

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

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

---

## **Abstract:**

Aid agencies spend an extraordinary amount of time, money, and effort trying to improve public sector institutions in developing countries. We investigate why the record of promoting institutional development has been broadly unsuccessful, even as aid agencies focus attention on achieving observable targets. We show that it is easier to achieve and maintain targets that measure the form of institutions, in contrast with targets that measure the functionality of institutions. While recipient countries often receive additional rewards for achieving function targets, only the countries that can achieve these targets with a high probability are likely to choose them. Recipient countries most in need of institutional development focus on form targets that are less likely to improve institutional function, because they seek to maximize the risk-adjusted reward for achieving targets. 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 ..... 5
- 3. Difficulty of Institutional Development Targets ..... 7
- 4. Rewards for Achieving and Maintaining Institutional Function Targets..... 8
- 5. Theoretical Model: the Choice of Function of Form Targets ..... 10
- 6. Measuring the Choice, Achievement, and Persistence of Institutional Development Targets ..... 13
- 7. Can Targets Be a Strategic Choice?..... 15
- 8. The Political Economy of Institutional Targets ..... 20
- 9. Validating Models with Out-of-Sample Prediction ..... 24
- 10. Conclusion ..... 29
- References..... 32

# 1. Introduction

Donor agencies spend an incredible amount of time, money, and energy trying to improve the performance of government institutions in developing countries. Common wisdom states 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, enforce regulations, and respond to citizen concerns (Esty and Porter 2005; Keohane and Levy 1996; VanDeveer and Dabelko 2001). Yet, in spite of the consensus that strong public sector institutions facilitate economic development and environmental management, little evidence exists that development assistance systematically contributes to improved institutional performance in developing countries (see Andrews 2013: 376-381; Booth 2011).

The lack of success is puzzling since both donor agencies and recipient countries have faced growing pressure in recent decades to measure and deliver results as part of the “results agenda” (Barder 2012; Birdsall 2008; Clemens et al. 2007; Ebrahim 2013; Khagram et al. 2009; Natsios 2010; Ramalingam 2013). The World Bank, for example, requires the specification of observable targets and a plan to monitor them before any project is approved. Continued flows of development assistance are partly based on the achievement of these targets. We explain in part why the “results agenda” has not led to substantial institutional improvements even as donor agencies emphasize institutional targets. Pressured to achieve specific and measurable institutional targets, recipient countries and their donor agency counterparts strategically choose targets that can be achieved quickly and maintained easily. In countries where it is difficult to improve public sector institutions or where rewards for improving public sector institutions are low, governments and their donor agencies counterparts receive greater payoffs by choosing easy and shallow institutional targets that do not correspond to improved functioning of government agencies.

To test this argument, we distinguish between targets that measure the public sector’s ability to achieve policy outcomes versus targets that measure how institutions are organized. We call these function and form targets, respectively. We expect that form targets are easier to achieve and maintain (Andrews and Pritchett 2012: 7). For example, it is easier to create a new law or agency than to ensure that an agency properly implements a law. We argue that in response to pressure to achieve targets, recipient countries often adopt the forms of institutions proposed by donors without addressing the underlying factors that

perpetuate institutional dysfunction.

We develop and test a theory about when and why recipient governments and donor agencies jointly pursue institutional form targets despite their dubious value. We particularly examine the choice of institutional targets in World Bank projects related to environmental management, an area of programming that focuses heavily on improving public sector institutions (Keohane and Levy 1996; VanDeveer and Dabelko 2001). Environmental management is also a sector strongly driven by donor country preferences, which should heighten external signaling pressures (Keohane and Levy 1996; Heltberg and Nielsen 2003; Hicks et al. 2008). We first develop a formal model to illustrate our argument. We also develop two original datasets by extracting institutional targets from post-project evaluations and pre-project appraisal reports, which we use respectively for an in-sample explanatory model and an out-of-sample predictive model about the choice of targets.

Our formal, explanatory, and predictive models all point to the same conclusion: the countries most in need of stronger public sector institutions are least likely to choose institutional targets that focus on functionality. While the payoff is often higher for achieving function targets, the recipient countries most in need of improved institutions have a lower probability of achieving such targets. These countries place greater emphasis on selecting targets they know they can achieve and maintain, since they have a greater need to maintain access to foreign assistance. The pressure to achieve targets therefore undermines the improvement of institutional function in these countries, which helps explain why externally financed institutional development is so difficult to achieve (Andrews 2013; Eubank 2012; Pritchett et al. 2013).

## **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 2008b; UNECA 2003; World Bank 1997; World Bank 2003). Yet the best available evidence suggests that institutional development programs for the public sector often fall short of expectations (Berg 1993; IEG 2008b; Levy and Kpundeh 2005; Meagher 2005; van de Walle 2001; World Bank 2010b). Progress is usually slow and discontinuous (IEG 2008b; World Bank 2010b), and many developing countries that received institutional development support in previous decades are still receiving support for the same activities today (Andrews 2013). Leading development economists who study foreign aid effectiveness, such as Birdsall (2008) and Easterly (2008), are most critical of donor-sponsored institutional development programs. Birdsall (2008: 517), for example, points to the "limited success of donors over the past three ...[or] more decades in supporting institutional development in the

many more countries that are not or have not ‘failed’ but are now variously labeled as weak or poorly performing.”

The primary explanation for the poor track record 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 worldwide experience with public sector programs offers the same diagnosis: “[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 2008b: 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 solutions for a recipient country (Arruñada 2007, 2012; Grindle 2004; Haggard et al. 2008; Pritchett and Woolcock 2004; Rodrik 2007).

Left unchecked, “blueprint” and “best practice” pressures from donors and international organizations can also result in dysfunctional recipient government behavior 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, together with donor agency staff can declare success without actually improving the functions that institutions are supposed to perform (Samuel 2013). Given reputational and material incentives to achieve observable and quantitative results, these form targets offer cheap and easy demonstrations of institutional development “progress” without costly improvements to long-run state capacity (Arruñada 2007; Cullen and Randall 2006; Hood 2006; Jacob 2005; Wynia et al. 2000).

The alternative to “best practice” or “blueprint” institutional development activities is context-specific projects to improve institutional function, as measured by the achievement of policy outcomes (Evans 2004; Andrews et al. 2012b; Chong et al. 2012; Woolcock 2013). But improving function-based measures of institutional performance is more difficult than adopting new institutional forms. 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 have the policy space to iteratively adapt to local constraints and opportunities (Andrews 2013; DFID 2003; Grindle 2004; Rodrik 2007; Pritchett et al. 2012). In light of these difficulties, we develop and test a theory about when and why recipient governments and donor agencies jointly pursue institutional form targets despite their dubious value. In particular, we provide and test a formal model that illustrates how differences in the difficulty of achieving

targets across countries interacts with differences in rewards for achieving targets.

### 3. Differences in the Difficulty of Institutional Development Targets

The first part of our argument is that form targets are easier to achieve and maintain than function targets. To investigate differences in the difficulty of achieving and maintaining institutional targets, we develop two datasets of institutional targets related to environment and natural resource management from World Bank post-project evaluations and project appraisal reports. We assign 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. 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 neither form nor function. These additional targets typically measure actions (e.g., personnel training) without measuring results. From a theoretical perspective, we do not consider action targets to be substantially different from form targets, but we test this assumption in the first stage of our analysis. After demonstrating that form and action indicators provide similar signaling opportunities, we collapse our discussion to form and function targets, combining indicators classified as form and action together and asking what causes countries to choose function targets.<sup>1</sup>

We expect that targets that can be achieved through one-off actions (e.g., a new law, task force, or

---

<sup>1</sup> 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.

training program) are easier to achieve and maintain than targets that require sustained action over time. If this is correct, then the ability of recipient governments and World Bank staff to act strategically regarding target selection is strong. As such, we propose the following hypothesis:

**H1: Institutional form targets are more likely to be achieved and maintained than institutional function targets.**

## **4. Rewards for Achieving and Maintaining Institutional Function Targets**

If form targets were easier to achieve and maintain than function targets and offered higher payoffs, recipient governments would always choose them. Since recipient governments and donor agencies choose function targets regularly, achieving and maintaining function targets must bring additional rewards. Our background interviews with World Bank officials suggest that formal 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 every project. Therefore, both World Bank staff and officials in recipient countries are likely to select more difficult targets when the rewards for doing so outweigh the risks of failure to meet targets. To explain the selection of targets, therefore, we must explain the joint incentives of recipient officials and World Bank staff who prepare projects for approval.

