targets used to measure success, we coded them based on whether they measure the achievement of a policy outcome or the way that a public sector institution is organized. We call these institutional function and form targets, respectively. We show that if a country is more dependent on concessional finance from the World Bank, it is more likely to choose form-based targets, which have limited value for institution building and reform. To rule out the possibility that these countries simply have fewer institutions in place and thus need to pursue form targets as precursors to function targets, we show that the timing of the establishment of national environmental ministries and environmental legislation does not change our estimates of how targets are chosen.

We find that the countries most in need of institutional development are least likely to select targets that can lead to improved institutional performance. Our argument and findings have several important implications. First, our results contribute to a growing body of literature on how bargaining dynamics between states and international organizations influence the design and implementation of aid contracts (Marcoux and Urpelainen 2012; Helmke and McLean 2014; Bayer et al. forthcoming; McLean forthcoming). We show that by setting targets, international development organizations can create strong incentives that affect important choices about the development of domestic institutions, even when this outcome is not purposeful.

Second, we provide an explanation about why externally-financed institutional development is so difficult to achieve where it is most needed (Birdsall 2007; Haggard et al. 2008; Eubank 2012; Andrews 2013; Pritchett et al. 2013). Targets have the ability to shape bureaucratic and organizational actions by focusing effort and resources on certain outcomes, but our findings also demonstrate that they can create perverse incentives and unintended consequences in substantively important areas of public administration (Jacob 2005; Hood 2006, 2012; Seabrooke 2012; Bush 2014; Hoey 2015).

Third, our findings speak to an important debate about the mid- to long-term implications of externally-financed institution building efforts (Andrews 2013; Pritchett et al. 2013; Samuel 2013; Krasner and Risse 2014; Lake and Farris 2014). We demonstrate that aid-dependent countries have strong incentives to select form-based targets of institutional performance that are relatively easy to maintain, but they face substantially weaker incentives to pursue targets that measure the functional capabilities of public sector institutions.

2. Institutional Targets in Development Assistance

Over the last several decades, donor agencies have devoted significant time, energy, and funding to "capacity building" and "institutional strengthening" activities in developing countries (Andrews 2013; Berg 1993; IEG 2008; UNECA 2003; World Bank 1997; World Bank 2003). Yet the best available evidence suggests that institutional development programs for the public sector usually fall short of expectations (Berg 1993; IEG 2008; Levy and Kpundeh 2005; Meagher 2005; van de Walle 2001; World Bank 2010b). Progress is often slow and discontinuous, and many developing countries that received institutional development support in previous decades are still receiving support for the same activities today (Birdsall 2008: 517; Easterly 2008; IEG 2008; World Bank 2010b; Andrews 2013).

The primary explanation put forward for the poor track record of programs that promote institutional development is the preoccupation of donor organizations with "blueprint" and "best practice" reforms—for example, the creation of anti-corruption commissions or streamlined agencies for business registration—that work in some countries but not in others (Booth 2011; Evans 2004; Grindle 2011; Haggard et al. 2008; Rodrik 2000; World Bank 2008a). A World Bank evaluation summing up its global experience with programs for public sector development arrives at the same conclusion: "[t]he Bank's approach was too technocratic; it relied on small groups of interlocutors within core ministries and promoted one-size-fits-all [civil service and administrative] reform blueprints in diverse country settings" (IEG 2008: 2). Political scientists, economists, and organizational sociologists generally agree that a focus on institutional forms rather than institutional function is problematic because it can crowd out tailored, country-specific solutions

to governance problems (Grindle 2004; Pritchett and Woolcock 2004; Rodrik 2007; Arruñada 2007, 2012; Haggard et al. 2008; Swift-Morgan 2014).

Left unchecked, “blueprint” and “best practice” pressures from donors and international organizations can also result in dysfunctional behavior by recipient governments that Pritchett et al. (2013) call isomorphic mimicry. Governments adopt “the camouflage of organisational forms that are deemed successful elsewhere to hide their actual dysfunction” (Pritchett et al. 2013: 1-2). After recipient countries adopt these institutional forms, they and their donor agency counterparts can declare success without improving the functions that public sector institutions are able to perform (Samuel 2013). Given reputational and material incentives to achieve observable and quantitative results, form-based targets of institutional development offer cheap and easy demonstrations of “progress” without requiring costly actions to improve long-run state capacity (Arruñada 2007; Cullen and Randall 2006; Hood 2006; Jacob 2005; Wynia et al. 2000).

The alternative approach to institutional development is context-specific projects that seek to improve institutional function, as measured by the ability of government agencies to solve public problems and improve policy outcomes (Evans 2004; Andrews et al. 2012; Chong et al. 2012; Woolcock 2013). But improving public sector institutions in these ways is substantially more difficult; in many cases, it requires an “incremental search for solutions to problems that local agents care about” (Andrews 2013: 86). Also, insofar as efforts to improve institutional performance disrupt the domestic political status quo, these activities usually require that public officials are authorized to iteratively adapt to local constraints and opportunities (Andrews 2013; Grindle 2004; Rodrik 2007; Pritchett et al. 2012).

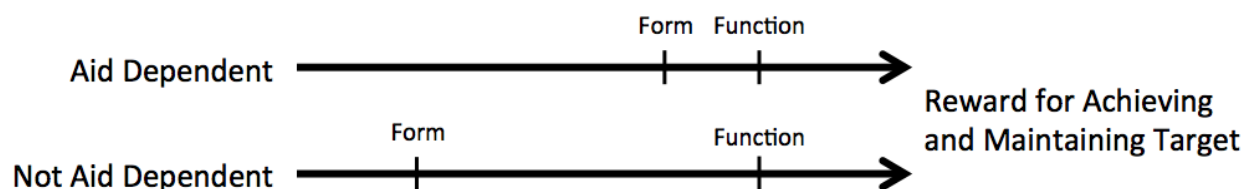
Given these difficulties, we develop and test a theory about when and why recipient governments and donor agencies jointly pursue form-based targets of institutional development despite their limited value. In particular, we test a theory based on differences in rewards for achieving targets in countries that are more or less dependent on foreign aid.

3. The Selection of Institutional Targets is a Strategic Choice

We argue that recipient countries choose targets in order to maximize the payoff that achieving and maintaining these targets will deliver. The expected payoff is a function of the rewards for achieving and maintaining different targets and the difficulty of achieving and maintaining the targets. When different targets offer similarly-sized rewards — in terms of access to aid, signals to investors, and support from domestic constituencies — recipient countries will choose the easier target that requires less costly effort.

We argue that the aid-dependent countries most in need of institutional development receive substantial rewards from donor agencies in the form of continued access to aid for achieving and maintaining form-based targets of institutional development proposed by donors. In contrast, countries that are less dependent on aid do not value funding from donors as highly and often their access to donor funds is not conditional on the achievement of targets. Given that domestic constituencies and international investors reward countries for institutional function rather than new institutional forms, less aid-dependent countries face a larger *relative* difference in the rewards that they can reap by achieving form and function targets.

Figure 1. Relative reward for targets based on country type

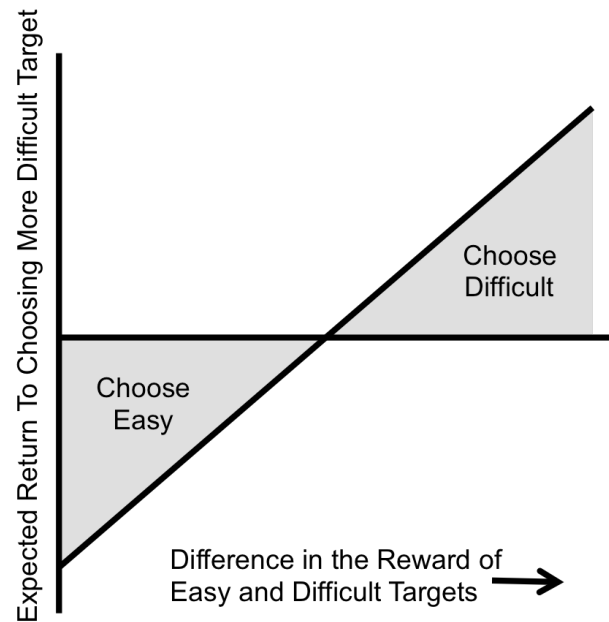


We argue that this difference in relative rewards across different types of recipient countries explains why the countries most in need of institutional development do not choose the more difficult targets that measure improved institutional function. Recipient countries are only likely to choose these more difficult targets (an assumption we verify in robustness checks below) when the rewards for achieving and maintaining more difficult targets are significantly higher, thereby compensating them for the added difficulty and effort required. Aid-dependent countries receive significant rewards for both easy form and more difficult function targets, which often prompts them to choose easier form targets. By contrast, countries that are not aid dependent do not receive rewards for achieving and maintaining easier form targets, so they are more likely to choose function targets despite their added difficulty. Taken together, relative differences in the rewards for achieving targets generate a perverse incentive for the recipient countries most in need of institution building to avoid targets that strengthen public sector institutions.

More formally, consider two hypothetical borrower countries—Country A and Country B—that have secured a certain amount of financing from the World Bank for a project. These countries and their World Bank counterparts must now negotiate the targets to be achieved and later maintained to maximize benefits from domestic constituencies and donor agencies. Country A is more aid dependent and thus receives similar rewards for achieving both form and function targets. Given that the payoffs between these two types of targets are not very different, the political leadership in Country A will choose easier

targets that require less effort. In contrast, Country B receives almost no reward for achieving form targets, either from external donor audiences or from domestic constituencies. This country is more likely to choose function targets than form targets because they offer higher risk-adjusted payoffs. In either scenario, given that function targets are consistently more difficult to achieve, they will only be chosen by countries that receive significantly higher rewards for achieving them (Figure 1).

