

## **Why the “Results Agenda” Produces Few Results: An Evaluation of the Long-Run Institutional Development Impacts of World Bank Environmental Projects**

**Mark T. Buntaine**  
UC-Santa Barbara  
[buntaine@bren.ucsb.edu](mailto:buntaine@bren.ucsb.edu)

**Benjamin P. Buch**  
Stanford University  
[bpbuch@stanford.edu](mailto:bpbuch@stanford.edu)

**Bradley C. Parks**  
College of William & Mary  
[bcpark@wm.edu](mailto:bcpark@wm.edu)

**14 January 2014**

### **Abstract**

Aid agencies have focused an incredible amount of time, money, and energy on improving formal institutions in developing countries, but broad improvements to institutional performance remain elusive. We investigate why institutional results have been so difficult to achieve with development finance, even as aid agencies have increased their monitoring and evaluation of observable targets as part of the “results agenda.” We show that recipient governments and project staff are more likely to achieve and maintain institutional targets when they choose form-based targets rather than functional targets as part of World Bank environment and natural resource management projects. This sets in motion a strategic logic where only countries that have a high likelihood of achieving functional targets are likely to choose them, because they can receive additional rewards as compared to when they achieve form-based targets. This leads to a separating equilibrium where the recipient countries most in need of institutional reform focus on outward, form-based targets that are less likely to improve the functionality of formal institutions.

## Introduction

Donor organizations like the World Bank have devoted an incredible amount of time, money, and energy to improving government institutions in developing countries. It is now common wisdom that “institutions matter” and that development rarely occurs in the absence of government agencies that have the capacity and incentives to fulfill basic functions of the state (Herbst 1990; Haggard 1990; Evans 1995; Callaghy 1988; Rodrik 2000; Acemoglu and Robinson 2012; Besley and Persson 2010). Strong government agencies provide law and order, enforce the rules of economic exchange, raise revenue, and deliver essential public services. In the environment sector, the protection of the public from pollution and the rational use of natural resources depends on government agencies that are capable of technical planning, enforcing regulations, and responding to citizen concerns. Despite the consensus that strong formal institutions facilitate economic development and environmental management, there is little evidence that aid has systematically contributed to improved institutional performance and functionality in developing countries. Although donor agencies have widely supported institutional development and government reform, around half of recipient countries have experienced a decline in indices of institutional performance since the late-1990s (Andrews 2013, KL. 376-381).

This is puzzling, since both donor agencies and recipient countries have faced more pressure to measure and deliver results during this period, with the inception of the “results agenda” in development finance. The “results agenda” attempts to make donor organizations and recipient countries more accountable by requiring them to specify measurable goals of development assistance when designing projects.<sup>1</sup> At aid organizations like the World Bank, the specification of observable goals and a plan to measure them is required before any project can be approved. Multilateral donors and their counterparts in recipient governments are expected to meet these goals to maintain flows of development assistance.

In this paper, we explore one explanation for why the “results agenda” in development finance has led to disappointing institutional outcomes. We argue that it encourages recipient countries that are most in need of institutional development and their counterpart staff at the

---

<sup>1</sup> For background on the “results agenda” see Clemens et al. 2007; Khagram et al. 2009; Natsios 2010; Barder 2012; Ramalingam 2013; Ebrahim 2013.

World Bank to focus on shallow, but observable targets that can be achieved quickly and maintained easily when it is difficult to achieve more functional targets. For example, it is much easier to create a new law or agency than to ensure that a government agency properly implements the law to meet policy goals. Under pressure to demonstrate results quickly, recipient countries may outwardly adopt the institutional features proposed by donors, but not actually take advantage of these outward reforms to alter power relationships in ways that improve functionality of formal institutions. As Matt Andrews writes, “many reforms are introduced as short-term signals that ensure developing countries attain and retain external support and legitimacy” (Andrews, 2013, KL. 144). But, the focus on the appearance or existence of formal institutions can “crowd out efforts to introduce reforms with less-demanding, more fitted content; that could actually enhance functionality” (Andrews, 2013, KL. 2208). To this point, however, research has not delineated and tested the sources of these signaling pressures and their relationship to the results agenda.

We specifically argue that aid recipients and project staff face a strategic choice when selecting institutional targets. Targets that measure “form,” whether or not a given organization or institution is present, are often easier to achieve and maintain than targets that measure “function,” whether or not a given institution performs its duties adequately (Andrews and Pritchett 2012, p. 7). Although the payoff is often higher for achieving targets that measure function, the recipient countries most in need of institutional reform have a low probability of achieving and maintaining such targets. To receive any payoff and to maintain access to foreign assistance, these countries choose easier, form-based institutional goals that make questionable contributions to the functionality of formal institutions. Thus, the countries most in need of institutional reform are least likely to benefit under the “results agenda.”

To test this theory, we examine institutional developments related to environmental management supported by the World Bank, an area of lending that is perhaps most often driven by the preferences of donor countries (Keohane and Levy 1996; Heltberg and Nielsen 2003; Hicks et al. 2008). By examining targets of institutional development related to environmental management — from establishing new pollution control laws, to conducting regular pollution monitoring research, to decreasing actual levels of pollution — we focus on targets that are likely to have a strong signaling dynamic between project teams and their recipient country counterparts and

member states to the World Bank. Additionally, most targets in the environmental management sector relate to institution building, which makes our study empirically tractable.

We develop an original dataset by extracting observable indicators of institutional development related to environmental management from World Bank operational and Independent Evaluation Group post-project evaluations. We code the indicators on whether they measure a form, activity, or function of the relevant government agency. We then remeasure those indicators several years after projects have closed by eclectically sourcing from current documentation. Using this dataset, we find that “form” targets are both easier to achieve and easier to maintain than “function” targets, confirming the viability of a strategic choice of targets for signaling purposes. We find that recipient countries which have a lower probability of achieving functional targets or a higher dependence on development assistance are more likely to select form-based targets. These results support our theory and help diagnose why the “results agenda” has produced few results for the countries most in need of institutional reforms.

### **Institutional Targets in Development Assistance**

Since the 1980s, development agencies have increasingly focused on public sector reform as a key element of development. One way to measure the increase of these externally sponsored reform activities is to look at the number of World Bank projects classified in the public administration, law, and justice sector (Mohoney and Thelen 2009). Of projects before 1980, less than 1% were classified as targeting this sector. By the 1980s, that number had risen to 20% and between 2000 and 2010 65% of projects had some public administration, law, or justice component (Andrews 2013, KL. 178-203). Over 140 countries have at least one World Bank project targeting public sector reform and administration (Andrews 2013, KL. 376).

This focus comes from the idea that formal institutions set the incentives and constraints for all actors in a country and therefore are crucial in shaping development outcomes. However, public sector reform programs have often fallen short of expectations. Several previous World Bank studies on the effectiveness of public sector reforms have used the World Bank Country Policy and Institutional Assessment (CPIA), which ranks countries on their economic

management, structural policies, policies for social inclusion and equality, and public sector management and institutions. A 2008 study on World Bank interventions between 1999 and 2005 found that around two-thirds of countries improved their overall governance scores after implementing World Bank-sponsored reform activities, although only 40% of countries improved in quality of administration and only 50% improved in transparency and accountability (World Bank 2008). A 2011 update showed less encouraging results. Of the 80 countries receiving World Bank support for public administration reform, less than 40% showed an increase in CPIA governance scores between 2007 and 2009. Over a quarter saw their scores decline (World Bank 2011).