For the World Bank staff and recipient officials, the incentives to select function targets are similar and reinforcing. At the World Bank, which is the focus of our empirical analysis, the allocation of grants and concessional loans under the International Development Association (IDA) is directly linked to a recipient country's performance in meeting function targets (ADB 2005; World Bank 2010b). 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). Thus, IDA borrowers understand that improvements (declines) on function targets will result in more (less) funding, both in terms of the outcomes of projects and more general improvements in institutional function over time.

Once an overall resource envelope is established, the World Bank provides funding with some discretion. Funding through development policy loans and results-based financing is significantly more discretionary than funding through other mechanisms such as project support and technical assistance. The amount of discretion is usually tied to the counterpart government's 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 environmental impacts of major infrastructure projects or paying forest titleholders for protecting sensitive watersheds, activities often supported in previous projects (World Bank 2008b, 2010a; World Bank Carbon Finance Unit 2013).

Major regional development banks (the Inter-American Development Bank, the Asian Development Bank, and the African Development Bank) and several other multilateral institutions (IFAD, the Global Environment Facility, the Caribbean Development Bank, the European Commission) use similar criteria to establish the size of country resource envelopes and determine discretionary funding (ADB 2005, 2007; European Commission 2008, 2014; GEF 2013a). This creates an additional incentive for recipient governments to improve function target performance. The Global Environment Facility, for example, conditions environmental assistance on if environmental function targets are met through their “System for Transparent Allocation of Resources”.

Recipient countries that achieve function targets may also garner indirect rewards. First, domestic political constituencies in developing countries reward the achievement of function targets and improved public services, while often caring little about changes to institutional forms (Lake and Baum 2001; Stasavage 2005). Second, a reputation for sharing the policy preferences of donor agencies and maintaining functional government agencies can generate more international attention and support (Chwieroth 2013; Flores et al. 2013).

The incentives of World Bank staff reinforce the selection of function targets when recipient counterparts have a high probability of achieving them. 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 (Easterly 2001; IEG 2011; Phillips 2009; World Bank 1992). Conditional on securing the approval for loans, staff does receive credit for designing and operating projects that improve outcomes in recipient countries (Whittle 2013). 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 is likely to push for feasible targets that are ambitious enough to pass review. To judge which countries are likely to achieve function targets, we expect that staff at the World Bank will rely on past success in the country as a guide (Buntaine 2011).

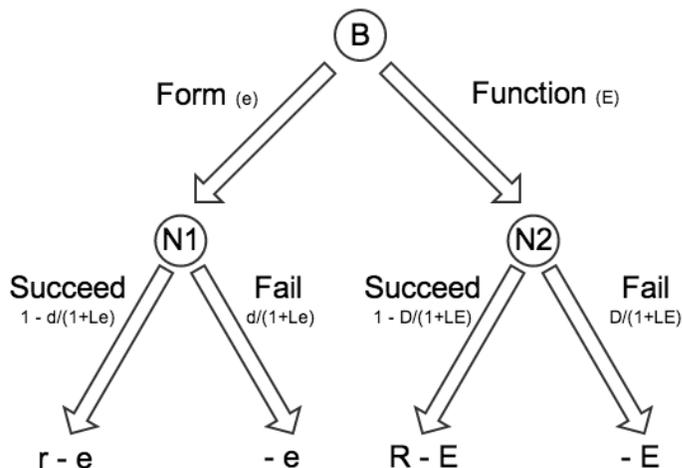
## 5. Theoretical Model: The Choice of Function or Form

### Targets

While achieving and maintaining function targets offers additional rewards, choosing these more ambitious targets is risky. In more formal terms, we expect the payoff for achieving and maintaining function targets to be greater than the payoff for achieving and maintaining form targets, which is greater than the payoff for failing to achieve and maintain either kind of target. The consideration then is whether the risk-adjusted payoff from function targets is greater than the risk-adjusted payoff from form targets, since the failure to achieve either will not lead to rewards. The result of failure may be negative, but for our purposes, all that matters is that the payoff for failure is lower than the payoff for success.

Consider a borrowing country that has secured a certain amount of financing for a World Bank project. This country and the World Bank team now must negotiate the targets to be achieved. Each side of the game tree has an identical structure, though the optimal levels of effort chosen by the borrowing country might vary according to the different rewards associated with achieving and maintaining form or function targets. Thus, the player first optimizes the allocation of effort to maximize the value of each branch of the game tree, and then chooses the form or function branch that has a higher expected payoff.

**Figure 1. The decision tree for selecting institutional targets**



A difficulty factor ( $d$  or  $D$ ) defines the probability of success in each branch of the tree. Intuitively, this is the inherent difficulty of achieving the particular target. We assume this value is strictly positive, but not more than 1, at which point the probability of success is zero without some effort expended by the borrowing government. We also assume that it is more difficult to achieve a function target, but we make

no assumptions about the relative degree of difficulty.

In both branches, the player allocates effort ( $e$  or  $E$ ) to achieve the target, which reduces the difficulty factor and increases the probability of success, but also reduces the payoff by the cost of the effort that is exerted. We assume that the reward of achieving a form target ( $r$ ) is less than the reward for achieving a function target ( $R$ ), though we do not make further assumptions about the difference of these rewards or their absolute values. In the case where form targets provide higher rewards, countries would always choose them because we assume (and test below) that they are easier to achieve and maintain. To maintain flexibility, we do not make assumptions about the return to effort. Instead, the effort expended is scaled by the “return to effort” parameter  $L$ . The parameters of this relatively flexible model that aims to outline our theory and derive our hypotheses are thus:

$e$ or $E$	Effort	$[0, \text{Inf})$
$d$ or $D$	Difficulty	$(0, 1)$
$L$	Return to Effort	$[0, \text{Inf})$
$r$ or $R$	Reward	$(0, \text{Inf})$

Solving the game in each branch of the tree in an identical manner produces the following expected values:

$$E[N1|e, d, L, r] = r - e - \frac{rd}{1 + Le}$$

$$E[N2|E, D, L, R] = R - E - \frac{RD}{1 + LE}$$

Maximizing the expected value of each branch with respect to effort ( $e$ ) gives the following (substituting upper and lower-case letters):

$$e = \frac{\sqrt{rdL} - 1}{L} \text{ if } \sqrt{rdL} > 1$$

$$e = 0 \text{ otherwise}$$

As the size of the reward for success decreases, the effort allocated by the player decreases. When the reward for achieving a target approaches zero, no effort is allocated to achieve the target. This will have important implications for middle-income countries that do not receive many external or domestic rewards for achieving form targets, as the expected value of the form branch is likely to be zero or close to it in many cases.