Figure 2. Choice of targets based on relative rewards of achievement and persistence



3.1 Rewards for Achieving and Maintaining Institutional Form and Function Targets

We derive expectations about the relative rewards of achieving and maintaining different types of targets — and thus their expected payoffs — from primary interviews and previous research about donor allocation practices. In terms of the choice of targets, we first confirmed that recipient governments and project leaders at the World Bank have significant discretion when selecting targets. Our background interviews with World Bank officials suggest that formal organizational rules do not significantly constrain the choice of targets, which are negotiated between teams at the World Bank and staff in line ministries of borrower countries for each individual project. To explain the selection of targets, we must then explain the incentives of officials from recipient governments and World Bank staff who prepare projects for approval.

For the World Bank staff and recipient government officials, the incentives to select different types of targets are similar and reinforcing. Bilateral and multilateral aid agencies provide a wide variety of incentives for recipient governments to improve their performance on function targets (Savedoff 2011; Parks and Rice 2015). The Global Environment Facility, for example, conditions environmental assistance on an indicator that measures “the extent to which [a country’s] environmental policies and institutions foster the protection and sustainable use of natural resources and the management of pollution” (World Bank 2011: 35). Countries can also secure more *discretionary* sources of funding (e.g. budget support and results-based financing) by improving their performance on various measures of institutional function (Clist et al. 2012; Perakis and Savedoff 2015; Parks and Rice 2015).¹

Recipient countries can also garner indirect rewards if they improve institutional function, regardless of their level of aid dependence. Whereas domestic political constituencies in developing countries care little about the adoption of new institutional forms, they reward the achievement of function targets and improved public services (Ferraz and Finan 2008; Banerjee et al. 2011).²

In contrast to function targets, the incentives to select form targets are substantially different among governments with high and low levels of aid dependence. Donors often encourage and incentivize aid-dependent governments with weak institutions to pursue “best practice” institutional forms that are observed in wealthy, industrialized countries (Haggard et al. 2008; Booth 2011; Andrews 2013). New evidence from a survey of nearly 7,000 development policymakers and practitioners suggests that these pressures are less acute in lower-middle income and upper-middle income countries where aid dependence is less severe, than in low-income countries (Parks and Rice 2015).

Countries that rely heavily upon aid face at least three incentives to select form-based targets of institutional development that countries with low levels of aid dependence do not. First, the selection of form-based targets can hasten the approval of new projects from donors (Andrews 2013), which aid-dependent governments often need to stay in power (Morrison 2009; Masaki 2015). Second, when aid-dependent governments agree to include the form-based targets that donors favor in general budget support or sector budget support agreements, they reduce the risk of major funding disruptions that could threaten regime stability (Nielsen et al. 2011; Steinwand 2014).

Third, aid dependent countries can secure more funding from donors that allocate aid with performance-

¹ The amount of discretion that the World Bank yields to the counterpart government is usually tied to the performance of certain functions. In the environment sector, for example, use of these more discretionary financing modalities may depend on whether a government ministry is assessing the environmental impacts of major infrastructure projects or paying forest titleholders for protecting sensitive watersheds (World Bank 2008b, 2010a; World Bank Carbon Finance Unit 2013).

² Additionally, a reputation for sharing the policy preferences of donor agencies and maintaining functional institutions can help the authorities in developing countries to generate more international attention and support (Chwieroth 2013; Flores et al. 2013).

based formulae if they achieve and maintain form-based targets of institutional development. Consider the World Bank, which is the focus of our empirical analysis. The allocation of grants and concessional loans through the World Bank's International Development Association (IDA) — a funding window for poor, aid-dependent countries that are insufficiently creditworthy to borrow on commercial terms — is directly linked to a recipient country's ability to meet project performance targets (ADB 2005; World Bank 2010b), and (as we will soon show) form-based measures of institutional development success figure prominently among these project performance targets.³⁴ Thus, IDA beneficiaries should understand that improvements (declines) on form-based targets of institutional development will result in more (less) funding over time. Many multilateral development banks (the Inter-American Development Bank, the Asian Development Bank, the African Development Bank, and the Caribbean Development Bank) and multilateral institutions (IFAD, the Global Environment Facility, the European Commission) have followed suit and established similar performance-based resource allocation systems that reward countries for achieving better project performance scores (ADB 2005, 2007; European Commission 2008, 2014; GEF 2013a, 2013b). This creates an additional incentive for aid dependent governments to achieve form-based targets of institutional development. In contrast, the less concessional sources of development financing that middle-income countries can access usually do not depend on performance-based allocation formulae. This is true of the World Bank as well as most of the major regional development banks.

The incentives of World Bank staff also reinforce the selection of form targets in aid-dependent countries. The existing literature suggests that World Bank staff is primarily rewarded — in terms of salary, promotion, and internal prestige — for project approval and loan disbursement rather than project quality (World Bank 1992; Easterly 2001; Phillips 2009; IEG 2011). Conditional on securing approval for loans, staff receive credit for designing and operating projects that improve outcomes in recipient countries (Whittle 2013; Malik and Stone 2015). Approving new loans involves inter-departmental and shareholder country review, including monitoring and evaluation specialists who focus on finding feasible targets that project staff can achieve (Cotlear and Kronick 2010). Under pressure to secure new lending, staff are likely to push for feasible targets that are ambitious enough *in a given country context* to pass review. Specifically, based on observed patterns in the achievement and persistence of different types of institutional development goals (reported later in this paper), we expect that World Bank staff will advocate for form targets in low-capacity governments and function targets in high-capacity governments.

³ Knack et al. (2014) demonstrate that crossing the IDA eligibility threshold substantially reduces the amount of aid that a country receives from many donors.

⁴ IDA's performance-based allocation system has used a formula and quantitative measures of need, project performance, and policy and institutional performance to determine the resource envelope that will be made available to countries since the 1970s (World Bank 2010b).

4. Measuring the Choice of Targets for Institutional Development

To test our hypotheses, we examine institutional development targets that were selected for inclusion in World Bank-sponsored environmental management projects. Aid projects that focus on human interactions with the natural environment deal almost exclusively with institution building, making them an ideal set of projects to test our theory. Scholars agree that environmental management is rarely successful without strong formal institutions (Keohane and Levy 1996; VanDeveer and Dabelko 2001; Weidner 2002). However, given that institution building in the environmental sector is often driven by the concerns of donor governments and aid agencies (Keohane and Levy 1996; Lewis 2003; Hicks et al. 2008), it also offers aid-dependent recipient countries strong incentives to signal success to external donor audiences.

For the purposes of our explanatory analysis, we created an original dataset of institutional targets in the environment and natural resource management sector from World Bank post-project evaluations. We identified 250 World Bank projects completed between 2003 and 2009 that allocated more than 10% of their financing to strengthening environmental institutions. We collected Implementation Completion Reports and Independent Evaluation Group evaluations for these projects and extracted targets from these documents that were used to measure institutional development at the completion of the World Bank projects. We identified 826 targets, around 3.3 per project. We assigned all institutional targets to “form,” “action,” or “function” categories according to the following coding rules:

Form: the target records that an institution, law, policy, or regulation exists or is organized in some way. There is no measure of activities or the results of activities. Examples include the establishment of a governmental unit or the passage of a law.

Action: the target records that an agency did something, though the intended results of the activity are not measured. Examples include training a certain number of people or regularly monitoring an environmental attribute.

Function: the target records the results of institutional development for a relevant environmental attribute or policy outcome. Examples include reducing wasted water or the number of days with severe air pollution.

While the primary theoretical focus of this paper centers on the distinction between institutional form and

function targets, development programs often include targets that measure actions (e.g., personnel training). From a theoretical perspective, we do not consider action targets to be substantially different from form targets, since both involve measuring whether an institution or process exists, rather than whether the institution or process successfully achieves some goal. We find that action indicators are achieved and maintained at a similar rate as form targets and thus for our main analysis we combine indicators classified as form and action together and focusing on what causes countries to choose function targets.⁵

Three research assistants independently coded each target. For 72% of targets, all three coders agreed. For the disputed codes, one of the co-authors arbitrated. Either all coders agreed or the arbitrator agreed with the majority of the coders 95% of the time. Of the 826 targets, 380 were coded as measuring function, 285 as measuring form, and 161 as measuring action. A large number of projects contained both form and function targets, which allows us to examine the average relative difficulty of targets within projects. Table 1 contains actual examples of targets to illustrate coding choices.

Table 1. Examples of institutional development indicators from study sample

Country	Completion Year	Project Name	Indicator	Form, Action, or Result?
Argentina	2007	Native Forests & Protected Areas Project	Does Argentina regularly update its national inventory of forest resources?	Action
Ghana	2008	Natural Resource Management Project	Does the Environmental Protection Agency maintain regional offices in all ten regions of Ghana?	Form
Senegal	2009	Long Term Water Sector Project	What percentage of water produced by Senegalaise des Eaux (SDE) is unaccounted for?	Function
India	2007	Industrial Pollution Control Project and Industrial Pollution Prevention Project (PPAR)	Does Maharashtra have a State Pollution Control Board (SPCB) laboratory?	Form
Guinea	2006	Third Water Supply and Sanitation Project	What is the billing/production ratio of the Guinean Water Operation Company (SEG)?	Function

⁵ In the empirical models of target choice reported in Table 4, we dropped all “action” targets from the dataset as a robustness check. In no case did the substantive or statistical significance change.

5. Model and Results: The Political Economy of Institutional Targets

We hypothesize that World Bank staff and recipient governments should favor function targets when the rewards for achieving form targets are low. This occurs most often for countries that are not dependent on flows of aid. In this case, the additional rewards for selecting and achieving function targets are large relative to form targets.

To indicate the relative difference in the reward for form and function targets, we use four different operationalizations that are closely tied to World Bank operational rules.⁶ First and second, we use the project-level and country portfolio-level proportions of funding from the World Bank's International Development Association (IDA) versus the International Bank for Reconstruction and Development (IBRD). Recipient countries gain access to IBRD lending as their income rises and creditworthiness improves. IDA borrowers tend to have less access to commercial bond markets, and they gain more rewards for signaling to donors that they have achieved form-based performance targets, which should make them less likely to choose function targets (Knack et al. 2014). Third, we use GDP per capita in the year prior to project approval, since richer countries are less likely to be dependent on international donors than countries with lower levels of income. Countries with higher per capita income often enjoy a larger and more effective tax base (Knack 2009), which reduces dependence on foreign assistance and diminishes the payoff for signaling to donor agencies. Fourth, we use the country-income classification of the World Bank. Countries that enter middle-income status have more access to non-concessional lending that is not dependent on the achievement of project targets, which should increase the selection of function targets.