Using the World Governance Indicators (WGI), a report by the Asian Development Bank finds that despite significant donor support for institutional reforms, “almost across the board” there has been a decrease for Asian countries in ratings of government effectiveness and regulatory quality (Stone 2008, p. 11-13). In his recent book, Matt Andrews compares government effectiveness scores from the WGI dataset between 1998 and 2008 for the 145 countries that received support for institutional reforms in that time period. He finds that of the countries that received support, 73 experienced an increase in government effectiveness scores while 72 experienced a decrease in government effectiveness scores (Andrews 2013, KL. 376-381).

Broad-based measures are imperfect gauges of institutional quality, but taken together this evidence suggests that support for institutional improvement in developing countries has produced mixed results at best. In fact, countries seem to be facing similar problems to the ones they faced before an inflow of donor resources. As a description of the issues facing public sector reform efforts, a 2011 IEG report concluded, “[c]hallenges that continue to undermine the functioning of the public sector include poor budget planning and execution, poorly motivated civil service cadres, legacies of complex or irrational bureaucratic structures, corruption, political resistance to reform, fiscal constraints, the use of public sector employment as a safety net for the job market, and unfair or unavailable judicial services.”

## *The pernicious effects of form-based institution building in aid*

How can this be, when donor agencies are increasingly focused on promoting strong formal institutions as a core objective? One explanation put forward for the mixed performance of institutional development projects has been the agenda of donor organizations to finance “best practice” reforms in developing countries— for example, the creation of anti-corruption commissions or one-stop shops for business registration (World Bank 2008a; Arruñada 2007, 2012; Andrews 2013). Among economists, political scientists, and organizational sociologists, there is growing skepticism about the ability of aid agencies and international organizations to easily export models of institutional development of the public sector from one country to another (Haggard et al. 2008; Grindle 2004; Evans 2004; Rodrik 2000; Andrews 2009, 2010; Booth 2010). Development agencies themselves have begun to question the wisdom of promoting “best practice” reforms (Grindle 2011). By way of illustration, consider a 2008 evaluation of World Bank-sponsored public sector reform programs writ large: “[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).

The focus on “best practice” reforms that are based on institutional forms can crowd out solutions that are more tailored to fit the specific context of the recipient country. In the last decade, scholars have been increasingly concerned that the quality of service delivery is being conflated with the form of bureaucracies in many contexts (Pritchett & Woolcock 2004). In fact, the focus on observable, form-based targets can crowd out other priorities in development planning (Piotrowski 2003). Under pressure from donors, recipient countries might engage in isomorphic mimicry, which Pritchett and colleagues (2013) define as “adopting the camouflage of organisational forms that are deemed successful elsewhere to hide their actual dysfunction.” Many times, the institutions that foster broad-based development and the delivery of public services must be unique to the social, economic, and political context where they are established. There are often multiple pathways to the same functional goals for developing countries (Rodrik 2007).

This is one reason why a focus on “best practices” can be harmful. The reforms that work

in one country might not succeed in other countries. In their push for “institutional monocropping,” donors can ask for “blueprint” style institutions that have worked elsewhere, perhaps even very successfully, without fully considering the possibility that the institutions are not appropriate to new conditions (Evans 2004). Donor organizations face considerable uncertainty, making “blueprint” approaches to institutional development, well-meaning as they are, a limitation on flexible choices focusing on functional outcomes that could prove more successful.

Isomorphic mimicry is pervasive in countries that rely heavily upon foreign assistance from bilateral and multilateral donors (Pritchett et al. 2010; Andrews 2013). Jordan Smith (2003: 711) provides a detailed anthropological account of how isomorphic mimicry works in a family planning program managed by UNFPA and Nigeria’s Ministry of Health:

“Nigerians who work in the Ministry of Health have become adept at adjusting to and accommodating donor expectations, even as they manage to address local, often competing, priorities. To assure that they are able to continue to reap and distribute the benefits that accrue from managing internationally-funded population programs, Nigerian health personnel must stay attuned to the seemingly endless shifts in donor policies and approaches. They change their vocabularies, adjust their programs, and rework their reports, all the time preserving their ability to build and feed the patronage networks that are essential for survival and success in Nigeria.”

Similarly in Mauritania, Samuel (2013) finds that formal practices were closely followed, enabling international organizations and donor agencies to declare victory, but compliance with form indicators masked a set of informal practices that contributed to widespread institutional dysfunction. Quantitative target systems applied to public agencies, when combined with strong reputation or resource-based incentives, can induce agencies to find cheap and easy ways to show achievement on chosen indicators, perhaps at the expense of deeper functionality (Hood 2006; Arruñada 2007; Cullen and Randall 2006; Jacob 2005; Wynia et al. 2000).

The alternative is context-specific projects that focus on improving institutional function, measured by the results of service delivery (Andrews et al. 2012). Much of the research on more functional reforms deals with small, incremental steps that take into account context-specific factors (Evans 2004). The problem with this approach to reform is that staff in donor

organizations face considerable uncertainty about the outcomes of tailored or discretionary programs, especially when there are accepted models in the profession about improving functionality of government agencies. Donor agencies are often large, working in many countries around the world simultaneously. Thus, there are efficiencies to be gained by some standardization in process (Shirley 2005). However, these pressures do not always, or even frequently, result in better development outcomes for recipient countries. The problem of uncertainty plagues “best practice” reforms, just like other more context-specific reforms that focus on results, though it might always not be recognized as such (Krause 2013).

Improving functional targets through institution building is far more difficult than adopting institutional forms. It requires an “incremental search for solutions to problems that local agents care about” (Andrews, 2013, KL. 2268). Previous research suggests that fundamentally changing the domestic political status quo usually requires that the executive branch has the policy space to experiment with and iteratively adapt to local constraints and opportunities (DFID 2003; Grindle 2004; Rodrik 2007; Pritchett et al. 2012; Andrews 2013).

### *The strategic logic of target selection at the World Bank*

The “results agenda” increases pressure on recipient governments and staff at development agencies to achieve measurable and often quantifiable targets. Form-based institutional targets should be easiest to achieve and maintain. But, sometimes countries do attempt more difficult functional reforms with development assistance. Before shifting to the question of what countries choose to undertake functional reforms, we need to ensure that form-based targets are indeed easier to achieve and maintain. If correct, this creates the conditions necessary for a strategic choice about targets within donor organizations and recipient countries.

To investigate differences in the difficulty of achieving and maintaining targets, we code institutional targets in World Bank projects related to environmental management, an issue area that is greatly focused on improving institutional performance. We assign the observable institutional targets in World Bank projects to either “form,” “action,” or “function” categories according to the following rules.