Having solved for the optimal level of effort, we compute the expected payoff for each branch of the tree by substituting the solved value of  $e$  into the payoff function. For the form branch:

$$E[N1|\sqrt{rdL} > 1] = r - \frac{2\sqrt{rdL} - 1}{L}$$

$$E[N1|\sqrt{rdL} \leq 1] = r - rd$$

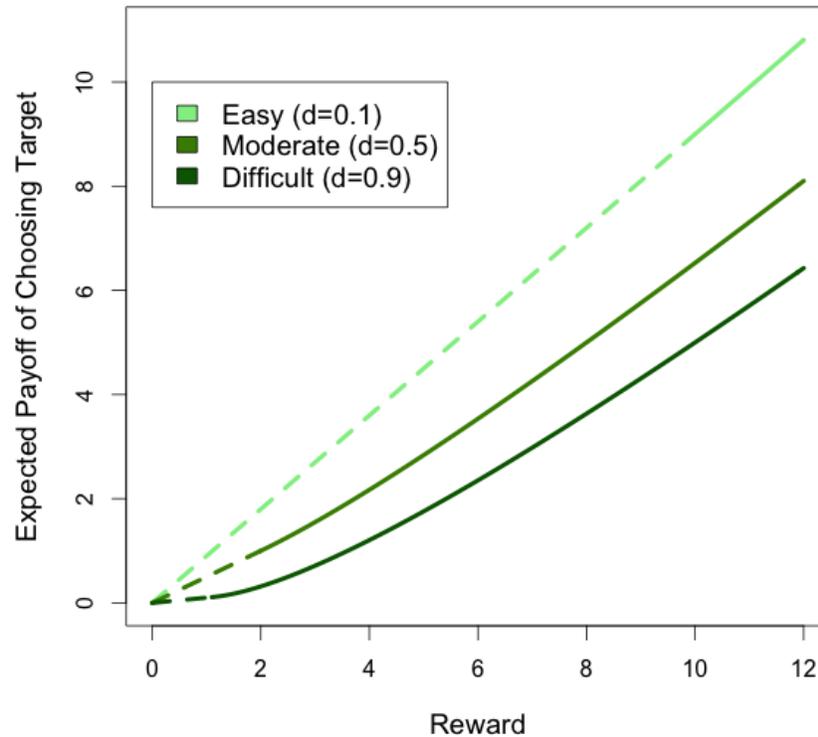
And symmetrically for the function branch:

$$E[N2|\sqrt{RD\bar{L}} > 1] = r - \frac{2\sqrt{RD\bar{L}} - 1}{L}$$

$$E[N2|\sqrt{RD\bar{L}} \leq 1] = R - RD$$

These payoff functions yield two testable hypotheses (Figure 2). First, both payoff functions are strictly decreasing in  $d$ . This means as the difference in difficulty between form and function targets increases, it becomes more likely that the payoff will be higher for form targets. Second, both payoff functions are strictly increasing in  $r$ , conditional on the optimal choice of effort. This means that as the difference of the reward between form and function targets increases, the payoff for choosing a function target will also increase. Importantly, if the reward is small, the borrowing country will quickly have incentives to choose no effort, so both the relative and absolute sizes of  $r$  and  $R$  matter for empirical predictions.

**Figure 2. Expected payoffs predicted by theoretical model for targets of different difficulty**



Notes: Dashed line indicates no effort exerted and solid line indicates effort exerted consistent with theoretical model.

In light of the fact that the ability to successfully execute institutional development programs varies significantly among borrower governments (Denizer et al. 2011; IEG 2008b), we expect utility-maximizing governments with low levels of implementation capacity to press for more easily implemented form targets. In the terminology of our theoretical model, recipient countries with strong pre-existing public sector institutions should have lower values on the difficulty factor.

**H2 (Difficulty): Recipient countries with higher pre-existing levels of governance and government capacity are more likely to choose function targets.**

We expect differences among recipient countries in the relative rewards for achieving function or form targets, even though achieving function targets will generally offer higher rewards.<sup>2</sup> Aid flows can make government budgets more fungible (Whitfield and Fraser 2010) and constitute a stable source of revenue (Samuel 2013). For recipient countries without access to alternative revenue sources, the benefit of securing continued aid is likely to take precedence over adopting riskier targets that might result in changes to aid flows.<sup>3</sup> Alternatively, for middle-income borrowing countries that have less access to concessional aid and instead borrow on commercial markets, the reward for achieving donor-supported form targets is likely to be low, since these countries gain little among donor or domestic constituencies for achieving form targets. Thus, we expect a country's level of dependence on concessional aid will be a key determinant of the relative difference between the rewards of achieving form and function targets.

**H3 (Rewards): Recipient countries with lower levels of dependence on foreign assistance are more likely to choose function targets than recipients that are more heavily dependent on foreign assistance.**

## 6. Measuring the Choice, Achievement, and Persistence of Institutional Development Targets

To test these hypotheses, we examine targets for World Bank projects related to environmental

---

<sup>2</sup> In countries where institutional dysfunction provides an opportunity for rent seeking by elites (in the urban environment sector, for example, see Davis 2004) or in situations where form reforms allow patronage opportunities for political leaders (Mwenda and Tangri 2005), it may be the case that function targets have lower rewards. In this case, the prediction is unambiguous. Countries will always choose form targets that are both less difficult and have higher payoffs. As a robustness check of our empirical models of target choice reported in Table 4, we remove all countries (11 or 87) in the sample that choose only form targets. The substantive and statistical results are unchanged.

<sup>33</sup> Isomorphic mimicry is purportedly more pervasive in countries that rely heavily upon foreign assistance from bilateral and multilateral donors (Pritchett et al. 2010; Andrews 2013).

management. 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). Thus, efforts in this sector focus heavily on building the capacity of state agencies to plan, monitor, and enforce. However, institution building in the environmental sector is often driven by the concerns of donor governments and aid agencies (Lewis 2003, Marcoux et al. 2011). Thus, it also offers aid-dependent recipient countries strong incentives to signal success to external donor audiences.

To test our hypothesis about the relative difficulty of achieving and maintaining form and function targets, we created an original dataset on the achievement and persistence of institutional targets used by the World Bank to measure the success of projects. 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 specific, observable indicators from these documents that were used to measure institutional development during the life of the project. We also used these project documents to measure the baseline and completion values of the targets. In cases where both baseline and completion data are 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 identified 826 indicators, around 3.3 per project, and re-measured post-completion data for 347 of them, a success rate of around 42%.

We coded each target as a form, action, or function indicator according to the definitions outlined in the previous section. 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 indicators, 380 were coded as measuring function targets, 285 as measuring form targets, and 161 as measuring action targets. Table 1 contains examples of targets to illustrate coding choices.

**Table 1. Examples of institutional development indicators from study sample**

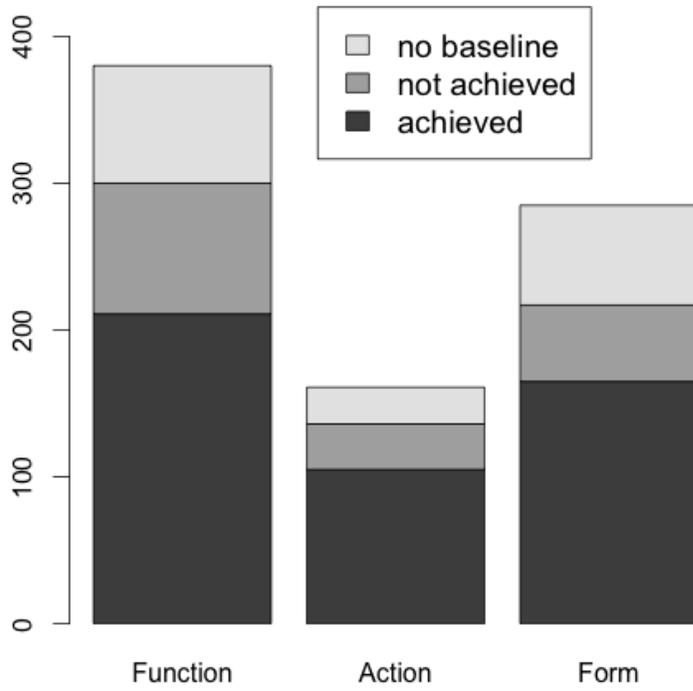
Country	Completion Year	Project Name	Indicator	Form, Action, or Result?
Argentina	2007	Native Forests & Protected Areas Project	Does Argentina regularly update its national inventory of forest resources?	Action
Ghana	2008	Natural Resource Management Project	Does the Environmental Protection Agency maintain 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
China	2009	Liao River Basin Project	Does the Liaoning Provincial Government have a river basin plan?	Form
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