To account for access to alternative sources of revenue, we use the share of GDP from natural resource rents for each recipient country as a control variable. Natural resources rents provide governments with an independent revenue stream and might blunt the incentive to signal to donor organizations (Girod and Tobin 2011). We use a measure developed by Hamilton and Clemens (1999) that sums rents from fuel and nonfuel natural resources.

To rule out the possibility that low-income countries that depend on concessional financing simply have fewer institutions in place and need to pursue these targets as precursors to more function, we control for the the establishment of a national environmental ministry (Aklin and Urpelainen 2014) and an omnibus,

⁶ We chose not to use official development assistance as a percentage of gross national income because of very significant missing data in our sample and because this total official development assistance is further removed from more proximate measures of dependence upon concessional finance from the World Bank.

national environmental law (Longhofer and Schofer 2010).⁷ In neither case does the presence and age of formal environmental institutions change our estimates about how aid dependence shapes the choice of targets.⁸

To ensure that our model of the choice of indicators accounts for country-specific factors and the level of implementation, we specify a random-effects model where the intercept varies by country and level of implementation (e.g. municipal, regional, national). We also control for variables that are likely to be independent from country and implementation level random effects. First, in order to account for the growing measurement pressures that the “results agenda” may have brought to bear on the World Bank, we include a count of years since the first year in our sample to account for trends in the selection of indicators. Second, we expect that projects with an explicitly environmental justification will likely induce signaling behavior and the choice of easy targets, since they are often less preferred by borrowers and implemented with less success than projects that are a mix of development and environmental objectives (Buntaine and Parks 2013). Thus, we use an indicator for whether the Environment Sector Board at the World Bank was responsible for reviewing the project.⁹

Our model results support our main hypotheses that poorer, more aid-dependent countries that receive primarily concessional finance from the World Bank are less likely to choose function targets, even after controlling for the presence of major public institutions in the environment sector (Table 2). The results are insensitive to the choice of indicator to measure dependence upon concessional World Bank financing. The results are also insensitive to controlling for the presence of environmental ministries, or alternatively, major environmental legislation. The results do not change when adjusting the time frames that define new or established public institutions in the environment sector. Likewise, the results are insensitive to the removal of “action” targets from the sample.

Across all of our model specifications, we find that countries with more natural resource rents as a proportion of GDP are also more likely to choose function targets. We interpret the model results to mean that countries with more access to alternative sources of non-aid revenue reap fewer rewards for accepting form-based performance targets that primarily signal performance to donor audiences. Our prediction study presented below calls this result into question, however.

⁷ These two indicators are closely correlated. Aklin and Urpelainen (2014: APP-6) note that “[t]he diffusion of environmental ministries follows a curve that is almost identical to environmental framework laws” (Aklin and Urpelainen 2014: APP-6).

⁸ We thank Evan Schofer for sharing an updated version of the environmental legislation data used in Longhofer and Schofer 2010.

⁹ Sector boards at the World Bank are cross-cutting units, typically made up of managers in a specialized field who have responsibilities to ensure that projects are designed and managed according to the prevailing knowledge and practice in a given sector.

Table 2. Choice of function institutional development targets (at project approval)

DV: Function Target	1	2	3	4
Project IDA proportion	-0.78 (0.32) [0.01]			
Portfolio IDA proportion (AY)		-0.67 (0.39) [0.04]		
GDP per capita (\$1k @ AY-1)			0.19 (0.10) [0.02]	
Lower-Middle Income Category				0.77 (0.31) [0.01]
Upper-Middle Income Category				1.13 (0.51) [0.01]
Resource Rents / GDP (% @ AY-1)	0.05 (0.02) [0.01]	0.04 (0.02) [0.02]	0.05 (0.02) [0.01]	0.06 (0.02) [0.01]
New Environmental Ministry (≤ 5 yrs.)	0.14 (0.34)	0.09 (0.34)	0.20 (0.34)	0.02 (0.34)
Established Environmental Ministry (≥ 6 yrs.)	0.10 (0.42)	0.06 (0.43)	0.18 (0.43)	0.10 (0.42)
Approval Year (centered linear)	-0.00 (0.00)	-0.00 (0.04)	-0.01 (0.04)	-0.01 (0.04)
Environment Sector Board	-1.54 (0.33)	-1.44 (0.33)	-1.49 (0.33)	-1.49 (0.33)
<i>Level R.E.</i>	Yes (4)	Yes (4)	Yes (4)	Yes (4)
<i>Country R.E.</i>	Yes (85)	Yes (81)	Yes (85)	Yes (85)
Observations	806	801	806	806

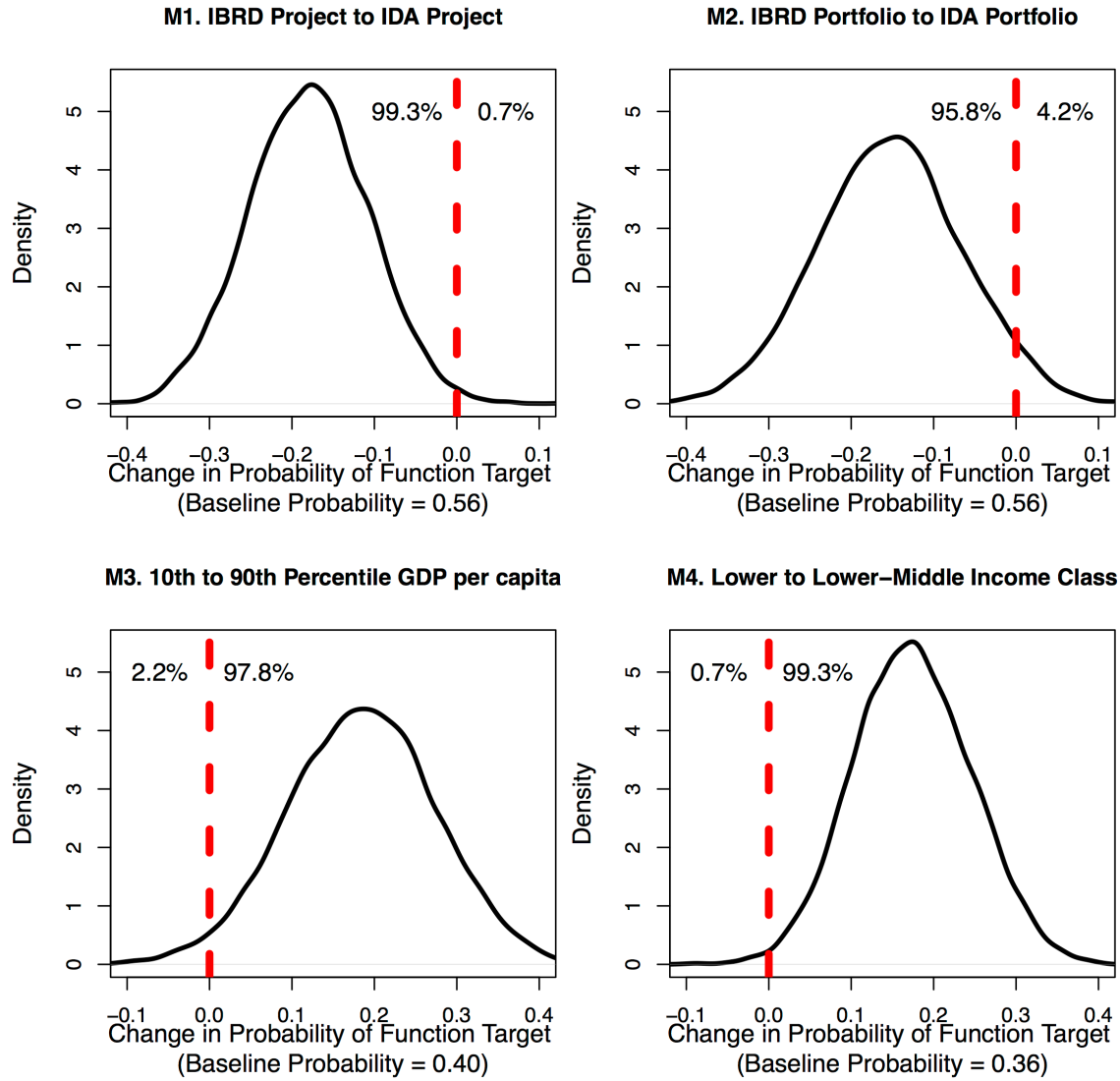
Model cells list: Parameter estimate; (Standard Error); [p-value of one-sided z-test]

All models are random-intercept logit fitted by Laplace approximation with levels as indicated

To aid substantive interpretation of our in-sample explanatory models, Figure 3 contains simulation

results that display changes to the probability of choosing a function target based on our four operationalizations of aid dependence. We randomly draw from all the coefficient distributions in each model to capture total model uncertainty, and then compute pairs of predicted probabilities varying only the main predictor variable. The top row of Figure 3 shows that changing from a non-concessional IBRD project or country-level portfolio to a concessional IDA project or portfolio decreases the probability that a function target will be chosen by approximately 20%. The bottom row of Figure 3 shows that moving from the 10th to the 90th percentile of per capita income in our sample, or moving from a low-income to a lower-middle income World Bank country classification increases the probability that a function target will be chosen by approximately 20%. Each operationalization leads to a finding that countries more dependent on concessional finance from the World Bank are less likely to choose function targets, even after controlling for the level of institutionalization in the environment sector. The countries that choose function targets are least in need of donor-supported institutional development, which helps explain why institutional results have been disappointing for poor aid recipients. These countries favor easy, shallow, form-based targets.