**Form:** the target records that an agency 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.

**Activity:** 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 monitoring an environmental attribute on a regular basis.

**Function:** the target records the results of institutional development on the relevant environmental attribute. Examples include reducing the percentage of unaccounted water or the number of days with severe air pollution.

Our expectation, according to the research we have reviewed, is that targets that can be maintained through inertia alone (e.g., a law exists, task force exists, etc.) are easier to achieve and maintain than targets that require continued action and choice over time. In turn, we expect that targets that measure the cumulative results of action are more difficult to achieve and maintain than targets that measure actions alone. If correct, then the ability of recipient governments and World Bank staff to act strategically with respect to target selection is strong. More formally, we have the following hypothesis:

**H1: Function targets are less likely to be achieved and persist than both action and form targets.**

#### *Benefits of achieving functional targets*

If this hypothesis is correct, then we can proceed to investigate the conditions under which different types of targets are chosen for strategic reasons. The first question to be answered is: if form-based targets are easier to achieve and maintain, why do some countries select functional targets? This requires information on how project targets are chosen. There is no formal process

that dictates which targets are chosen to measure the success of the institution building components of a World Bank project. These are selected through negotiations between project teams at the World Bank and staff in line ministries in the borrowing countries. As such, both World Bank staff and officials in recipient countries have the opportunity to influence the types of targets that are selected to measure project success. This means that we must explain the joint incentives of recipient officials and World Bank staff who prepare projects for approval.

For both World Bank staff and recipient officials, there are a range of benefits for achieving functional targets that exceed the benefits of achieving form-based targets. 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 both environmental and non-environmental function targets (ADB 2005; World Bank 2010b). IDA's performance-based allocation system uses cross-country measures of need, project performance, and policy and institutional performance to determine the overall resource envelope that will be made available to countries. Thus, improvements on environmental functional targets — at the project level or the country level — result in access to more World Bank funding for IDA countries.

Once an overall resource envelope has been established, the World Bank can provide funding through modalities that provide more or less discretion to counterpart governments. Programmatic funding through so-called development policy loans (DPLs) and results-based financing is significantly more discretionary than funding through other mechanisms like project support and technical assistance, but programmatic disbursements are usually tied to the counterpart government's performance vis-a-vis functional targets in the environmental sector, such as whether a government ministry is systematically reviewing environmental impact assessments of major infrastructure projects or paying titleholders of forests for the protection of sensitive watersheds (World Bank, 2008b, 2010; World Bank Carbon Finance Unit 2013).

The 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 very similar if not identical criteria to both establish the size of country-specific

resource envelopes and determine how discretionary funding is allocated (ADB 2005, 2007; GEF 2013a; European Commission 2008, 2014). This creates an additional incentive for developing country governments to improve their performance on both environmental and non-environmental functional targets; however, improved performance on these targets does not necessarily result in more *environmental* assistance. The Global Environment Facility is an exception. It only provides environmental assistance and through its so-called its System for Transparent Allocation of Resources<sup>2</sup> countries receive more or less assistance based on their performance on functional targets in the environmental sector (GEF 2013b).

Additionally, developing country governments are able to capture important reputational benefits from improving their performance on functional targets in the environmental sector. Consider the “Joint Staff Advisory Notes” produced by the World Bank and IMF. These assessments not only evaluate progress vis-a-vis specific functional targets established in a country’s poverty reduction strategy, but they are also widely used by donors who supply (highly discretionary) budget support to determine whether and to what extent recipient countries are meeting environmental and non-environmental functional targets. As such, World Bank-IMF Joint Staff Advisory Notes can indirectly influence the amount of discretionary revenue that a developing country government is able to access from a range of donors, including DFID, the European Commission, and several Northern European bilateral donors.

For recipient governments, domestic constituencies are likely to care little about projects and programs that change institutional forms without improving the delivery of public services or the management of environmental resources. Research finds that competitive democracies, where officials do not have a monopoly on political power, are more likely to expend resources on the delivery of frontline public services (Lake and Baum 2001; Stasavage 2005). Ross (2006) challenges whether this spending always results in better outcomes for the poor, arguing instead that increased spending on public services focused on politically powerful constituencies of middle and higher income voters. Nonetheless, the consistent conclusion is that spending on frontline public services is driven in part by domestic constituencies. In contrast, it is not clear that public

---

<sup>2</sup> STAR replaced the GEF’s previous performance-based resource allocation framework, which was adopted in 2005.

support can be generated by adopting institutional forms preferred by donor organizations.

Finally, it is worth considering the incentives of the individuals on both the recipient government side and the World Bank side on choosing projects. Neither officials or World Bank staff want to fail to meet targets. For counterpart officials in borrowing countries in particular, having a reputation of being sympathetic with donor policy preferences and able to achieve functional targets can generate more attention and support from donor institutions (Chwieroth 2013). These working relationships can promote access to additional sector level financial revenues and may increase later access to job opportunities at international organizations.

### *Choice of function or form targets*

While there are benefits to achieving targets associated with functional goals that are not available when achieving form-based goals, there are also risks involved with choosing these more ambitious targets. In more formal terms, we expect the payoff for achieving and maintaining functional targets to be greater than the payoff for achieving and maintaining form-based targets, which is greater than the payoffs for failing to achieve and maintain either kind of target. The consideration then is whether the risk-adjusted payoff of choosing functional targets is greater than the risk-adjusted payoff of choosing a form-based target, since the failure to achieve either is likely to be similarly negative. Following this logic, a recipient country will choose a functional target when:

$$P_{\text{func}} * B_{\text{func}} > P_{\text{form}} * B_{\text{form}}$$

This choice generates a number of predictions about the kinds of recipient countries that will choose to focus on easy, but shallow form-targets versus difficult, but deeper functional targets. Borrower governments vary significantly in terms of their ability to successfully execute institutional development programs (IEG 2008; Denizer et al. 2011). As such, one would expect utility-maximizing governments with low levels of implementation capacity to press for more form / action indicators than functional indicators, as projects that define success as de jure improvements in institutional performance will be easier to implement. The donor agency and development bank counterparts of borrower governments also have professional incentives to set

low standards for success in countries that lack significant implementation capacity. Thus, poor previous project ratings can also serve as an indicator of low implementation capacity, which we expect will encourage the selection of “form” indicators. In terms of the probability of achieving targets, recipient countries that already have strong formal institutions and that are able to successfully implement World Bank projects should have a lower difference in the probability of achieving functional targets as compared to form-based targets.

**H2: Recipient countries with higher pre-existing levels of governance are more likely to choose functional targets than recipients with lower levels of governance.**

In terms of the benefits of achieving these different types of targets, we also expect differences between different types of recipients. Resource flows from aid donors can be extremely important for recipient governments, since they make government budgets more fungible (Whitfield and Fraser 2010) and help to maintain access to monetary support (Samuel 2013). For recipient countries without access to other sources of revenue and who are more sensitive to shocks in aid flows over time, the benefits of successfully maintaining access to aid is likely to take precedence. These countries are likely to be risk-averse and seek to maximize the minimum payoff, resulting in more stable flows by choosing form-based targets. For recipient countries, pressures to achieve targets are pronounced when aid flows are an essential part of government revenue. As a consequence of these pressures, we expect political leaders who have fewer alternative sources of revenue to pressure their ministries into negotiating easier institutional targets, so that they maintain flows of aid and are less susceptible to the often wide fluctuations that come with results frameworks.