## 7. Can Targets be a Strategic Choice?

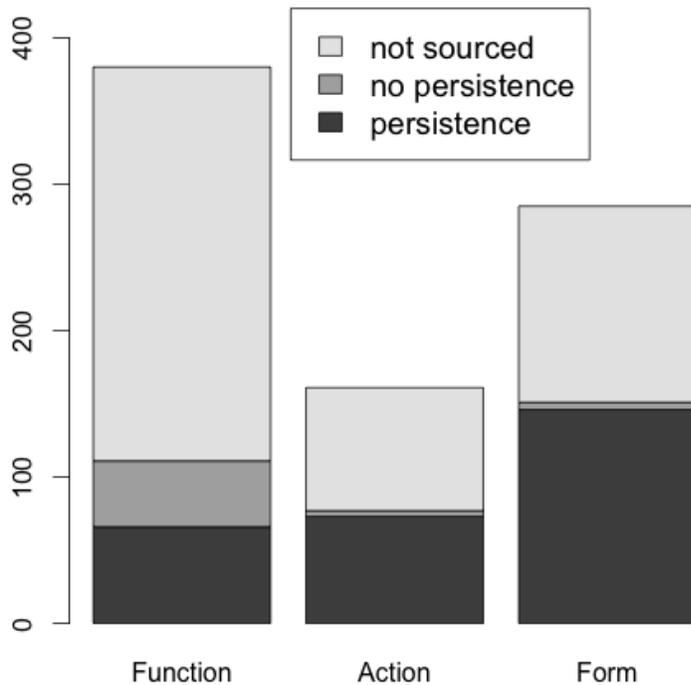
To demonstrate that the choice of form and function targets can be strategic, we must first demonstrate that form targets are easier to achieve and maintain than function targets. For targets with baseline data in the evaluation documents, 70% of function targets were achieved during project implementation, while 77% of action targets and 76% of form targets were achieved during project implementation (Figure 3). The potential for countries to maintain 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 4).

**Figure 3. Descriptive data on achievement of different types of targets**



**Figure 4. Descriptive data on persistence of different types of targets**



But these descriptive statistics risk omitted variable bias and case selection bias. Therefore, we present a multivariate analysis that controls for other variables that might be associated with the achievement and persistence of institutional targets or the missingness of data. In this analysis, we are forced to assume that the outcome variables are missing at random (MAR) — that missingness does not depend on the value of the outcome variable conditional on the predictor variables. This is a strong assumption, since it may be the case that even conditional on the predictor variables, indicators that do not persist are less likely to be observed during the re-measurement wave. Vach (1994) notes that when only the outcome variable is missing in logistic regression, the parameter estimates (not the intercept) will be unbiased if the proportional odds assumption holds. Thus, we can only interpret the parameter values in terms of their effect on odds, rather than their effect on predicted probabilities even if the MAR assumption cannot be sustained. 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).

In light of these modeling assumptions, we specify 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 our main predictor variables of target

type. The resulting estimates about the effect of target type on achievement and persistence therefore reflect the relative difficulty of targets *within* particular countries, levels, and projects. This specification also has the advantage of ensuring that our results are not biased by missingness, to the extent that missingness is fully predicted by these broad factors.

For both the models related to the achievement of institutional targets and the models related to the maintenance 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. Thus our models seek to identify systematic differences in the difficulty of achieving and maintaining targets within countries, levels, and projects. 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. Finally, in our second model of target persistence, we include an indicator variable for whether the target was achieved during implementation, since this might correspond to borrower commitment or capacity, but again, this is not necessarily a project-specific effect. Table 2 summarizes the results. The results confirm the hypothesis that form and action targets are easier to achieve and maintain than function targets, which is the reference type in the models.

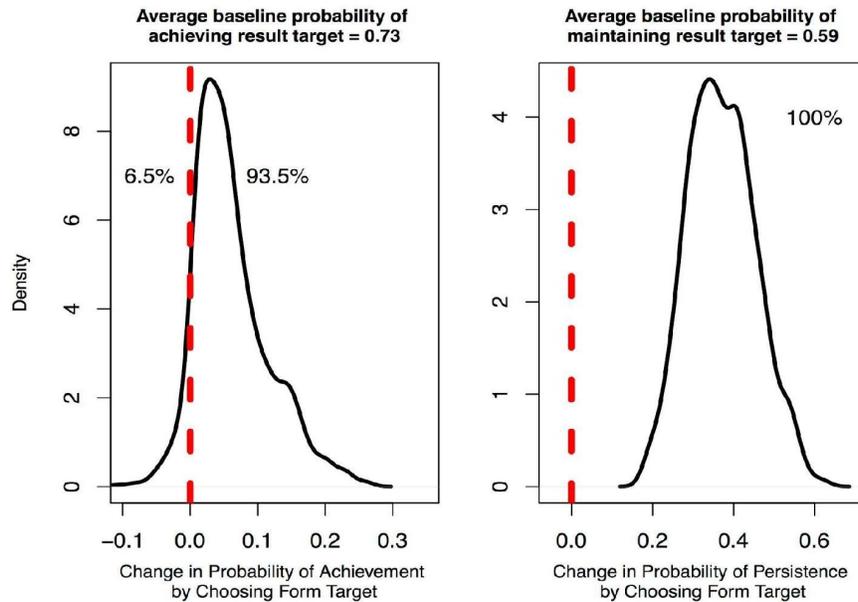
**Table 2. Achievement and persistence of institutional development targets**

DV:	Achievement	Achievement	Persistence	Persistence
<b>H1: Form Indicator</b>	<b>0.41</b> (0.28) [0.07]	<b>0.43</b> (0.28) [0.06]	<b>3.03</b> (0.51) [0.00]	<b>3.04</b> (0.59) [0.00]
<b>H1: Action Indicator</b>	0.35 (0.32) [0.13]	0.40 (0.32) [0.10]	<b>2.55</b> (0.56) [0.00]	<b>2.60</b> (0.59) [0.00]
<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)
<b>Project</b>	Yes (208)	Yes (204)	Yes (147)	Yes (126)
<b>Country</b>	Yes (80)	Yes (80)	Yes (73)	Yes (66)
<b>Level</b>	Yes (4)	Yes (4)	Yes (4)	Yes (4)
Observations	653	645	339	281
Deviance Reduction w/ Fixed Parameter	0.01	0.01	0.18	0.18

*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 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 3). 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 5. The effect of target type choice on achievement and persistence of institutional developments**



## 8. The Political Economy of Institutional Targets

The results from the previous section indicate that target choice is a strategic option for aid agency staff and recipient governments. We hypothesized that World Bank staff and recipient governments should favor function targets when there is a high probability they will achieve them (H2). Additionally, World Bank staff and recipient governments should favor function targets when the rewards for achieving form targets are low because aid dependence is not a major concern (H3). In this case, the additional rewards for selecting and achieving function targets are relatively high in comparison.

To measure the difficulty of achieving function targets for recipient countries, we use two operationalizations. First, we use a binary variable that is positive whenever at least 75% of the recipient country's post-project evaluations in the previous five years completed by the Independent Evaluation Group rates the overall outcome of the project as successful (H2a). Second, we use the Worldwide Governance Indicators variable for government effectiveness, which measures the ability of government agencies to successfully implement policies (H2b).

To measure the difference in the reward for form and function targets, we use three operationalizations.<sup>4</sup> First, we use the project-level proportion of funding from the World Bank's International Bank for

<sup>4</sup> We have chosen not to employ the conventional measure of net official development assistance as a percentage of gross national income because of substantial missing data.

Reconstruction and Development versus the International Development Association (H3a). Recipient countries gain access to IBRD lending through creditworthiness and sound fiscal management, rather than performance-based allocation standards that use data on project outcomes. IBRD borrowers tend to have access to commercial bond markets, and they gain little reward for signaling the achievement of form target to donors. Second, 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 (H3b). Countries with higher per capita income are likely to enjoy a larger tax base, which reduces dependence on foreign assistance and the rewards for signaling to donor organizations. Third, we use the share of GDP from natural resource rents for each recipient country (H3c). Natural resources rents decrease dependence on foreign aid and thus the reward for signaling to donor organizations. We use a measure developed by Hamilton and Clemens (1999) that sums rents from 14 fuel and nonfuel mineral resources.