Figure 3. Effect of predictor variables on choosing function target



6. Validating Models with Out-of-Sample Prediction

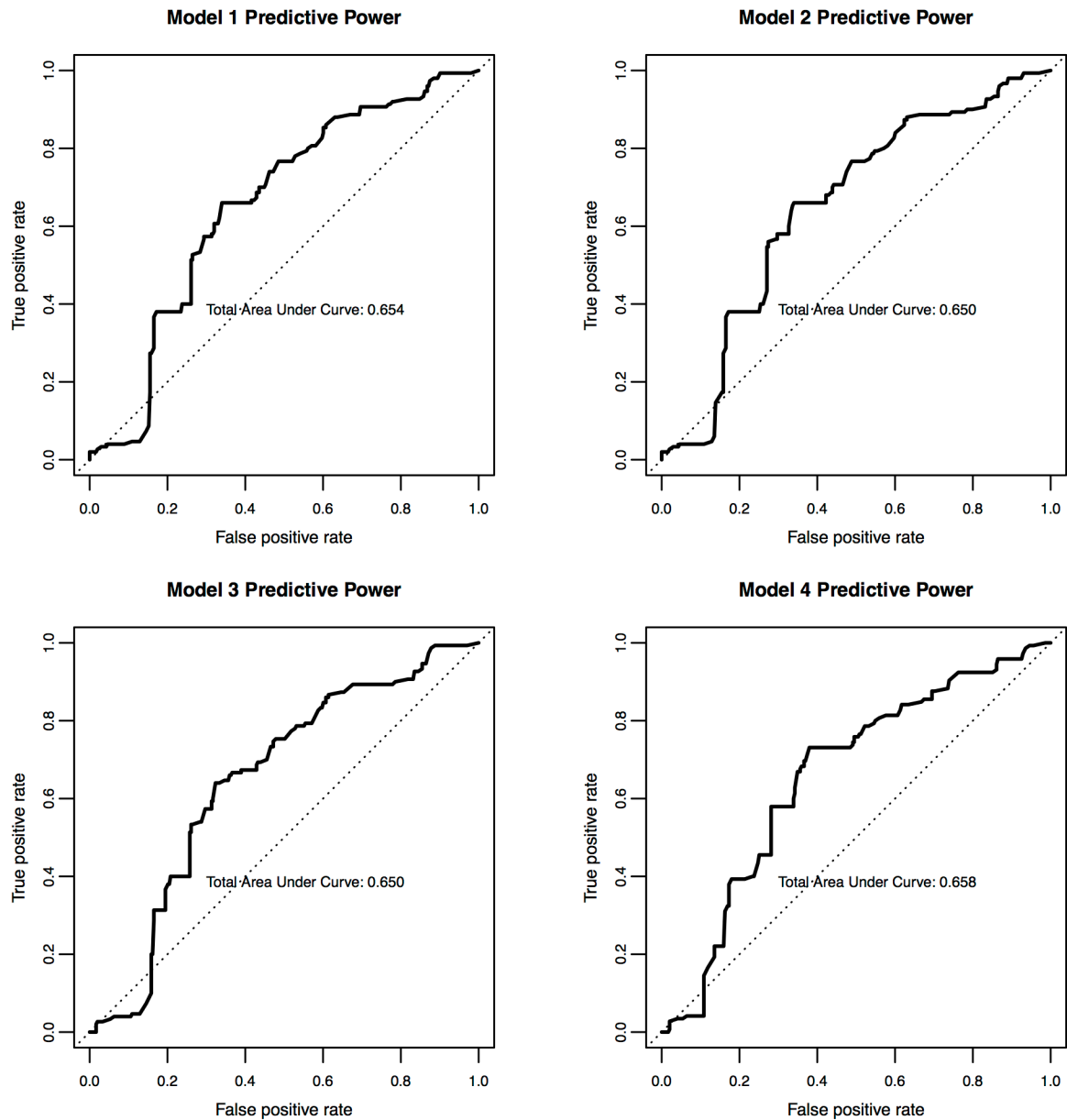
To assess whether our models can predict the selection of targets out-of-sample and thus alleviate concerns that our results hinge on particular modeling assumptions, we collected a new dataset of targets from 79 World Bank projects approved between 2009 and 2011 that focused on environmental institutions. To limit our ability to search through model specifications and report only those specifications that fit our hypotheses, we have not updated the models in the previous section in light of this prediction

study (see Humphreys et al. 2013).¹⁰ In total, we extracted and coded 454 institutional targets from project appraisals using the same procedure reported above. We used each of the four models in Table 4 to predict whether a function target would be chosen in the new sample given the characteristics of the country, project, and level of implementation.

To validate and assess the predictive power of these models, we calculated Receiver Operator Characteristic (ROC) plots (see Ward et al. 2010). Because logistic models produce a probability estimate between 0 and 1 for each observation, a rule to convert these estimates to binary predictions is necessary to assess the predictive power of a model. ROC plots show the percentage of correctly predicted true positives against incorrectly predicted false positives along different threshold values that convert probabilistic predictions into binary predictions. The area under the resulting curve (AUC) will be higher when more true positives than false positives are predicted at each threshold value, indicating correct sorting by the model. A model with no predictive power will produce true and false positives at the same rate at each threshold, resulting in a baseline AUC of 0.5. Predictive results for our four models are displayed in Figure 4. All four models perform better than random guessing. While these graphs give a sense of the predictive power of each model as a whole, they do not show which variables add most to the predictive power.

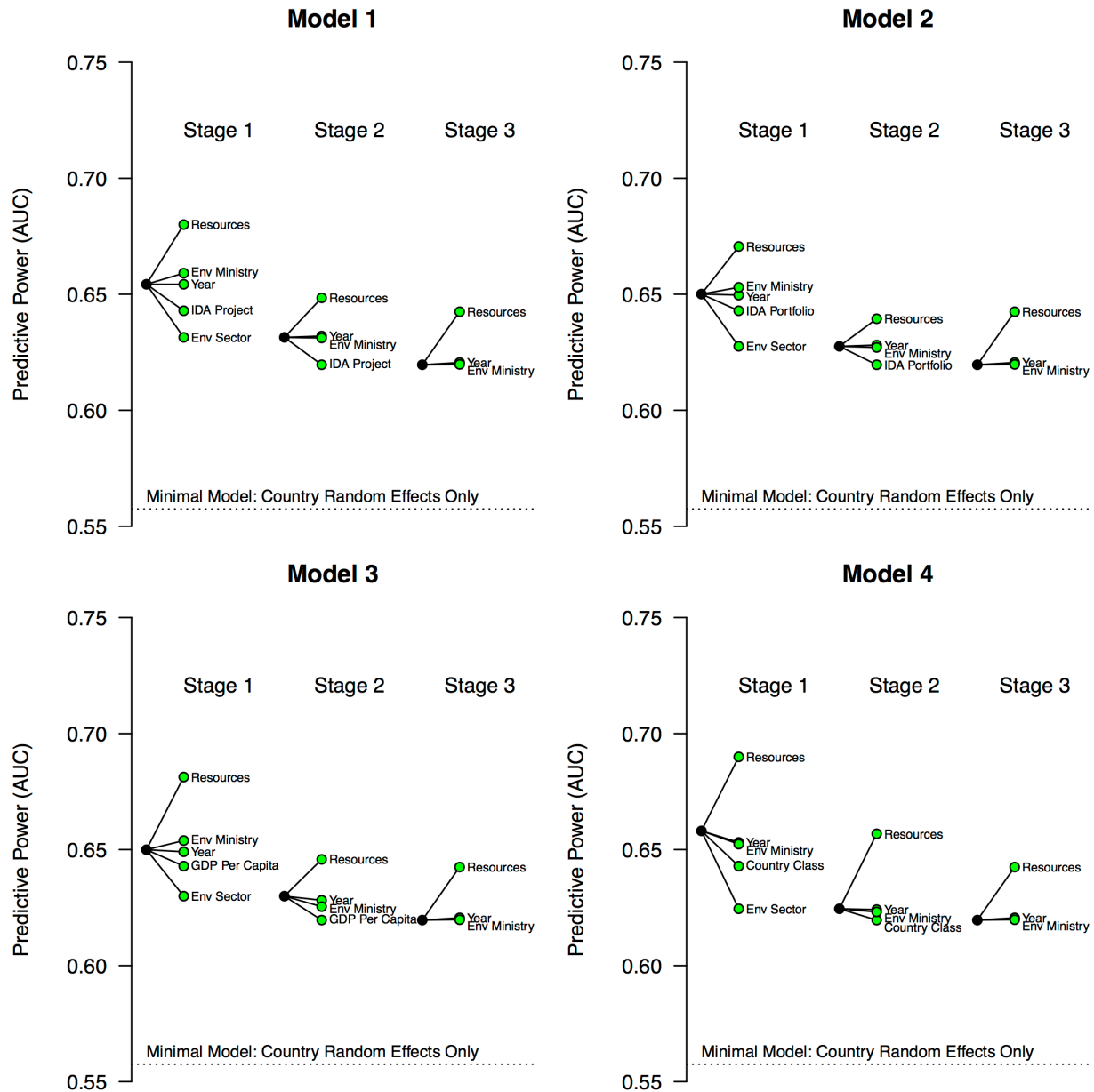
¹⁰ The models presented above do differ slightly from earlier versions of the paper, particularly in terms of the inclusion of control variables for environmental ministries that measure institutionalization in our study sector. We have not updated these models to optimize the prediction results below, though as we explain that could actually be done.

Figure 4. AUC curves for out of sample prediction using the four models



To assess how each independent variable contributed to predictive power, we constructed partial models by removing one predictor variable at a time, fitting the model to the in-sample data reported above, and then predicting the out-of-sample outcome for our second dataset. We then calculated the AUC for each of these partial models and compared it to the AUC of the full model. We then repeated this process for the partial model that removed the most predictive variable, resulting in a new set of partial models. We repeated the process for a third stage (Figure 5).