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

But, should not staff at the World Bank, in protection of the results agenda, seek to deflect the choice of “easy” targets? For staff who are directly involved with the negotiation of project targets, the incentives do not push in this direction. In fact, staff at the World Bank are rewarded for being involved with successful projects and punished for being involved with unsuccessful projects, in terms of salary, promotion, and internal prestige (IEG 2010; Whittle 2013). Under these conditions, it might make sense to choose ambitious targets — but, only with countries that are more likely to achieve them successfully. To judge which countries are likely to achieve more results driven indicators within the life of the project, we expect that staff at organizations like the World Bank will rely on past project success in the country as a guide, potentially even looking to previous project ratings (Buntaine, 2011). When recipient countries have not been successful at previous targets, staff who are rotated in to work in that sector should be under more pressure to choose easily achievable targets. These staff are under pressure to have new operations approved quickly, which will reinforce the incentives described in the decision-theoretic framework.

### **Measuring the choice, achievement, and persistence of institutional targets**

#### *Building institutions for environmental management*

In order to test these theoretical expectations, we turn to World Bank projects in the environmental management sector. 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 generally agree that the management of environmental issues is rarely successful in the absence of strong formal institutions (VanDeveer and Dabelko 2001; Weidner 2002). As a consequence, aid agencies have spent a greater deal of their resources in the sector building the capacity of state agencies to carry out the planning, monitoring, and enforcement involved with environmental management. However, institution building in the environment sector is perhaps an area of aid portfolios most driven by the concerns of donor states and aid agencies (Lewis 2003, Marcoux et al. 2011). Thus, it offers a set of projects that should yield the signaling dynamics that we theorize above.

To test our hypotheses, we created an original dataset on the achievement and persistence of indicators used by the World Bank to measure project success in post-project evaluations. First, we identified 250 World Bank projects completed between 2003 and 2009 that had more than 10% of their financing allocated 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 levels of these indicators. In cases where both baseline and completion data are available, around 85% of indicators, this enables us to compute an ordinal measure about whether progress was made on achieving the target during the project.

To measure the persistence of these targets, we developed an eclectic sourcing methodology and compiled a new measure from official documents or news sources at the beginning of 2013. We turned to a variety of publicly available news sources, such as LexisNexis and Google News. In addition, many government agencies issue annual reports on staffing and appropriations, allowing us to resample some indicators at present. Government websites and NGO reports also proved to be valuable resources. In all cases, we recorded the source and publication date of the re-sampled indicators. In total we identified 826 indicators, around 3.3 per project, and were able to remeasure post-completion data for 347 of them, a success rate of around 42%.

After collecting indicators, each one was coded for whether it measured a form, activity, or function according to the definitions outlined in the previous section. Each indicator was coded independently by three research assistants. For 72% of indicators, all three coders were in agreement. For the disputed codes, one of the co-authors arbitrated. 95% of the time either all the coders were in agreement or the arbitrator agreed with the majority of the coders. For an example of indicators with form, action, and function codings, see Table 1.

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

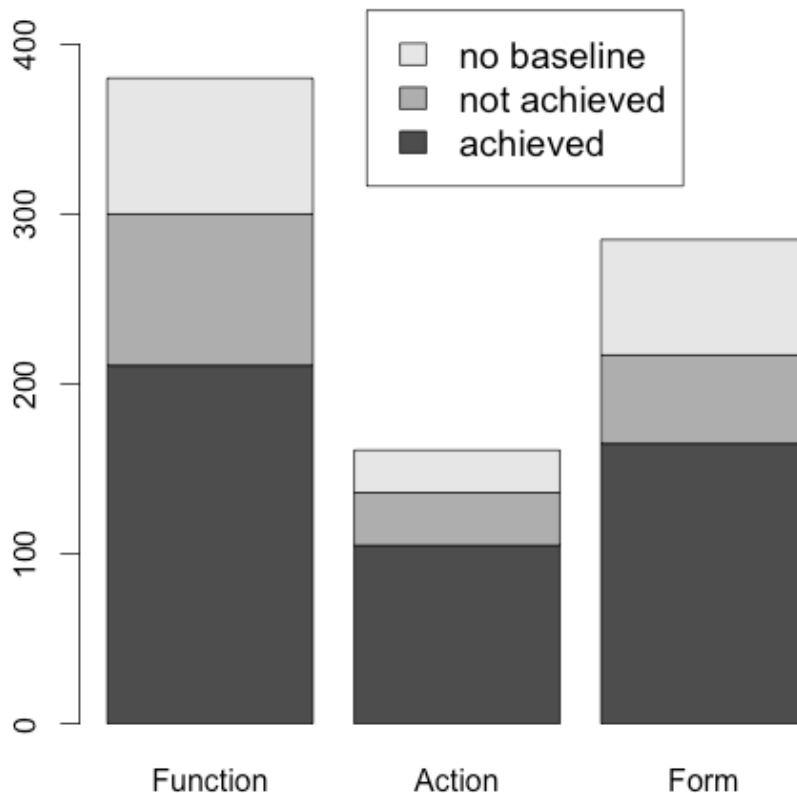
<b>Country</b>	<b>Completion Year</b>	<b>Project Name</b>	<b>Indicator</b>	<b>Form, Action, or Result?</b>
Argentina	2007	Native Forests & Protected Areas Project	Does Argentina regularly update its national inventory of forest resources?	Action
Ghana	2008	Natural Resource Management Project	Does the 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



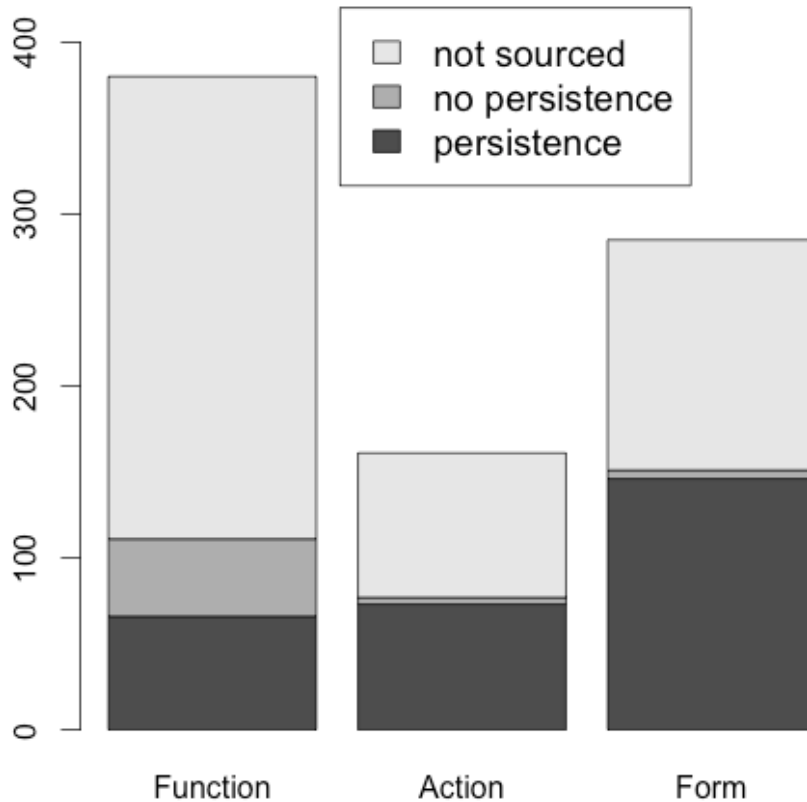
## Can targets be a strategic choice?