Like our models about the difficulty of achieving and maintaining targets, we want to ensure that any effect of difficulty and reward on the choice of indicators accounts for country specific factors and the level of implementation. We therefore specify a random-effects model where the intercept varies by country and level of implementation. We also control for variables that are likely to be independent from country and level random effects. First, because the results agenda and the use of the results frameworks has grown more common at 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 be more likely to elicit 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.<sup>5</sup>

Our model results largely support our hypotheses that countries that can achieve function targets with greater ease and countries with options for non-aid revenue are more likely to choose function targets (Table 3). Countries that have a record of achieving satisfactory outcomes in at least 75% of their projects during the five years prior to project approval are more likely to choose function targets (H2a). As these countries are likely to be in the good graces of aid agencies, they face less of a need to signal achievement to external audiences. Furthermore, World Bank staff assigned to these projects should be more confident about choosing ambitious institutional development targets.

---

<sup>5</sup> 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.

Although we would like to distinguish between the effects of government effectiveness (H2b), the proportion of IBRD lending (H3a), and higher levels of development as measured by per capita income (H3b) on the choice of function targets, these variables are highly correlated. Thus, we cannot completely disentangle whether these variables correspond to a decreased difficulty of achieving and maintaining targets or a decreased reward for achieving form targets. We include them in successive models to demonstrate robustness across specifications. All of these variables have a positive effect on the probability that the country will choose a function target, which supports our hypotheses. Across all of our model specifications, we find that countries with greater access to natural resource rents as a proportion of GDP are also more likely to choose function targets. We interpret this to mean that countries with less dependence on aid donors have lower rewards for accepting form targets that only signal to donor audiences. When running the same models on the subset of the countries that do not only select form targets, the results do not change substantively.

This suggests the model is not sensitive to our assumption that the rewards for achieving function targets are larger than form targets, since we remove the subset of countries that might have only chosen form targets because they are both easier and higher reward. Likewise, the results are insensitive to the removal of “action” targets from the sample.

**Table 3. Choice of function institutional development targets (at project approval)**

DV:	“Function” Target	“Function” Target	“Function” Target
<b>H2a: Successful Evaluation Record</b> (At least 75% satisfactory AY-1 to AY-5)	<b>1.39</b> (0.46) [0.00]	<b>0.68</b> (0.31) [0.01]	<b>0.41</b> (0.29) [0.08]
<b>H2b: Government Effectiveness (AY-1)</b>	<b>0.94</b> (0.53) [0.04]		
<b>H3a: IBRD proportion</b>		<b>0.84</b> (0.34) [0.01]	
<b>H3b: GDP per capita (\$1k @ AY-1)</b>			<b>0.17</b> (0.09) [0.02]
<b>H3c: Resource Rents / GDP (% @ AY-1)</b>	<b>0.04</b> (0.03) [0.07]	<b>0.05</b> (0.02) [0.01]	<b>0.05</b> (0.02) [0.01]
Approval Year (centered linear)	<b>0.18</b> (0.09)	0.02 (0.03)	0.02 (0.03)
Environment Sector Board	<b>-1.92</b> (0.66)	<b>-1.62</b> (0.37)	<b>-1.45</b> (0.33)
<i>Level R.E.</i>	Yes (4)	Yes (4)	Yes (4)
<i>Country R.E.</i>	Yes (61)	Yes (85)	Yes (85)
Observations	364	773	805

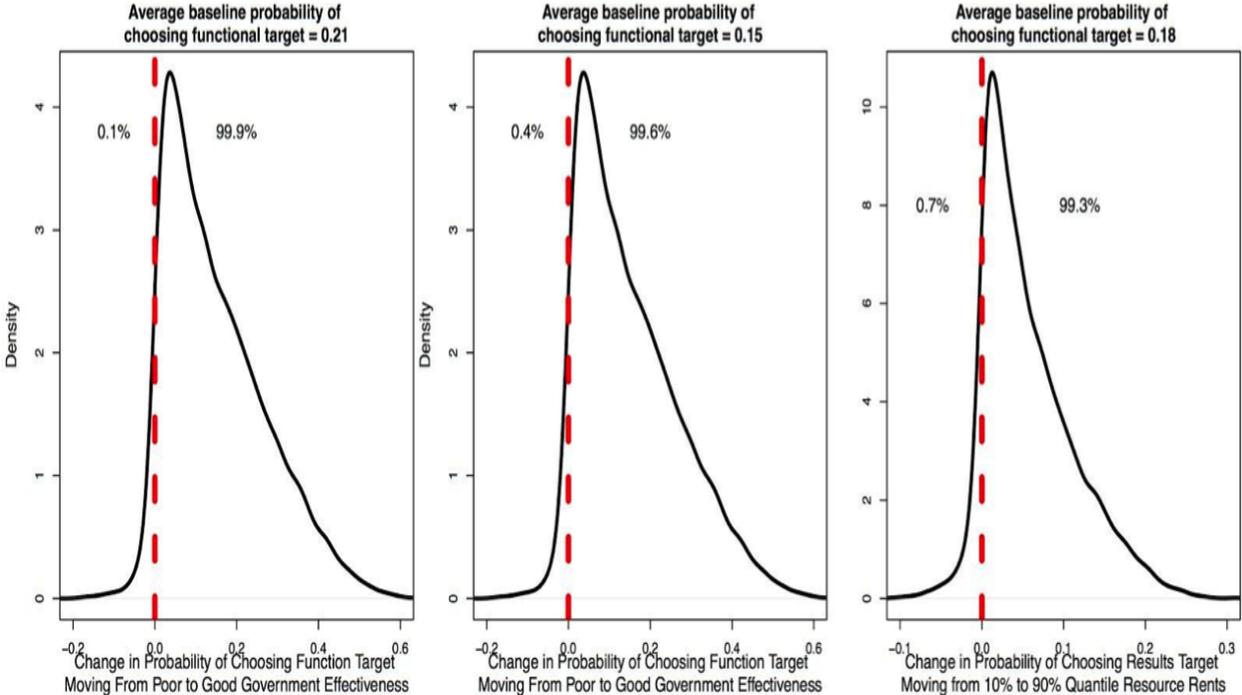
*Model cells list: Parameter estimate; (Standard Error); [p-value of one-sided z-test].  
All models are random-intercept logit fitted by Laplace approximation with levels as indicated.*

To aid substantive interpretation of these models, we produce first difference simulations based on the first model in Table 4. These simulations display the increased probability of choosing a function target based on changes in the main predictor variables (Figure 4). We randomly draw from the coefficient distributions, including the varying intercept, and then compute pairs of predicted probabilities, varying only one predictor variable at a time.

Several of our predictor variables influence the probability that a country will select targets that measure institutional function. For the average project allocated to a country without a satisfactory evaluation record, the probability of choosing a function target increases from 21% to 44% after having achieved a

satisfactory record. For the average country with a government effectiveness score at the 10% quantile of the sample, the probability of choosing a function target increases from 15% to 30% when rising to the 90% quantile of the sample on this score. Likewise, for the average country at the 10% quantile of natural resource revenue in the sample, the probability of choosing a function target increases from 19% to 24% when rising to the 90% quantile in natural resource revenue. Together, these results indicate that borrowing countries with a higher probability of achieving results and lower dependence on revenues from development assistance are more likely to choose function targets. These are precisely the countries that are least in need of donor-supported institutional development, which helps explain why institutional results have been disappointing for poor aid recipients. These countries instead favor easy, shallow, form targets.

**Figure 6. Effect of predictor variables on choosing function-based target**



In addition, we ran specifications that included other variables that might be expected to influence the choice of targets such as project length, total project funding, and strength of environmental NGO community. These variables did not predict the choice of targets and did not otherwise influence the results of our models.