Figure 5. Representation of the drop in AUC accompanying the removal of each IV



Whether the project was routed through the environment sector board is most predictive of the target that is chosen in all of the models. However, in each of the models the second most predictive set of variables are our measures of aid dependence and reliance on concessional finance. In each model, these variables continue to add predictive power even after the environmental sector board is removed from the model, as displayed in the stage 2 results. No other variables add out-of-sample predictive power. These results add further support to our theory and show that our results generalize across time periods and are

not produced by idiosyncratic modeling choices.

Removing access to natural resource rents from the model fit almost always *improves* predictive power, indicating that while resource rents may have been statistically significant in our initial sample this effect does not hold in our new set of data. Our finding about natural resource rents is likely an artifact of our sample period or model specification reported in Table 4, which illustrates the importance of out-of-sample procedures for model validation. This result also shows the importance of a pre-commitment to out-of-sample testing. If we had updated our in-sample results in light of this prediction study, we would have found that our main variables of theoretical interest add significantly more predictive power in partial models without access to natural resource rents included as a variable, while our main independent variables are both significant in-sample and predictive out-of-sample.

6.1 Testing the “Difficulty of Form and Function Target Achievement” Assumption

A key assumption underpinning our argument is that form targets are easier to achieve and maintain than function targets. We expect that targets that can be achieved merely by an institution or policy existing (e.g., a new law, task force, or training program) are easier to achieve and maintain than targets that require improvement in how well environmental institutions perform their intended goals (e.g. quality of forest management, reduction in pollution, or reduction in water waste). If this is correct, then the ability of recipient governments and World Bank staff to act strategically regarding target selection is strong.

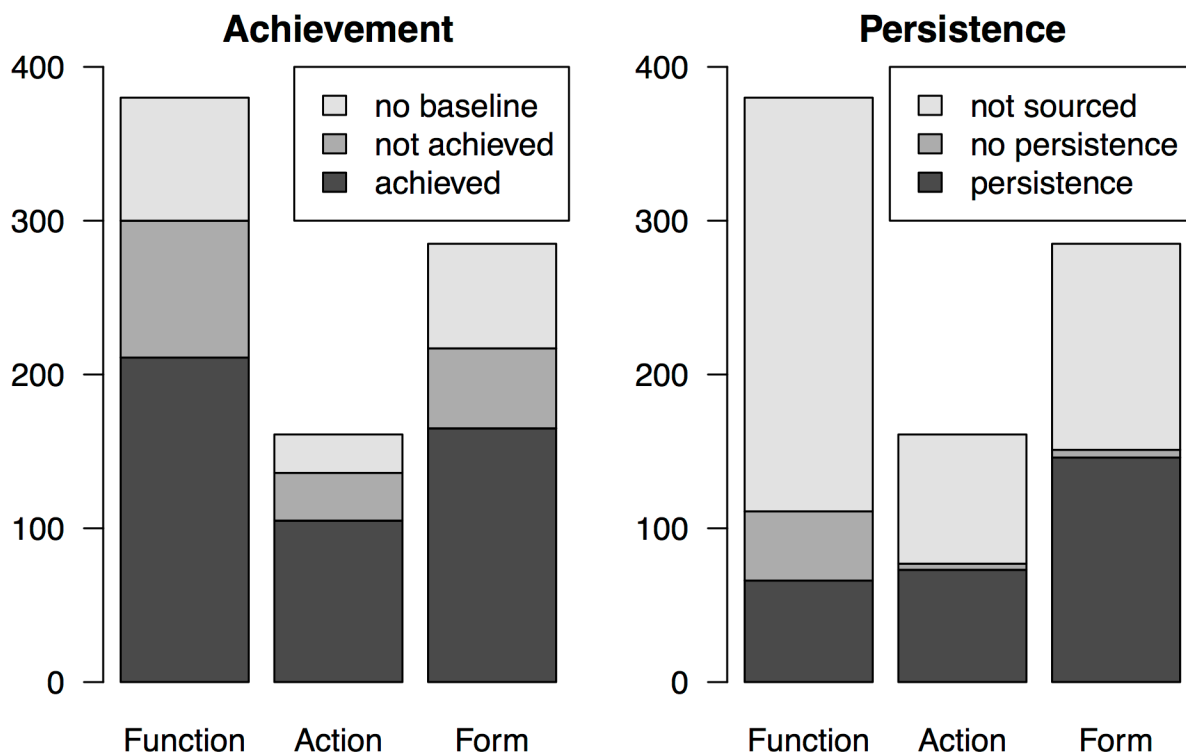
In order to test this assumption, we need additional data on whether the targets selected to measure project success were actually achieved by the projects in our sample and if these achievements persisted after the project was completed. Baseline and completion values of the targets were often included in the post project evaluations that were used to initially code the institutional targets. In cases where both baseline and completion data were available (around 85% of indicators), we computed an ordinal measure of the progress made towards the target during the project.

To monitor the persistence of institutional development gains beyond the close of projects, we compiled new target measures from a comprehensive search of official documents, government websites, NGO reports, and news sources at the beginning of 2013. We used a variety of public news sources, such as LexisNexis and Google News. In addition, many government agencies issue annual reports on staffing and appropriations, allowing us to re-measure some targets. Government websites and NGO reports also offered valuable data. In all cases, we recorded the source and publication date of the re-measured target. In total we re-measured post-completion data for 347 of our indicators, a success rate of around

42%.

An initial examination of the descriptive data seems to demonstrate that form targets are easier to achieve and maintain than function targets (Figure 6). For targets with baseline data in the evaluation documents, 70% of function targets were achieved during project implementation, while 77% of action targets and 76% of form targets were achieved during project implementation. The apparent differences in the difficulty of maintaining targets following project completion is clearer. Whereas only 59% of function targets were maintained following project completion, 95% of action targets and 97% of form targets were maintained following project completion.

Figure 6. Descriptive data on achievement and persistence of different types of targets



If our results about the strategic choice of targets are correct, however, these descriptive data are likely to suffer from country-level selection effects. We find that countries that need to signal to donor audiences are more likely to select form targets. In contrast, countries that have less of a need to signal donor audiences are more likely to select function targets (Table 4). These selection effects might work in two directions. Countries with higher levels of aid dependence might have lower capacity and thus be less likely to achieve and maintain targets of any kind. However, these same countries may have greater incentives to signal success to donors, making them more likely to achieve and maintain targets.

Another potential problem with these descriptive data is omitted variable bias. We were not able to collect data about achievement and persistence for all of the targets in our sample. To interpret regression results about the relative difficulty of targets in light of missing outcome data, we must make the strong assumption that the outcome variable is missing at random (MAR) — that missingness does not depend on the value of the outcome variable conditional on the predictor variables. It may be the case that even conditional on the predictor variables, indicators that are not achieved or do not persist are less likely to be observed.

To address both of these potential problems, we developed a set of models to estimate *within-project* differences in achievement and persistence between form and function targets. If targets offer a strategic choice, then we should observe different rates of achievement and persistence even after we hold countries and projects constant. Fortunately, our sample includes many projects where countries select a mix of form and function targets. By estimating the effect of target type of the rate of achievement and persistence from *within-project* variation, we rule out the possibility that country-level selection effects drive the results.

Controlling for project-level variation in rates of achievement and persistence has the added advantage that we are more likely to satisfy the MAR assumption. To satisfy this assumption, we want to control for all factors that affect missingness. Controlling for project-level variation ensures that our results are not biased by missingness, to the extent that missingness is fully predicted at the project-level and with other included control variables. Even more reassuringly, when only the outcome variable is missing in logistic regression, the parameter estimates (not the intercept) will still be unbiased if the proportional odds assumption holds, even when the MAR assumption does not (Vach 1994). Note that multiple imputation would not help, since multiple imputation analysis and complete case analysis converge when missingness is only in the outcome variable (Snijders and Bosker 1999). By controlling for variation through project-level fixed and random effects, rather than other independent variables that are subject to missingness, we remove bias caused by casewise deletion and decrease uncertainty associated with imputation of missing data.

We specify random-effects and fixed-effects models for both the achievement and persistence of targets to justify our primary theoretical assumption that there are consistent differences in the difficulty of achieving and maintaining these targets. Our random-effects model is a hierarchical logistic regression model with project, country, and implementation level (municipal to national) random effects. The modeling strategy is a conservative approach to omitted variable bias and case selection bias, since we seek to account for all of the time-invariant country effects (e.g., statistical capacity), project-specific

effects (e.g., difficulty of context for project implementation), and implementation level effects (e.g., greater difficulty of supervising implementation by municipal agencies) that are independent from target type.

For both models related to the achievement and persistence of institutional targets, we include one minimally specified model that includes only the random-effects and the type of target (i.e., “form”, “action”, “function”) as the main predictor variable. In our second model for both achievement and persistence, we include an indicator for whether borrower performance was satisfactory during project implementation, as rated by the World Bank Independent Evaluation Group. In our model of persistence, we add a variable for the number of years between completion and our re-measurement data, since targets achieved long ago might be less likely to persist and the gap to re-measurement is not necessarily a project-specific characteristic. In our second model of target persistence, we also include an indicator variable for whether the target was achieved during implementation, since this might correspond to borrower commitment or capacity, but this is not necessarily a project-specific effect.

Our fixed-effects model controls for project-level variation through project indicators. Because fixed-effects models must include variation on other predictor variables within the fixed-effects indicators, these models are restricted to a sample of projects where (1) both form and function targets were chosen; and (2) outcome data on either achievement or persistence exists for at least two targets. Table 2 summarizes the results. The results justify our main theoretical assumption that form and action targets are easier to achieve and maintain than function targets, as estimated from within-project variation.