In order to demonstrate that the choice of form versus result indicators can be strategic, we must first demonstrate that form-based targets are easier to achieve and maintain than other types of targets that require either actions or results. As shown in Figure 1 and 2, the descriptive statistics make an initial impression that this strategic choice is available. For targets that had baseline data available in the evaluation document, 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 1). The potential for countries to maintain targets following project completion is more clear. Whereas only 59% of functional targets were maintained following project completion, 95% of action targets and 97% of form targets were maintained following project completion (Figure 2). It appears that functional targets are harder to achieve and maintain than action and form targets, consistent with Hypothesis 1.

*Figure 1. Descriptive data on achievement of different types of targets*



*Figure 2. Descriptive data on persistence of different types of targets*



But, these descriptive statistics run the risk of 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 developments or the missingness of data. In this analysis, we are forced to make the assumption that the outcome variables are missing at random (MAR), which means that missingness does not depend on the value of  $y_i$  after conditioning 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 remeasurement 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 the situation, since multiple imputation

analysis and complete case analysis converge when missingness is only in the outcome variable (Snijders and Bosker, 2012).

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. In this way, we seek to ensure that it is easier to achieve and maintain form targets versus functional targets across different types of countries, projects, and size of implementing agencies. 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 models on the achievement of institutional targets and 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. In a second model for both achievement and persistence, we add a dummy variable that indicates that borrower performance was satisfactory during project implementation, as rated by the World Bank Independent Evaluation Group. In our model of persistence, we additionally add a variable for the number of years between completion and our remeasurement data, since targets achieved long ago might be more likely to be overturned and the gap to remeasurement is not necessarily a project-specific characteristic. Finally, in our second model of persistence, we also include a dummy variable for whether the target was achieved during implementation, since this might correspond to borrower commitment or capacity, but again is not necessarily a project-specific effect. The results of our analysis are summarized in Table 2. In general, they confirm the hypothesis that form and action indicators are more likely to be achieved and to persist beyond the life of a project than are function indicators.

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

<b>DV:</b>	<b>Achievement</b>	<b>Achievement</b>	<b>Persistence</b>	<b>Persistence</b>
<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

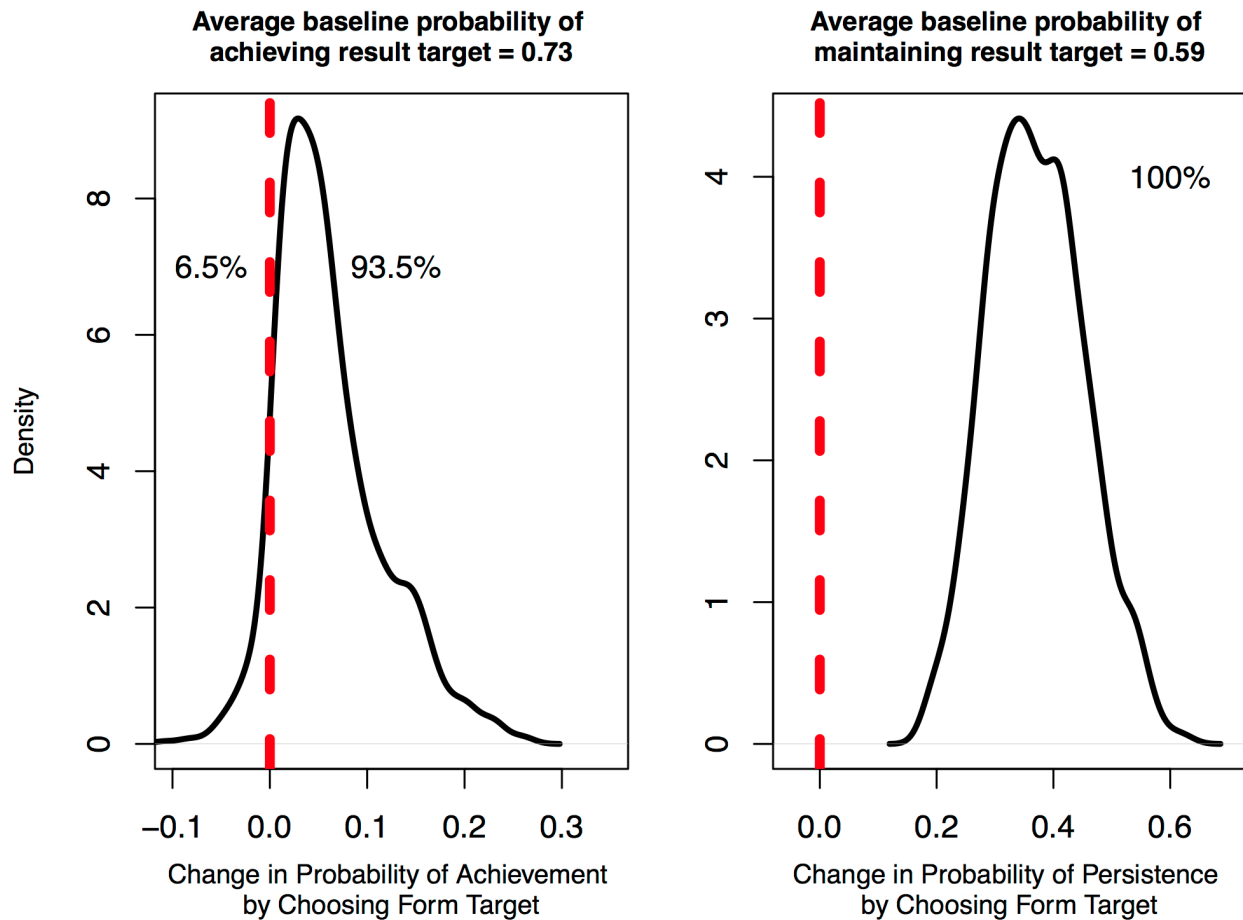
Model cells list: Parameter estimate; (Standard Error); [p-value of one-sided z-test]

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 results target for the hypothetical average project and country based on our models (Figure 3). The models estimate, but did not find any implementation level effects. We take 1000 draws from the distributions of model coefficients, including the average random effect levels for project and country and then vary only whether the hypothetical target measured a result 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

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



### Political Economy of Setting Institutional Targets in World Bank Projects

The results from the previous section indicate that the choice of targets is a strategic option for development staff and recipient governments alike. But, under what conditions are staff and governments likely to avail themselves of this strategic option? As we argue above, we should be

more likely to observe the choice of function indicators for countries that have a higher probability of achieving these targets (H2). Additionally, we should observe countries that are less dependent on development assistance and have other sources of government revenue to choose functional targets (H3).