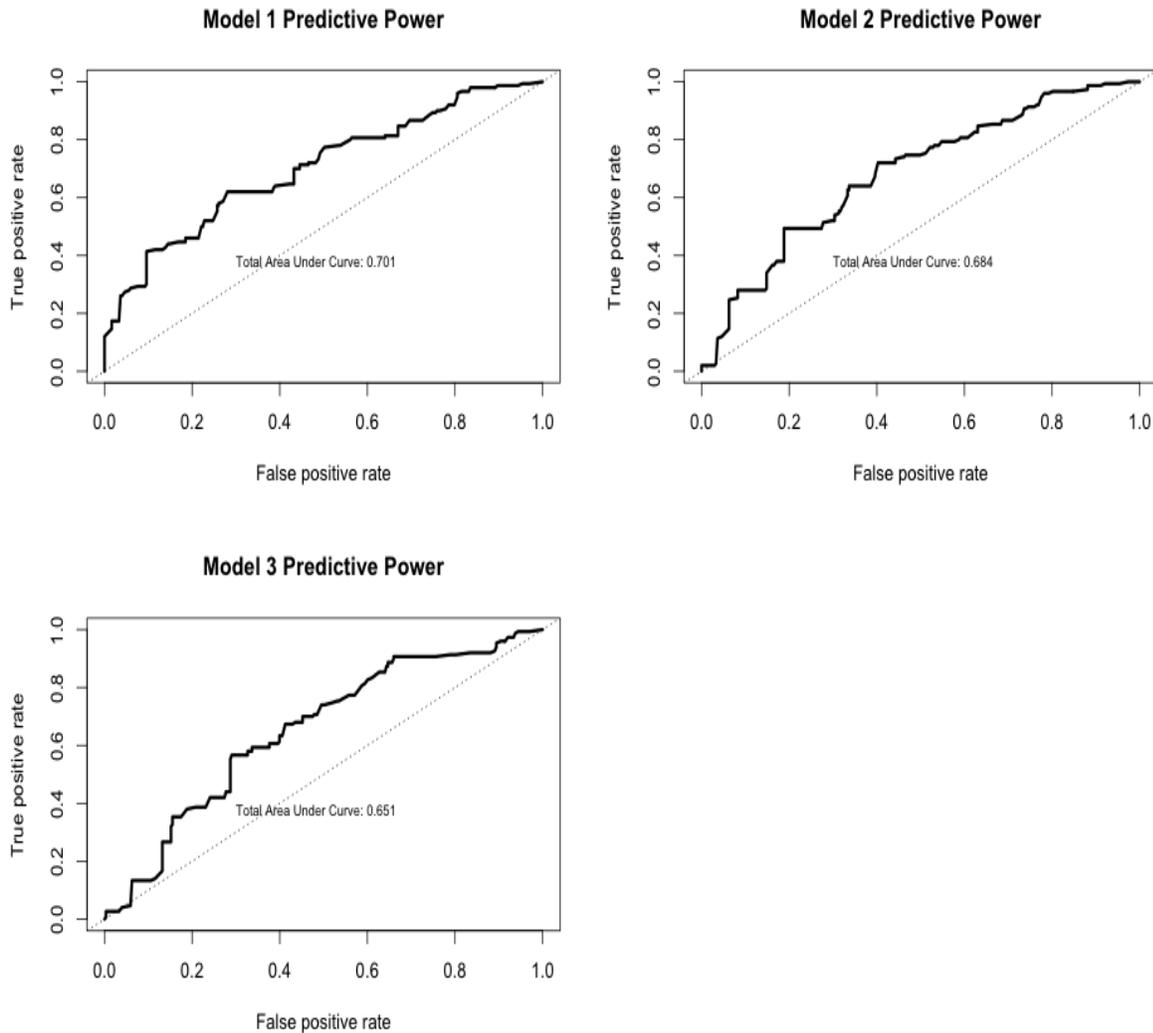
## 9. Validating Models with Out-of-Sample Prediction

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. Prior to collecting these data, we reported the results in Table 4 in working paper form at various conferences and workshops. 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). In total, we extracted and coded 454 institutional targets from project appraisals using the same procedure reported above. We used each of the three models in Table 4 to predict whether a function target would be chosen 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. 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 three models are displayed in Figure 7:

**Figure 7. AUC curves for out of sample prediction using the three models**

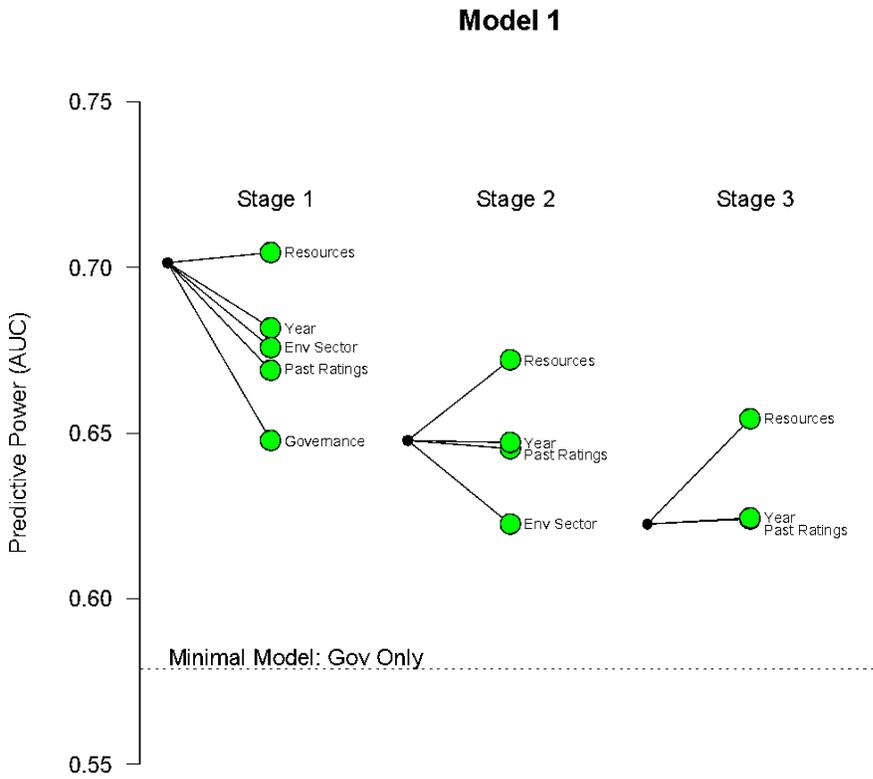


All three models perform better than random guessing. Model 1 performs the best, indicating that adding government effectiveness has a significant positive impact on predictive power. Models 2 and 3 have lower predictive power because they predict a high probability of function target selection for two projects when form targets were actually chosen. 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.

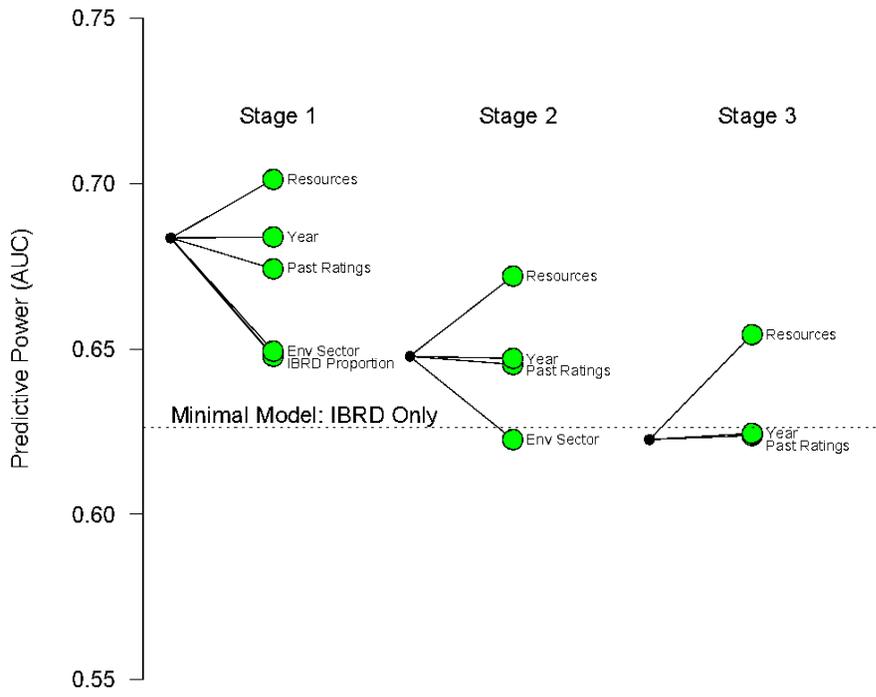
To assess how each independent variable contributed to predictive power, we constructed partial models by removing one predictor variable at a time. We then calculated the AUC for each of these partial models and compared it to the AUC of the full model. We then repeated this process for the partial model that removed the most predictive predictor variable, resulting in a new set of partial models. We repeated the

process for a third stage, leaving partial models with only two predictor variables (Figure 8).