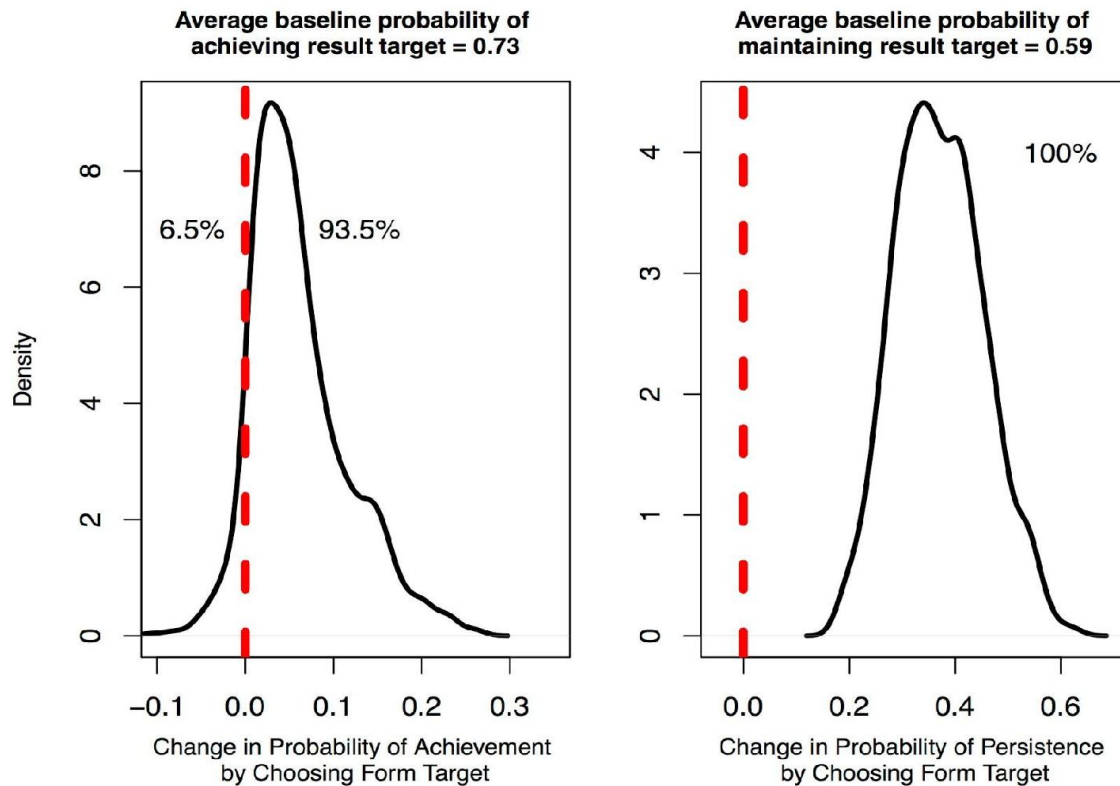
Table 2. Achievement and persistence of institutional development targets

DV:	Achieve	Achieve	Achieve	Persist	Persist	Persist
Form Indicator	0.41 (0.28) [0.07]	0.43 (0.28) [0.06]	1.06 (0.55) [0.03]	3.03 (0.51) [0.00]	3.04 (0.59) [0.00]	3.86 (1.42) [0.00]
Action Indicator	0.35 (0.32) [0.13]	0.40 (0.32) [0.10]	0.89 (0.60) [0.07]	2.55 (0.56) [0.00]	2.60 (0.59) [0.00]	4.18 (1.95) [0.02]
<i>Satisfactory Borrower</i>		0.43 (0.29)			-0.53 (0.50)	
<i>Gap to Remeasure</i>					0.11 (0.08)	
<i>Achievement</i>					-0.04 (0.49)	
Project	R.E. (208)	R.E. (204)	F.E. (81)	R.E. (147)	R.E. (126)	F.E. (39)
Country	R.E. (80)	R.E. (80)		R.E. (73)	R.E. (66)	
Level	R.E. (4)	R.E. (4)		R.E. (4)	R.E. (4)	
Observations	653	645	345	339	281	132
Deviance Reduction w/ Estimated Parameters	0.01	0.01	0.02	0.18	0.18	0.26

Model cells list: Parameter estimate; (Standard Error); [p-value of one-sided z-test]

To aid substantive interpretation of these models, we display first difference simulations for the increase in probability of achieving and then maintaining a form target versus a function target for the hypothetical average project and country based on our models (Figure 7). We do not find implementation level effects in our models. We take 1000 draws from the distributions of model coefficients, including average random effect levels for project and country, and then vary only whether the hypothetical target measured a function or a form. These simulations show that on average choosing a form target for a municipal project increases the probability of achievement from 73% to 79% and increases the probability of maintaining the target from 59% to 96%. Thus, the choice of targets, even within the confines of specific projects that are more or less challenging to implement, might serve as an important strategic choice for both recipient governments and World Bank staff.

Figure 7. The effect of target choice on achievement and persistence of institutional developments



7. Conclusion

Previous research suggests that a focus on form rather than on function is at least partially responsible for the mixed success of donor-financed projects that promote institutional development. The temptation to rely on form-based measures of success is strong for both donor agency staff and developing country officials because form targets are significantly easier to achieve and maintain than indicators that measure institutional function.

In this paper we take this argument further, outlining a theory of target selection with explicit predictions about when form-based targets are most likely to be selected based on a strategic interaction between external donors and developing countries. In particular, we expect that countries dependent on the World Bank for concessional financing are the most likely to choose form targets. The results from our in-sample explanatory and out-of-sample predictive models provide strong support for this conclusion. We find that countries with lower levels of reliance on concessional development finance are more likely to choose

function targets. Overall, these findings suggest that the selection of targets is a strategic decision made by donor agency staff and its developing country counterparts. This strategic logic leads the recipient countries most in need of institutional development to select easy form targets that can crowd out deeper, function-based reforms.

These findings call into question the conventional wisdom that building stronger monitoring procedures for aid projects will necessarily improve the impact of development assistance. While project success rates at the World Bank have apparently increased since the 1980s (Sud and Olmstead-Rumsey 2012), our findings suggest that it is not enough to measure performance vis-à-vis targets; monitoring and evaluation should place greater emphasis on measuring de facto institutional function.¹¹ To be clear, we are not arguing that the “results agenda” being promoted and pursued by aid agencies and development banks is fundamentally misguided. The need for more effective measurement of development results is both compelling and overdue. However, the ability to choose different types of performance targets, even within organizations that prioritize accountability and measurable results, can set in motion the strategic logic between developing countries and development organizations as described in this article.

This article also calls to attention the fact that donor agencies and development banks are urgently in need of better methods to track the persistence of institutional improvements supported by their projects and programs. The data necessary to examine how specific donor-supported institutions fare in the medium- to long-term simply do not exist in most cases.¹² As such, donor organizations trying to evaluate post-project impacts are forced to guess about the conditions that support sustainability after project completion.

We have attempted to overcome this knowledge barrier by (a) introducing a new method for evaluating the long-run institutional development effects of aid projects; and (b) updating hundreds of indicators of institutional development that were previously measured in ex-post evaluations of donor projects. We hope that this methodological contribution will stir discussion and debate about how success is defined and how the long-run institutional developments of externally financed projects and programs should be monitored and evaluated. The absence of good project- and program-specific institutional development data has also limited what researchers can say about how donor agencies, international organizations, and other external actors can most effectively support developing countries in their efforts to build functional government agencies. The fact that scholars analyze broad indices of institutional quality

¹¹ Notwithstanding recent progress at the World Bank to produce core indicators that more effectively measure project success, many of these indicators still measure institutional forms and actions that are several steps removed from improved institutional function.

¹² In an initial search of post-project evaluations completed by all OECD Development Assistance Committee members, we found that only the Japan International Cooperation Agency (JICA) has a program in place to monitor the impacts of its projects following completion.

instead of the project-level institutional development outcomes targeted by donors limits the way research can inform decisions (Knack et al. 2003; Chauvet and Collier 2008).

Three additional policy implications logically follow from our findings. First, in the same way that donors exercise due diligence during project appraisal to address social, environmental, and fiduciary risks, our findings suggest that donors need to put in place measures that detect, combat, and deter isomorphic mimicry and opportunistic behavior geared towards maintaining stable access to aid. They also need to confront internal organizational incentives for staff to collude with their government counterparts in developing countries.

Second, given that the "results agenda" has been taken up by a large number of bilateral and multilateral donors, the efforts of any one donor to confront this issue will likely be inadequate. The prospect of "forum shopping" points to the need for donors to coordinate on setting targets that provide appropriate incentives (Bourguignon and Platteau forthcoming).

Third, pay-for-performance aid delivery schemes may very well provide decision-makers in developing countries with the policy autonomy and maneuverability that they need to "crawl the design space" (Pritchett al. 2012) in pursuit of difficult-to-identify solutions that are fit for local purpose (Natsios 2010; Sjöstedt 2013; Perakis and Savedoff 2015). However, the success of such arrangements hinges critically on the availability of sound performance measures. Institutional development is essential for a strong, endogenously functional state that is capable of providing public services without continued external support, but we may not see substantial improvements unless or until better measures of institutional function are developed.

Our results also have important implications for research in international relations, more generally. We have shown that international development organizations can create important incentives that steer the development of domestic institutions, even when this occurs inadvertently through the specification of performance targets. In attempting to support institutional development, international actors must carefully consider the strategic logic they set in place through contracting with developing country governments (Marcoux and Urpelainen 2012; Helmke and McLean 2014; Bayer et al. forthcoming; McLean forthcoming). As has been shown at the domestic level with regard to legislative targets and public administration (e.g., Boyne and Chen 2007; Keldman and Friedman 2009), our results highlight how external actors can likewise focus attention and effort on certain types of reforms for the countries most dependent on concessional finance.

As the set of core indicators used to measure the success of development projects is codified and

expanded, indicators that measure institutional function, but leave open multiple pathways to this outcome, should be prioritized. Additionally, donors should generally favor customized indicators of institutional function over those that can be standardized across projects and countries and are more likely to measure institutional form or action. If aid agencies and development banks are serious about building strong and effective institutions in the countries where they work, they ought to select targets that more effectively measure whether public sector institutions are capable of discharging their core functions.