To measure the ability of recipient countries to achieve functional targets, we use two different operationalizations. First, we include a binary variable that is positive whenever at least 75% of the recipient's post-project evaluations in the previous five years are rated successful, as determined by the independent evaluation unit at the World Bank (H2a). Second, we include the Worldwide Governance Indicators variable for government effectiveness, which measure the ability of government agencies to successfully implement policies (H2b).

To measure the dependence of the recipient countries on foreign assistance, we use three operationalizations. First, we use the project-level proportion of funding from the World Bank's International Bank for Reconstruction and Development versus the International Development Association (H3a). Since the IBRD provides loans at commercial interest rates, it is more driven by the demand of borrowers and it is not subject to the same performance-based allocation standards. We expect that borrowing countries that use IBRD will have less of a need to signal success to donor states and the World Bank because of access to commercial markets. Second, we use GDP per capita in the year prior to approval, since poor countries are more likely to be dependent on international donors than countries that have higher levels of income (H3b). Additionally, countries with higher per capita income are likely to enjoy a larger tax base, which decreases dependence on foreign assistance. Finally, we look at the share of GDP made up by natural resource rents for each country in our sample, since countries that have a high income from natural resources can use that money as an alternative to aid. We employ a variable developed by Hamilton and Clemens (1999) that measures natural resource rent revenue as a percentage of national income. This indicator sums all rents from 14 fuel and nonfuel mineral resources.

Like the model specification in the previous section, we want to ensure that any effect of expected project performance and aid dependence on the choice of indicators is consistent across different types of countries and different levels of target implementation. We therefore again specify a random effects model where the intercept varies by country and level of implementation.

In addition to these varying-intercepts, we also control for variables that are likely to be independent from country and level factors in some of the model specifications. First, because the results agenda and the use of logical results frameworks has grown more common over time at the World Bank, we include a linear year count variable since the first year approval year in our sample to account for variation across time 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 preferred less and implemented with less success than projects that are a mix of development and environmental objectives (Buntaine and Parks 2013). Thus, we expect that project sent to the Environment Sector Board at the World Bank will be less likely to include functional targets, as borrowers will want targets with lower implementation and monitoring costs.<sup>3</sup> Descriptive data on these variables is displayed in Table 3.

---

<sup>3</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.

**Table 3. Descriptive data on variables in target choice model**

	<b>Mean</b>	<b>Range</b>	<b>Standard Deviation</b>	<b>Obs</b>
<b>DV: Result Target</b>	0.46	0 - 1	0.50	826
<b>H2a: Successful Evaluation Record</b> (At least 75% satisfactory AY-1 to AY-5)	0.28	0 - 1	0.45	826
<b>H2b: Government Effectiveness (AY-1)</b>	-0.36	-1.50 - 0.67	0.44	370
<b>H3a: IBRD proportion</b>	0.53	0 - 1	0.48	793
<b>H3b: GDP per capita (\$1k @ AY-1)</b>	1.82	0.13 - 10.5	1.92	813
<b>H3c: Resource Rents / GDP (% @ AY-1)</b>	6.79	0 - 56.5	9.92	809
Approval Year (centered linear)	8.7	0 - 18	3.36	825
Environment Sector Board	0.19	0 - 1	0.39	822

Our model results largely support our hypotheses that countries with options for non-aid revenue and countries with records of good project performance are more likely to choose targets that measure results, rather than the form of institutions (Table 4). We find that countries that have a record of achieving satisfactory results in at least 75% of their projects during the five years prior to project approval are more likely to choose function targets. 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, staff at the World Bank who work on these projects should be more likely to feel confident in choosing ambitious targets.

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 functional targets, these variables are highly correlated.



This means that we cannot completely disentangle whether higher values of these variables increase the probability or decrease the spread in benefits of achieving a functional target. Thus, we include them in the core model one at a time, to demonstrate the robust nature of our results. All of these variables have a positive effect on the probability that the country will choose a functional target, which supports our hypotheses. Across all of our model specifications, we find that countries with higher amounts of natural resource rents as a proportion of GDP are also more likely to choose result targets. We interpret this result to mean that countries with less dependence on aid donors do not need to create as many positive signals for donor audiences.

**Table 4. Choice of functional institutional development targets (at project approval)**

<b>DV:</b>	<b>“Function” Target</b>	<b>“Function” Target</b>	<b>“Function” Target</b>
<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</b> (\$1k @ AY-1)			<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

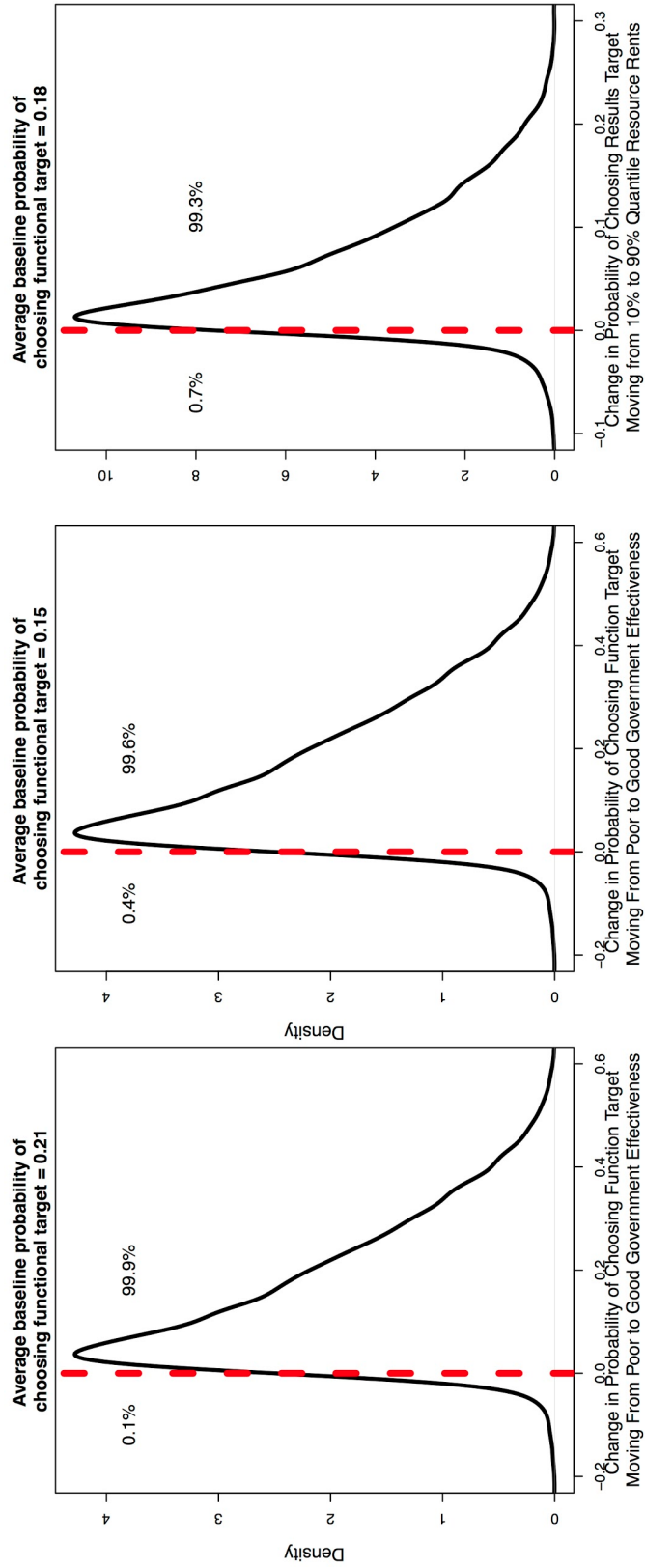
Model cells list: Parameter estimate; (Standard Error); [p-value of one-sided z-test]