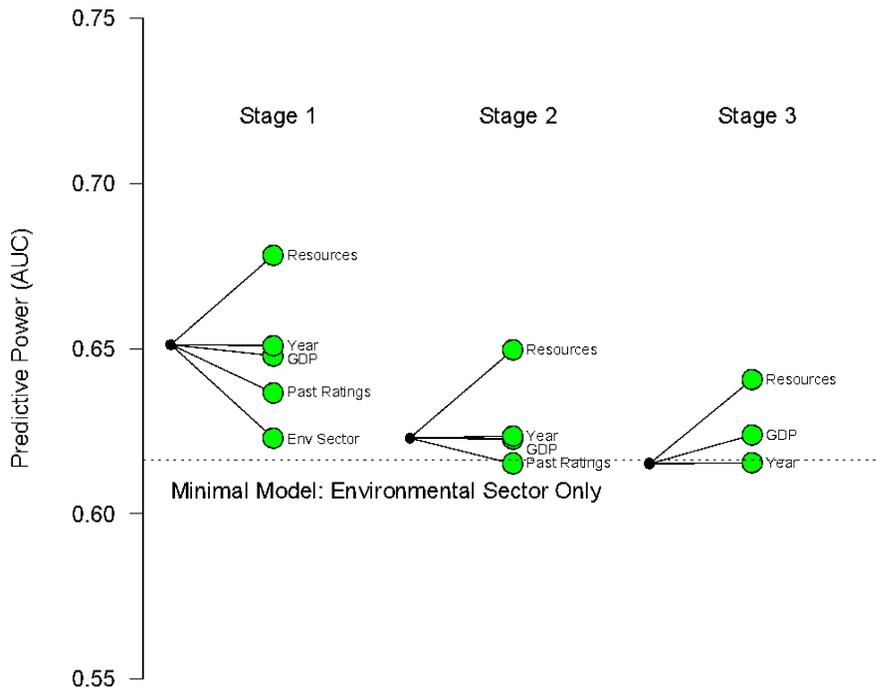
**Figure 8. Representation of the drop in AUC accompanying the removal of each IV**



### Model 2



### Model 3



Government effectiveness was the most important variable for predictive power. This supports the notion that countries that can achieve results with less difficulty are more likely to select function targets. Not only did the model that included government effectiveness have the most predictive power, but AUC dropped most significantly across all the models when government effectiveness was removed from the fitted model, falling from .701 to .648. Past project ratings have a smaller, but noticeable impact on predictive power across models, although once the most predictive variables are dropped its contribution to predictive power declines significantly.

The proportion of IBRD lending to the borrowing country also significantly increased the predictive power of Model 2; without it AUC fell from .684 to .648. This supports the notion that countries with access to other sources of financing receive comparatively smaller rewards for signaling to donors with form targets. GDP per capita as a measure of aid dependence contributes only modestly to predictive power in Model 3, but does not contribute to predictive power once past project ratings are removed. Removing 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. However, four of the five variables that we use to test our hypotheses add to predictive power out-of-sample, which generates further evidence for our theory.

## 10. 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 institutional development projects. 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 model of indicator selection with explicit predictions about when form based indicators of success are most likely to be selected. In particular, we expect that countries with low levels of institutional function and high aid dependency are the most likely to rely on form based reforms. The results from our in-sample explanatory and out-of-sample predictive models provide strong support for this conclusion. We find that countries with effective government agencies and low reliance on concessional development finance are more likely to choose function targets. More tenuously, we find that a successful record with World Bank projects and higher levels of per capita

income promote the choice of 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 functional reforms.

These findings call into question the conventional wisdom that building stronger monitoring procedures for aid projects will 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-a-vis targets; improved monitoring and evaluation should measure de facto institutional function.<sup>6</sup> 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 choice between stock form and functional indicators, even within organizations that prioritize accountability and measurable results, can set in motion the strategic logic 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.<sup>7</sup> 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 innovation will catalyze discussion and debate within donor agencies 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. Broad indices of institutional quality are analyzed

---

<sup>6</sup> 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.

<sup>7</sup> 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 project-level institutional outcomes that donors explicitly attempt to support (Knack et al. 2003; Chauvet and Collier 2008). This limits the way research can inform operational decisions. By going beyond country-level indicators of institutional development, we are better able to test a theory about how pressure to measure results affects the institutional development choices of recipient government and World Bank staff.

What should donor agencies do in light of incentives to adopt form targets that are easy to achieve and maintain? We would argue that, 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. The results agenda can promote institutional performance as measured by evaluators, but if aid agencies and development banks are serious about building strong and effective states they ought to select targets that more effectively track whether and to what extent states discharge their core functions.

## References

- Acemoglu, Daron and James Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York: Crown Business.
- Amin, Mohammad, and Simeon Djankov. 2009. "Natural Resources and Reforms." World Bank Policy Research Working Paper 4882. Washington D.C.: World Bank.
- Andrews, Matt. 2013. *The Limits of Institutional Reform in Development*. Cambridge: Cambridge University Press.
- Andrews, Matt, Lant Pritchett, and Michael Woolcock. 2012a. "Looking Like a State: Techniques of Persistent Failure in State Capability for Implementation." UNU-WIDER Working Paper 2012/63. Helsinki: UNU-WIDER.
- Andrews, Matt, Lant Pritchett, and Michael Woolcock. 2012b. "Escaping Capability Traps through Problem-Driven Iterative Adaptation (PDIA)." UNU-WIDER Working Paper 2012/64. Helsinki: UNU-WIDER.
- Archer, Candace C., Glen Biglaiser, and Karl DeRouen. 2007. "Sovereign Bonds and the "Democratic Advantage": Does Regime Type Affect Credit Rating Agency Ratings in the Developing World?" *International Organization* 61 (2): 341-365.
- Arndt, Christiane and Charles Oman. 2006. *Uses and Abuses of Governance Indicators*. Paris: OECD.
- 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
- Barder, Owen. 2012. "The Implications of Complexity for Development." The Kapuściński Lecture.
- Bates, Robert. 2009. *Prosperity & Violence: The Political Economy of Development*. 2nd Ed. New York, NY: W. W. Norton & Company.
- Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer, and Patrick Walsh, 2001. "New tools in comparative political economy: The Database of Political Institutions." *World Bank Economic Review* 15 (1): 165-176.
- 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.
- 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.
- Buntaine, Mark T. 2011. "Does the Asian Development Bank Respond to Past Environmental Performance when Allocating Environmentally Risky Financing?" *World Development* 39 (3), 336-350.
- Burki, Shahid Javed and Guillermo Perry. 1998. *Beyond the Washington Consensus: Institutions Matter*. Washington D.C.: World Bank.
- 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.
- Chwieroth, Jeffrey M. 2007. "Neoliberal Economists and Capital Account Liberalization in Emerging Markets". *International Organization* 61 (2): 443-463.
- 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.
- 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.
- Dai, Xinyuan. 2007. *International institutions and national policies*. Cambridge University Press.
- Denizer, Cevdet, Daniel Kaufmann, and Arat Kraay. 2011. "Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance." Background paper for IDA at 50 Report, World Bank.

- Department for International Development (DFID). 2003. *Promoting Institutional & Organizational Development: A Source Book of Tools and Techniques*. London, UK.
- Diez, Jordi. 2012. *Political Change and Environmental Policymaking in Mexico*. London, UK: Routledge.
- 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. <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](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](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.
- 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.
- Garbarino, Sabine and Jeremy Holland. 2009. "Quantitative and Qualitative Methods in Impact Evaluation and Measuring Results." Issues Paper.
- Geertson, Michael. 2010. "Achieving Quick Impact in Millennium Challenge Corporation Threshold Projects." *Yale Journal of International Affairs* 5: 34-48.
- 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.: GEF.
- Gelos, R. Gaston, Ratna Sahay, and Guido Sandleris. 2011. "Sovereign borrowing by developing countries: What determines market access?" *Journal of International Economic*. 83 (2): 243-25.
- Girod, Desha and Meir Walters. 2012. "Elite-Led Democratisation in Aid-Dependent States: The Case of Mauritania." *Journal of North African Studies* 17 (2): 181-193.