References

- Acemoglu, Daron and James Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York: Crown Business.
- Aklin, Michaël and Johannes Urpelainen. 2014. The Global Spread of Environmental Ministries: Domestic–International Interactions. *International Studies Quarterly* 58: 764–780.
- Andrews, Matt. 2013. *The Limits of Institutional Reform in Development*. Cambridge: Cambridge University Press.
- Andrews, Matt, Lant Pritchett, and Michael Woolcock. 2012b. “Escaping Capability Traps through Problem-Driven Iterative Adaptation (PDIA).” UNU-WIDER Working Paper.
- Arruñada, Benito. 2012. *Institutional Foundations of Impersonal Exchange: The Theory and Policy of Contractual Registries*. Chicago, IL: University of Chicago Press.
- Arruñada, Benito, 2007. “Pitfalls to avoid when measuring institutions: Is *Doing Business* damaging business?” *Journal of Comparative Economics* 35(4): 729-747.
- Asian Development Bank (ADB). 2007. Policy-Based Lending: Emerging Practices in Supporting Reforms in Developing Member Countries. Special Evaluation Study Update. August 2007. Manila: ADB.
- Asian Development Bank (ADB). 2005. Multilateral Development Bank Technical Meeting on Performance Based Allocation Methods. Meeting at ADB Headquarters on 24-25 January 2005. Manila: ADB.
- Banerjee, Abhijit V., Selvan Kumar, Rohini Pande, and Felix Su. 2011. “Do Informed Voters Make Better Choices? Experimental Evidence from Urban India.” Working Paper.
- Bates, Robert. 2009. *Prosperity & Violence: The Political Economy of Development*. 2nd Ed. New York, NY: W. W. Norton & Company.
- Bayer, Patrick, Christopher Marcoux, and Johannes Urpelainen. Forthcoming. When International Organizations Bargain: Evidence from the Global Environment Facility. *Journal of Conflict Resolution*.
- Berg, Elliot. 1993. *Rethinking Technical Cooperation: Reforms for Capacity Building in Africa*. New York, NY: UNDP.
- Besley, Timothy, and Torsten Persson. 2010. “State Capacity, Conflict, and Development.” *Econometrica* 78(1): 1-34.
- Birdsall, Nancy. 2008. “Seven Deadly Sins: Reflections on Donor Failings.” In *Reinventing Foreign Aid*, ed. William Easterly. Cambridge, Mass: The MIT Press, 515-553.
- Booth, David. 2011. “Aid, institutions and governance: what have we learned?” *Development Policy Review* 29(s1), s5-s26.
- Boyne, George A., and Alex A. Chen. 2007. Performance targets and public service improvement.

- Journal of Public Administration Research and Theory*, 17(3), 455-477.
- Bourguignon, François and Jean-Philippe Platteau. Forthcoming. "The Hard Challenge of Aid Coordination." *World Development*.
- Buntaine, Mark T., and Bradley C. Parks. 2013. "When Do Environmentally-Focused Assistance Projects Achieve their Objectives? Evidence from World Bank Post-Project Evaluations." *Global Environmental Politics* 13(2): 65-88.
- Bush, Sarah. 2015. *The Taming of Democracy Assistance: Why Democracy Promotion Does Not Confront Dictators*. Cambridge: Cambridge University Press.
- Callaghy, Thomas M. 1988. "The State and the development of capitalism in Africa: Theoretical, historical and comparative reflections." In *The Precarious Balance. State and Society in Africa*, eds. Donald Rothchild, and Naomi Chazan. Westview Press, 67-99.
- Chauvet, Lisa and Paul Collier. 2008. "Aid and reform in failing states." *Asian Pacific Economic Literature* 22(1): 15-24.
- Chong, Alberto., Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2012. "Letter Grading Government Efficiency." NBER Working Paper No. 18268. Cambridge, MA: NBER.
- Chwieroth, Jeffrey M. 2013. "The silent revolution: How the staff exercise informal governance over IMF lending." *The Review of International Organizations* 8 (2): 265-290.
- Clemens, Michael, Charles Kenny, and Todd Moss. 2007. "The Trouble with the MDGs: Confronting Expectations of Aid and Development Success." *World Development* 35 (5): 735-751.
- Clist, Paul, Alessia Isopi, and Oliver Morrissey. 2012. "Selectivity on Aid Modality: Determinants of Budget Support from Multilateral Donors." *The Review of International Organizations* 7 (3): 267-284.
- Cotlear, Daniel and Dorothy Kronick. 2010. "Monitoring Monitoring: Assessing Results Measurement at the World Bank." HNP Discussion Paper. Washington, D.C.: World Bank.
- Cullen, Julie B. and Reback, Randall. 2006. "Tinkering Toward Accolades: School Gaming under a Performance Accountability System." in *Improving School Accountability: Check-Ups or Choice, Advances in Applied Microeconomics*, eds. T. Gronberg and D. Jansen. Amsterdam: Elsevier Science, 1-35.
- Easterly, William. 2008. "Can the West Save Africa?" *Journal of Economic Literature* 47 (2): 373-447.
- Easterly, William. 2001. *The Elusive Quest for Growth: Economists Adventures and Misadventures in the Tropics*. Cambridge, MA: MIT Press.
- Ebrahim, Alnoor. 2013. "Let's Be Realistic About Measuring Impact." *Harvard Business Review* (blog). March 13, 2013. Available at: <http://blogs.hbr.org/hbsfaculty/2013/03/lets-be-realistic-about-measur.html>.
- Esty, Daniel and Michael Porter. 2005. "National Environmental Performance: An Empirical Analysis of Policy Results and Determinants." *Environment and Development Economics* 10 (4): 391-434.
- Eubank, Nicholas. 2012. "Taxation, Political Accountability and Foreign Aid: Lessons from Somaliland."

Journal of Development Studies 48 (4): 465-480.

- European Commission. 2014. MDG Contract. Accessed on 12 January 2014 at http://ec.europa.eu/europeaid/what/millennium-development-goals/contract_mdg_en.htm
- European Commission. 2008. Ghana MDG-Contract (MDG-C). Brussels: European Commission. Accessed on 12 January 2014 at http://ec.europa.eu/europeaid/documents/aap/2008/af_aap_2008_gha.pdf
- Evans, Peter. 2004. "Development as Institutional Change: The Pitfalls of Monocropping and the Potentials of Deliberation." *Studies in Comparative International Development* 38 (4): 30-52.
- Ferraz, C., and F. Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics* 123(2): 703–745.
- Flores, Thomas Edward, Gabriella Lloyd, and Irfan Nooruddin. 2013. "The Technocratic Advantage? Leadership and IMF Programs." Prepared for presentation at the 2013 Annual Convention of the International Studies Association, San Francisco, CA.
- Global Environment Facility (GEF). 2013a. "Mid-Term Review of the GEF System for the Transparent Allocation of Resources (STAR). Technical Document # 1: Comparison of the GEF STAR with other Performance-Based Allocation Systems." Washington D.C.: GEF.
- Global Environment Facility (GEF). 2013b. "System for Transparent Allocation of Resources. POLICY: PL/RA/01." Washington D.C.: Global Environment Facility.
- Girod, Desha M. and Jennifer Tobin. 2011. "Take the Money and Run: The Determinants of Compliance with Aid Agreements." Paper presented at the American Political Science Association Annual Meeting. Seattle, Washington.
- Grindle, Merilee. 2011. "Governance Reform: The New Analytics of Next Steps, Commentary" in *Governance: An International Journal of Policy, Administration, and Institutions* 24 (3): 415– 418.
- Grindle, Merilee. 2004. "Good Enough Governance: Poverty Reduction And Reform In Developing Countries." *Governance: An International Journal Of Policy, Administration, And Institution* 17 (4): 525-548.
- Haggard, Stephan. 1990. *Pathways from the periphery: The politics of growth in the newly industrializing countries*. Ithaca, NY: Cornell University Press.
- Haggard, Stephan, Andrew MacIntyre and Lydia Tiede. 2008. "The Rule of Law and Economic Development." *Annual Review of Political Science* 11: 205-34.
- Hamilton, Kirk and Michael Clemens. 1999. "Genuine savings rates in developing countries." *World Bank Economic Review* 13:333–356.
- Helmke, Gretchen and Elena V. McLean. 2014. Inducing Independence: A strategic model of World Bank assistance and legal reform. *Conflict Management and Peace Science* 31 (4): 383-405.
- Herbst, Jeffrey. 1990. "The Structural Adjustment of Politics in Africa." *World Development* 8 (7): 949-958.

- Hicks, Robert, Bradley Parks, Timmons Roberts and Michael Tierney. 2008. *Greening Aid? Understanding the Environmental Impact of Development Assistance*. Oxford: Oxford University Press.
- Hoey, Lesli. 2015. "Show me the Numbers": Examining the Dynamics Between Evaluation and Government Performance in Developing Countries. *World Development* 70: 1–12.
- Hood, Christopher. 2006. "Gaming in Targetworld: The Targets Approach to Managing British Public Services." *Public Administration Review* 66 (4): 515-521.
- Hood, Christopher. 2012. "Public management by numbers as a performance-enhancing drug: two hypotheses." *Public Administration Review* 72(1): S85–S92.
- Humphreys, Macartan, Raul Sanchez de la Sierra, and Peter Van der Windt. (2013). "Fishing, commitment, and communication: A proposal for comprehensive nonbinding research registration." *Political Analysis* 21 (1): 1-20.
- Independent Evaluation Group (IEG). 2011. "The Matrix System at Work: An Evaluation of the World Bank's Organizational Effectiveness." Washington D.C.: World Bank.
- Independent Evaluation Group (IEG). 2008. "Public Sector Reform: What Works and Why? An IEG Evaluation of World Bank Support." Washington D.C.: World Bank.
- Jacob, Brian A. 2005. "Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools." *Journal of Public Economics* 89 (5): 761-796 .
- Kelman, Steven, and John N. Friedman. 2009. Performance improvement and performance dysfunction: an empirical examination of distortionary impacts of the emergency room wait-time target in the English National Health Service. *Journal of Public Administration Research and Theory*, 19 (4): 917-946.
- Keohane, Robert and Marc Levy. 1996. *Institutions for Environmental Aid: Pitfalls and Promise*. Cambridge, MA: MIT Press.
- Khagram, Sanjeev, Craig Thomas, Catrina Lucero, and Subarna Mathes. 2009. "Evidence for Development Effectiveness." *Journal of Development Effectiveness* 1 (3): 247-70.
- Knack, Stephen. 2001. "Aid Dependence and the Quality of Governance: A Cross-Country Empirical Analysis." *Southern Economic Journal* 68 (2) 310-329.
- Knack, Stephen, Mark Kugler, and Nick Manning. 2003. Second-generation governance indicators. *International Review of Administrative Sciences*, 69(3), 345-364.
- Knack, Stephen. 2009. "Sovereign rents and quality of tax policy and administration." *Journal of Comparative Economics* 37(3), 359-371.
- Knack, Stephen, Lixin Colin Xu, and Ben Zou. 2014. Interactions Among Donors' Aid Allocations Evidence from an Exogenous World Bank Income Threshold. Policy Research Working Paper 7039. Washington DC: World Bank
- Krasner, Stephen D. and Thomas Risse. 2014. "External Actors, State-Building, and Service Provision in