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 functional 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. The results show that several of our predictor variables have an important influence on the probability that a country will select targets that measure function. For the average project allocated to a country without a satisfactory evaluation record, the probability of choosing a functional target would increase 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 functional 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 rents in the sample, the probability of choosing a functional target increases from 19% to 24% when rising to the 90% quantile in rents. 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 functional targets. These are precisely the countries that are least in need of institutional reforms, which helps explain why results have been so disappointing for so many of the world's neediest countries. These countries instead tend to focus on easy, shallow targets.

Figure 4. Effect of predictor variables on choosing results 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. These variables did not predict the choice of indicators and did not otherwise qualitatively influence the results of our models. We expected that a strong environmental NGO community in the borrower country would increase the likelihood that functional indicators are chosen. Previous research suggests that civil society oversight and pressure helps facilitate monitoring of institutional persistence or deterioration (Diez 2012). The civic sector often exerts pressure for governments to honor their international commitments. Environmental NGOs also help to keep reforms necessary for long-run institutional development in the public eye (Dai 2007; Raustiala 1997). In a specification not reported here, the number of environmental NGOs from various editions of the *Environment Encyclopedia and Directory*, did not predict the choice of results indicators and caused ~50% casewise deletion because of missing data.

## **Discussion and Implications**

The first section of our empirical analysis offers evidence that form indicators are easier targets to achieve and maintain. In our sample, form indicators were both significantly more likely than function indicators to improve over the life of World Bank projects and more likely to be maintained beyond the life of a project. Indicators measuring whether an action was taking place were not significantly more likely to improve than function indicators, but were significantly more likely to persist. This provides strong support for the contention that selection of targets is a strategic option for both recipient governments and donor staff.

The second section of empirical analysis provides evidence that this strategic choice is being used to deflect pressures from the results agenda. In particular, we show that countries with a lower probability of achieving functional targets and a higher dependence on foreign assistance are more likely to select shallow, form-based targets. We find that good previous project ratings were positively correlated with the selection of functional indicators. Because project managers in countries with poor track records would be more likely to select easier targets, we interpret this as evidence that pressure to select form indicators rather than function indicators occurs with the

donor as well. Staff at the World Bank are promoted on the basis of managing projects so that disbursements are not disrupted by performance problems. Under these pressures, they can be more confident about choosing deeper, functional targets when performance by the recipient counterpart has been satisfactory. We find a similar effect for an index of government effectiveness.

Related to the recipient's need for development assistance, we find that the percentage of funding coming from IRBD, which lends at a commercial rate rather than at a concessional interest rate, is more often associated with the choice of function indicators. In addition, both natural resource rents and GDP per capita were positively correlated with the selection of functional indicators. Overall, we provide evidence that donors and recipients have strategically selected indicators in our sample of projects that seek to build environmental institutions. This suggests that a push towards more measurable and objective outcomes can contribute to the adoption of shallow reforms, which can help explain the disappointing results of institution building and public sector reform programs in recent decades. While outside the scope of this project, these conclusions can be bolstered by looking both at different sectors and at different development agencies.

## **Conclusion**

Previous research suggests that a focus on institutional form rather than function is at least partially responsible for the mixed success of institutional development projects. The temptation to rely on form-based measures of success is strong for donor agency staff and developing country officials alike because form indicators are significantly easier to achieve and sustain beyond the life of a project than functional indicators. We extend research on this theory by examining the strategic logic behind the selection of targets. We examine the circumstances under which donors and recipient countries have incentives to focus on function. We find a stronger focus on functional targets in cases where the recipient country has access to sources of revenue other than aid. We observe a similar pattern in cases where the recipient country already has a track record of successfully implementing externally-financed projects. Overall, these findings suggest that the selection of indicators is a strategic decision made by donor agency staff and their developing

country counterparts with potentially far-reaching consequences for the effectiveness of institutional development efforts.

The convention wisdom also suggests that building stronger monitoring procedures for aid projects will improve success and increase development impacts. Indeed, rates of project success at the World Bank have increased since the 1980s. However, recent scholarship has called into question the veracity of this trend (Sud and Olmstead-Rumsey, 2012). Adding support to this skepticism, the results we present suggest that it is not enough to measure targets; improved monitoring and evaluation must measure functional results. Although there has been more work at the World Bank in recent years to produce core indicators to measure projects success, including for the environment sector, many of these indicators still measure forms and actions that are several steps removed from results. The choice of different indicators, even within organizations that prioritize accountability for measurable results, can give rise to the strategic logic that we explain here. The recipient countries that are most in need of institutional development because of poor governance are also likely to be the countries most dependent on aid donors and thus most likely to select, easy form-based targets.

Apart from the multivariate analysis that we offer, we introduce a new method for evaluating the long-run institutional development impact of aid projects. Donor agencies and development banks are urgently in need of better methods to track the durability of the institutional development gains 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. 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. As such, donor organizations that seek to understand their post-project impacts are forced to guess about the conditions that support sustainability at project completion.

We have attempted to overcome this knowledge barrier by 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 development

effects 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 (Besley and Persson 2010; World Bank 2011; Acemoglu and Robinson 2012). Broad indices of institutional quality are analyzed 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 typical, country-level indicators of institutional development, we are better able to test a theory about how pressure to measure results impacts the choices recipient government and World Bank staff make about measuring results.

What should donor agencies do in light of incentives to adopt form-based targets that are easy to achieve and maintain? Most importantly, as the set of core indicators used to measure the success of development projects is codified and expanded, indicators that measure function, but leave multiple pathways open to achieve functionality should be prioritized. For example, the World Bank environment sector core indicators include several items like “area under improved management” for natural resource management, without actually measuring what improved management should accomplish. In general, donors should prefer customized indicators of function to those that can be standardized across projects and are likely to measure form. The results agenda may indeed promote better performance as measured by evaluators, but these gains are not likely to translate to institutional impacts until targets measure functionality.