- 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.
- Haas, Peter M. 1992. "Introduction: Epistemic Communities and International Policy Coordination." *International Organization* 46 (1): 1-35.
- Haas, Peter M. 1990. *Saving the Mediterranean: the politics of international environmental cooperation*. New York: Columbia University Press.
- 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.
- Hallward-Driemeier, Mary and Lant Pritchett. 2011. "How Business is Done and the 'Doing Business' Indicators: The Investment Climate when Firms have Climate Control." World Bank Policy Research Working Paper 5563. Washington D.C.: World Bank.
- Hamilton, Kirk and Michael Clemens. 1999. "Genuine savings rates in developing countries." *World Bank Economic Review* 13:333–356.
- Harford, Tim. 2011. *Adapt: Why Success Always Starts With Failure*. New York, NY: Macmillan.
- Heckelman, Jac and Stephen Knack. 2008. "Foreign aid and market-liberalizing reform." *Economica* 75 (299): 524–548.
- Heltberg, Rasmus and Uffe Nielsen. 2000. "Foreign Aid, Development and the Environment." In *Foreign Aid And Development: Lessons Learnt And Directions For The Future*, eds. Finn Tarp and Peter Hjertholm. London: Routledge, 271-289.
- 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.
- Hood, Christopher. 2006. "Gaming in Targetworld: The Targets Approach to Managing British Public Services." *Public Administration Review* 66 (4): 515-521.
- 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). 2010. “World Bank Group Cooperation: Evidence and Lessons from IEG Evaluations.” IEG Evaluation Brief 11. Washington D.C.: World Bank.
- Independent Evaluation Group (IEG). 2008a. “Environmental Sustainability: An Evaluation of World Bank Group Support.” Washington D.C.: World Bank.
- Independent Evaluation Group (IEG). 2008b. “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 .
- Jordan Smith, Daniel. 2003. ‘Patronage, Per Diems and the “Workshop Mentality”’: The Practice of Family Planning Programs in Southeastern Nigeria.’ *World Development* 31 (4): 703–715
- 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.
- Killick Tony.1998. *Aid and the political economy of policy change*. New York: Routledge.
- Knack, Stephen. 2009. “Sovereign rents and quality of tax policy and administration.” *Journal of Comparative Economics* 37 (3): 359-371.
- Knack, Stephen. 2007. “Measuring Corruption: A Critique of Indicators in Eastern Europe and Central Asia.” *Journal of Public Policy* 27 (3): 255-291
- Kragelund, Peter. 2012. “The Revival of Non-Traditional State Actors' Interests in Africa: Does it Matter for Policy Autonomy?” *Development Policy Review* 30 (6): 703-718.
- Krause, Philipp. 2004. “From Epistemic Change to Action: The Case of State Reform at the World Bank.” Paper presented at the annual meeting of the International Studies Association, Le Centre Sheraton Hotel, Montreal, Quebec.
- Lake, David A., and Matthew A. Baum. 2001. “Th Invisible Hand of Democracy Political Control and the Provision of Public Services.” *Comparative Political Studies* 34 (6): 587-621.
- 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.
- Mahoney, James and Kathleen Thelen (eds.). 2009. *Explaining Institutional Change: Ambiguity, Agency, and Power*. Cambridge, UK: Cambridge University Press.
- Marcoux, Christopher, Claire Peeters, and Michael J. Tierney. 2011. Principles or Principals? Institutional

- Reform and Aid Allocation in the Global Environment Facility (GEF). Presented at the 2011 Political Economy of International Organizations Conference.
- 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.
- Mosley Paul, Jane Harrigan, and John Toye. 1991. *Aid and power: The World Bank and policy-based lending, vol. 1*. London, UK: Routledge.
- Mwenda, Andrew M., and Roger Tangri. 2005. "Patronage politics, donor reforms, and regime consolidation in Uganda." *African affairs* 104 (416): 449-467.
- Natsios, Andrew. 2010. "The Clash of the Counter-Bureaucracy and Development." Center for Global Development Essay. July 2010. Washington DC: Center for Global Development.
- 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.
- Phillips, David A. 2009. *Reforming the World Bank: Twenty years of trial-and error*. Cambridge University Press.
- Pritchett, Lant. 2002. "It pays to be ignorant: a simple political economy of rigorous program evaluation." *Journal of Policy Reform* 5 (4): 251-269.
- 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.
- Raustiala, K. 1997. "States, NGOs, and international environmental institutions." *International Studies Quarterly* 41 (4): 719-740.
- 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.
- Rodrik, Dani. 1996. "Understanding Economic Policy Reform". *Journal of Economic Literature* 34: 9-41.
- Ross, Michael 2006. "Is democracy good for the poor?" *American Journal of Political Science* 50 (4): 860-874.

- Ruggie, John. 1975. "International responses to technology: concepts and trends." *International Organization* 29 (3): 557 - 583.
- Saiegh, Sebastian M. 2005. 'Do Countries Have a "Democratic Advantage"?: Political Institutions, Multilateral Agencies, and Sovereign Borrowing.' *Comparative Political Studies* 38 (4): 366-387.
- 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.
- Shirley, Mary. 2005. "Can aid reform institutions?." *Chevy Chase, Md: The Ronald Coase Institute, Working Paper W 6*.
- Smith, Alastair. 2008. "The Perils of Unearned Income." *Journal of Politics* 70 (3): 1-14.
- Snijders, Tom, and Roel Bosker. 1999. "Multilevel analysis: An introduction to basic and advanced multilevel modeling." London, UK: Sage Publishing.
- Stasavage, David 2005. "Democracy and education spending in Africa." *American Journal of Political Science* 49 (2): 343-358.
- 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.
- 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.
- Vamvakidis, Athanasios. 2007. "External Debt and Economic Reform: Does a Pain Reliever Delay the Necessary Treatment?" IMF Working Paper.
- 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.
- Weymouth, Stephen and J. Muir Macpherson. 2012. "The Social Construction of Policy Reform: Economists and Trade Liberalization around the World." *International Interactions* 38 (5): 670-702.
- Winters, Matthew S. 2012. "The Obstacles to Foreign Aid Harmonization: Lessons from Decentralization Support in Indonesia." *Studies in Comparative International Development* 77 (3): 316-341.

- 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.
- Whitfield, Lindsay (ed.). 2009. *The Politics of Aid: African Strategies for Dealing with Donors*. Oxford, UK: Oxford University Press.
- Whitfield, Lindsay and Alastair Fraser. 2010. "Negotiating Aid: The Structural Conditions Shaping the Negotiating Strategies of African Governments." *International Negotiation* 15 (3): 341–366.
- Woolcock, Michael. 2013. "Using Case Studies to Explore the External Validity of 'Complex' Development Interventions." *Evaluation* 19 (3): 229-248.
- World Bank. n.d. "Harmonized Evaluation Criteria for ICR and OED Evaluations." Washington D.C.: World Bank. <http://ieg.worldbankgroup.org/Data/HarmonizeEvalCriteria.pdf>
- World Bank. 2014. "Learning and Results in World Bank Operations: How the Bank Learns." Washington DC: World Bank. [https://ieg.worldbankgroup.org/Data/reports/learning\\_results\\_eval.pdf](https://ieg.worldbankgroup.org/Data/reports/learning_results_eval.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://message.worldbank.org/www-ext-404http://siteresources.worldbank.org/EXTWBDEBTSTA/Resources/3561369-471255619840053/Note\\_on\\_Statistical\\_Capacity\\_Indicator\\_2009\\_BBSC.pdf.html](http://message.worldbank.org/www-ext-404http://siteresources.worldbank.org/EXTWBDEBTSTA/Resources/3561369-471255619840053/Note_on_Statistical_Capacity_Indicator_2009_BBSC.pdf.html)
- 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.