- Areas of Limited Statehood: Introduction." *Governance* 27 (4): 545–567.
- Lake, David, and Christopher J. Farris. 2014. "Why International Trusteeship Fails: The Politics of External Authority in Areas of Limited Statehood." *Governance*. 27 (4): 569–587.
- Levy, Brian and Sahr Kpundeh, (eds.). 2005. *State Capacity in Africa: New Approaches, Emerging Lessons*. Washington, DC: World Bank Institute.
- Lewis, Tammy L. 2003. "Environmental Aid: Driven by Recipient Need or Donor Interests?" *Social Science Quarterly* 84 (1): 144-161.
- Longhofer, Wesley, and Evan Schofer. 2010. "National and global origins of environmental association." *American Sociological Review* 75(4), 505-533.
- Malik, Rabia and Randall Stone. 2015. Private Politics in World Bank Lending. Presented at the 8th Annual Conference on The Political Economy of International Organizations. Berlin, Germany. February 12-14, 2015.
- Marcoux, Christopher, and Johannes Urpelainen. 2012. Capacity, not constraints: A theory of North-South Regulatory Cooperation. *Review of International Organizations* 7 (4): 399-424.
- Masaki, Takaaki. 2015. The Political Economy of Aid Allocation in Africa: Evidence from Zambia. AidData Working Paper 5. Williamsburg, VA: AidData.
- McLean, Elena. forthcoming. A Strategic Theory of International Environmental Assistance. *Journal of Theoretical Politics*.
- Meagher, Patrick. 2005. "Anti-corruption agencies: Rhetoric versus reality." *The Journal of Policy Reform* 8 (1): 69-103.
- Morrison, Kevin M. 2009. "Oil, Nontax Revenue, and the Redistributive Foundations of Regime Stability." *International Organization* 63 (1): 107-138.
- Natsios, Andrew. 2010. "The Clash of the Counter-Bureaucracy and Development." Center for Global Development Essay. July 2010. Washington DC: Center for Global Development.
- Nielsen, Richard A., Michael G. Findley, Zachary S. Davis, Tara Candland, and Daniel L. Nielson. 2011. "Foreign Aid Shocks as a Cause of Violent Armed Conflict." *American Journal of Political Science* 55:219–232.
- North, Douglass C., John Joseph Wallis, Steven B. Webb, and Barry R. Weingast. 2007. "Limited Access Orders in the Developing World: A New Approach to the Problems of Development." World Bank Policy Research Working Paper 4359. Washington D.C.: World Bank.
- Parks, Bradley C. and Zachary J. Rice. 2015. "The Marketplace of Ideas for Policy Change: Who Do Developing World Leaders Listen To And Why?" Williamsburg, VA: The College of William and Mary.
- Perakis, Rita and William Savedoff. 2015. "Does Results-Based Aid Change Anything? Pecuniary Interests, Attention, Accountability and Discretion in Four Case Studies." CGD Policy Paper 053. Washington DC: Center for Global Development.

- Phillips, David A. 2009. *Reforming the World Bank: Twenty years of trial-and error*. Cambridge University Press.
- Pritchett, Lant, and Michael Woolcock. 2004. "Solutions When the Solution is the Problem: Arraying the Disarray in Development." *World Development* 32 (2): 191-212.
- Pritchett, Lant, Salimah Samji and Jeffrey Hammer. 2012. "It's All About MeE: Using Structured Experiential Learning ('e') to Crawl the Design Space." UNU-WIDER Working Paper 012/104. Helsinki: UNU-WIDER.
- Pritchett, Lant, Michael Woolcock, and Matt Andrews. 2013. "Looking like a state: Techniques of persistent failure in state capability for implementation." *The Journal of Development Studies*, 49 (1), 1-18.
- Ramalingam, Ben. 2013. *Aid on the Edge of Chaos*. Oxford, UK: Oxford University Press.
- Rodrik, Dani. 2007. *One Economics, Many Recipes: Globalization, Institutions and Economic Growth*. Princeton, NJ: Princeton University Press.
- Rodrik, Dani. 2000. "Institutions for High-Quality Growth: What They Are and How to Acquire Them." *Studies in Comparative International Development* 35 (3): 3-31.
- Samuel, Boris. 2013, "Economic Calculations, Instability, and (In)formalization of the State in Mauritania, 2003-2011." Presented at the 2013 Conference on African Economic Development: Measuring Success And Failure.
- Savedoff, William, D. 2011. "Incentive Proliferation? Making Sense of a New Wave of Development Programs." Center for Global Development Essay. Washington D.C.: Center for Global Development.
- Seabrooke, Leonard. 2012. "Pragmatic numbers: the IMF, financial reform, and policy learning in least likely environments." *Journal of International Relations and Development* 15(4): 486-505.
- Sjöstedt, Martin. 2013. "Aid Effectiveness and the Paris Declaration: A mismatch between ownership and results-based management?" *Public administration and development* 33(2), 143-155.
- Snijders, Tom, and Roel Bosker. 1999. "Multilevel analysis: An introduction to basic and advanced multilevel modeling." London, UK: Sage Publishing.
- Steinwand, Martin. 2014. "Foreign Aid and Political Stability." *Conflict Management and Peace Science*, doi: 10.1177/0738894214541227.
- Sud, Inder and Jane Olmstead-Rumsey. 2012. "Development Outcomes of World Bank Projects: Real or Illusory Improvements?" Institute for Global and International Studies Working Paper. The George Washington University.
- Swift-Morgan, Jennifer. 2014. Confronting the Informal in Good Governance: The Case of Decentralised Education-System Reform in Guinea. *Development Policy Review* 32 (4): 399–425.
- UNECA. 2003. *Public Sector Management Reforms in Africa: Lesson Learned*. Addis Ababa: United Nations Economic Commission for Africa.

- Vach, Werner. 1994. *Logistic regression with missing values in the covariates*. New York: Springer-Verlag.
- van de Walle, Nicolas. 2001. *African Economies and the Politics of Permanent Crisis, 1979- 1999*. Cambridge, UK: Cambridge University Press.
- VanDeveer, Stacy D., and Geoffrey D. Dabelko. 2001. "It's capacity, stupid: International assistance and national implementation." *Global Environmental Politics*, 1 (2): 18-29.
- Ward, Michael D., Brian D. Greenhill, and Kristin M. Bakke. 2010. "The perils of policy by p-value: Predicting civil conflicts." *Journal of Peace Research* 47 (4): 363-375.
- Weidner, Helmut. 2002. "Capacity Building for Ecological Modernization Lessons from Cross-National Research." *American Behavioral Scientist* 45 (9): 1340-1368.
- Whittle, Dennis. 2013. "How Feedback Loops Can Improve Aid (and Maybe Governance)." Center for Global Development Essay. August 2013. Washington D.C.: Center for Global Development.
- Woolcock, Michael. 2013. "Using Case Studies to Explore the External Validity of 'Complex' Development Interventions." *Evaluation* 19 (3): 229-248.
- World Bank. 2011. The World Bank Group CPIA 2011 Criteria. Washington DC: World Bank.
<http://www.worldbank.org/ida/papers/CPIACriteria2011final.pdf>
- World Bank. 2010a. "Peru - Third Programmatic Environmental Development Policy Loan Program." Washington, DC: World Bank.
<http://documents.worldbank.org/curated/en/2010/07/12540891/peru-third-programmatic-environmental-development-policy-loan-program>
- World Bank. 2010b. "IDA's Performance Based Allocation System: Review of the Current System and Key Issues for IDA16." Washington DC: World Bank.
<http://siteresources.worldbank.org/IDA/Resources/Seminar%20PDFs/73449-1271341193277/PBAIDA16.pdf>
- World Bank. 2008a. "A Note on the Statistical Capacity Indicator." Washington DC: World Bank.
http://siteresources.worldbank.org/EXTWBDEBTSTA/Resources/3561369-1255619840053/Note_on_Statistical_Capacity_Indicator_2009_BBSC.pdf
- World Bank. 2008b. "Mexico - Environmental Sustainability Development Policy Loan." Washington, DC: World Bank. <http://documents.worldbank.org/curated/en/2008/09/9846600/mexico-environmental-sustainability-development-policy-loan>
- World Bank. 2003. *World Development Report 2004: Making Services Work for Poor People*. Oxford, UK: Oxford University Press.
- World Bank. 1997. *World Development Report 1997: The State in a Changing World*. Washington, D.C.: World Bank.
- World Bank. 1992. "Effective Implementation: Key to Development Impact." [Report of the Portfolio Management Task Force, chaired by Willi Wapenhans], Working Paper R92-195, Washington,

November 3.

World Bank Carbon Finance Unit. 2013. "Results-based finance for REDD-plus: Lessons learned from the Forest Carbon Partnership Facility and the Bio Carbon Fund." Washington DC: World Bank Carbon Finance Unit.

Wynia, Matthew K., Deborah S. Cummins, Jonathan B. VanGeest, and Ira B. Wilson. 2000. "Physician manipulation of reimbursement rules for patients: between a rock and a hard place." *Journal of the American Medical Association* 283 (14): 1858–1865.