## 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, 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.
- Andrews, Matthew. 2013. *The Limits of Institutional Reform in Development*. Cambridge: Cambridge University Press.
- 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 (02): 341-365.
- Arruñada, Benito, 2007. Pitfalls to avoid when measuring institutions: Is Doing Business damaging business? *Journal of Comparative Economics* 35(4): 729-747.
- Arruñada, Benito. 2012. *Institutional Foundations of Impersonal Exchange: The Theory and Policy of Contractual Registries*. Chicago, IL: University of Chicago Press.
- 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.
- 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.
- 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.
- Besley, T., and T. Persson. 2010. State Capacity, Conflict, and Development. *Econometrica* 78(1): 1-34.
- Buntaine, M. T. 2011. Does the Asian Development Bank Respond to Past Environmental Performance when Allocating Environmentally Risky Financing?. *World Development*, 39(3), 336-350.
- 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.
- Burki, Shahid Javed and Guillermo Perry. 1998. Beyond the Washington Consensus: Institutions Matter. Washington D.C.: World Bank.
- Callaghy, T. M. (1988). The State and the development of capitalism in Africa: Theoretical, historical and comparative reflections. *The Precarious Balance. State and Society in Africa*, edited by Donald Rothchild, and Naomi Chazan. Westview, 67-99.
- Chauvet, Lisa and Paul Collier. 2008. Aid and reform in failing states. *Asian Pacific Economic Literature*. 22(1): 15-24.
- Chwieroth, Jeffrey M. 2007. Neoliberal Economists and Capital Account Liberalization in Emerging Markets. *International Organization* 61 (2): 443-463.
- Chwieroth, J. M. 2013. "The silent revolution." How the staff exercise informal governance over IMF lending. *The review of international organizations*, 1-26.
- Clemens, Michael, Charles Kenny, and Todd Moss. 2007. The Trouble with the MDGs: Confronting Expectations of Aid and Development Success. *World Development*. Vol. 35, No. 5, pp. 735–751.
- 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*, edited by T. Gronberg and D. Jansen. Amsterdam: Elsevier Science.
- Dai, X. 2007. *International institutions and national policies*. Cambridge University Press.
- Denizer, C., D. Kaufmann, and A. 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.
- 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.
- 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.
- 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* Winter 2010: 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. Accessed on 12 January 2014 at  
<http://www.thegef.org/gef/sites/thegef.org/files/EO/TD1%20-%20Comparison%20of%20STAR%20to%20other%20allocation%20systems.pdf>
- Global Environment Facility (GEF). 2013b. System for Transparent Allocation of Resources. POLICY: PL/RA/01. Washington D.C.: GEF. Accessed on 12 January 2014 at  
[http://www.thegef.org/gef/sites/thegef.org/files/documents/document/PL.RA\\_.01.System\\_for\\_Transparent\\_Allocation\\_of\\_Resources.031413.pdf](http://www.thegef.org/gef/sites/thegef.org/files/documents/document/PL.RA_.01.System_for_Transparent_Allocation_of_Resources.031413.pdf)

- Gelos, R. Gaston, Ratna Sahay, and Guido Sandleris. 2011. Sovereign borrowing by developing countries: What determines market access? *Journal of International Economics* 83 (2): 243-254.
- 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. 2004. Good Enough Governance: Poverty Reduction And Reform In Developing Countries. *Governance: An International Journal Of Policy, Administration, And Institution* 17(4): 525-548.
- 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.
- Haas, Peter M. 1990. *Saving the Mediterranean: the politics of international environmental cooperation*. Columbia University Press.
- Haas, Peter M. 1992. Introduction: Epistemic Communities and International Policy Coordination. *International Organization* 46(1): 1 - 35.
- 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*, edited by Finn Tarp and Peter Hjertholm, 271-289. London: Routledge.

- Herbst, Jeffrey. 1990. The Structural Adjustment of Politics in Africa. *World Development* 8(7).
- 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* (July/August): 515-521.
- Jacob, Brian A. 2005. Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics* 89: 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.
- Independent Evaluation Group (IEG). 2008. Environmental Sustainability: An Evaluation of World Bank Group Support. 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.
- IEG 2011.
- Keohane, Robert and Marc Levy. 1996. *Institutions for Environmental Aid: Pitfalls and Promise*. Cambridge, MA: MIT Press.
- Khagram, S., C. Thomas, C. Lucerno, and S. Mathes. 2009. Evidence for Development Effectiveness. *Journal of Development Effectiveness* 1(3): 247-70.
- Killick T, R. Gunatilaka R, and A. Marr.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.
- 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.
- Krause 2013.

- Lake, D. A., & Baum, M. A. 2001. The Invisible Hand of Democracy Political Control and the Provision of Public Services. *Comparative Political Studies* 34(6), 587-621.
- Lewis, T. 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, C., Peeters, C., & Tierney, M. J. 2011. Principles or Principals? Institutional Reform and Aid Allocation in the Global Environment Facility (GEF).
- Meagher, P. 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: 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.
- 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.
- Piotrowski 2003.
- 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, 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 2012/104. Helsinki: UNU-WIDER.
- Pritchett, L., Woolcock, M., & Andrews, M. 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. 1996. Understanding Economic Policy Reform. *Journal of Economic Literature* 34: 9-41.
- 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. 2007. *One Economics, Many Recipes: Globalization, Institutions and Economic Growth*. Princeton, NJ: Princeton University Press.
- Ross, M. 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, Conference on African Economic Development: Measuring Success And Failure. April 18 – 20, 2013, Simon Fraser University.
- Shirley, M. 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.
- Jordan Smith, D. 2003. Patronage, per diems and the “workshop mentality”: the practice of family planning programs in southeastern Nigeria. *World Development*, 31(4), 703-715.
- Snijders and Bosker, 2012
- Stasavage, D. (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. Available at: [http://www.gwu.edu/~igis/assets/docs/working\\_papers/igis\\_wp1\\_sud\\_olmstead-rumsey.pdf](http://www.gwu.edu/~igis/assets/docs/working_papers/igis_wp1_sud_olmstead-rumsey.pdf)
- Vach, W. 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.
- VanDeveer, S. D., & Dabelko, G. D. 2001. It's capacity, stupid: International assistance and national implementation. *Global Environmental Politics*, 1(2), 18-29.
- Weidner, H. 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: 341–366.
- 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](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. 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.
- World Bank. 2011.
- World Bank Carbon Finance Unit. 2013. Results-based finance for REDD-plus: Lessons learned from the Forest Carbon Partnership Facility and the BioCarbon Fund. Washington DC: World Bank Carbon Finance Unit. Accessed on 12 January 2014 at  
<https://www.forestcarbonpartnership.org/sites/fcp/files/2013/August2013/Results-based-finance-for-REDD-Plus.pdf>
- World Bank. n.d. Harmonized Evaluation Criteria for ICR and OED Evaluations. Washington D.C.: World Bank. <http://ieg.worldbankgroup.org/Data/HarmonizeEvalCriteria.pdf>
- Wynia, MK. Cummins, DS. VanGeest, JB. Wilson, IB. 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.