

# WORKING PAPER 100

July 2020

When Does Transparency Improve Institutional Performance? Evidence from 20,000 Aid Projects in 183 Countries

Dan Honig Johns Hopkins SAIS

**Ranjit Lall** London School of Economics

Bradley C. Parks The College of William & Mary

# Abstract

Access to information (ATI) policies are often praised for enhancing transparency, accountability, and trust in public institutions, yet evidence that they lead to better institutional performance is mixed. We argue that a key impediment to the effective operation of ATI policies is the failure of public officials to comply with information requests that could expose poor performance. Unless accompanied by reliable mechanisms for preventing noncompliance, therefore, ATI policies are unlikely to improve performance. We test our argument through a difference-in-differences analysis of a new dataset on the performance of approximately 20,000 foreign aid projects financed by 12 donor agencies in 183 countries between 1956 and 2016—the largest dataset of its kind. We find that enforcement matters: ATI policies are only associated with better project outcomes when they include independent appeals processes for denied information requests. In addition, we recover evidence for several microlevel observable implications of our argument.

# **Author Information**

#### Dan Honig

Johns Hopkins SAIS dhonig@jhu.edu

Ranjit Lall London School of Economics r.lall@lse.ac.uk

#### Bradley C. Parks

The College of William & Mary bparks@aiddata.wm.edu

The views expressed in AidData Working Papers are those of the authors and should not be attributed to AidData or funders of AidData's work, nor do they necessarily reflect the views of any of the many institutions or individuals acknowledged here.

# Acknowledgements

The authors owe a debt of gratitude to Muhammad Saad Imtiaz for his outstanding research assistance. We also thank Jennifer Brass, Alice Evans, Jonathan Fox, Osvaldo Feinstein, Christopher Kilby, John Marshall, Luciana Oliveira, Chris Wilson, and participants in the Annual Conference on the Political Economy of International Organization, the International Studies Association Annual Convention, the International Development Seminar at Georgetown University's Mortara Center for International Studies, the Political Science Seminar at Nuffield College, University of Oxford, and the Junior Scholars Forum at the Stanford Center on Philanthropy and Civil Society for helpful comments on earlier drafts. The College of William and Mary, Johns Hopkins SAIS, the London School of Economics, and the William and Flora Hewlett Foundation generously provided the funding that made this study possible.

#### Introduction

In 2011, a civil society activist from the Turkish city of Bartin approached the World Bank's country office in Ankara seeking information on a disaster recovery project financed by the organization and implemented by Turkish government entities a decade earlier.<sup>1</sup> Unsatisfied with the World Bank's response, the activist filed an Access to Information (ATI) request through a process established by the 2010 World Bank Policy on Access to Information.<sup>2</sup> This request was rejected by the World Bank, prompting the activist to make use of two independent appeals mechanisms provided by the policy.<sup>3</sup> Both appeals were also denied: the specific information requested did not exist in the World Bank's records. Yet this seemingly unsuccessful effort had significant ramifications for the organization's operations in Ankara. As one World Bank staff member involved in handling the request and appeals put it, "[W]e were all frightened—if someone requests, we have to address that."<sup>4</sup> Responding to the claims consumed "a tremendous amount of time and energy" within the Ankara office-staff had to translate and scan a large number of project-related documents to allow their counterparts at the World Bank's Washington D.C. headquarters to adjudicate the claim - and raised the salience of the ATI policy in day-to-day decision-making.<sup>5</sup> Perhaps most notably, the threat and eventual usage of the appeals process contributed to "a culture of caution and carefulness," prompting staff to pay significantly closer attention to internal rules and guidelines to ensure that they effectively discharged their responsibilities.<sup>6</sup>

In recent decades, ATI policies have emerged as the most prominent form of institutionalized transparency in governments and international organizations. By formally guaranteeing the right to request information from public institutions, ATI policies not only create new opportunities for stakeholders to monitor bureaucratic activities and to access public-sector

<sup>&</sup>lt;sup>1</sup> The project in question was the Turkey Emergency Flood and Earthquake Recovery Project, which was implemented between 1998 and 2003. Account draws on author interviews with a member of the World Bank's Turkey Country Management Unit, 26 March 2019, Washington D.C., and with a former member of the World Bank's Ankara office, 27 March 2019, Washington D.C.

<sup>&</sup>lt;sup>2</sup> World Bank Access to Information Request AI1362.

<sup>&</sup>lt;sup>3</sup> WB Access to Information Committee and Access to Information Appeals Board denials at (http://pubdocs.worldbank.org/en/735201433885011928/19-AIC-appeal-19-Case-AI1362.pdf) and

<sup>(</sup>http://pubdocs.worldbank.org/en/295211433883917207/2-CaseAI1362AIAppealsBoardDecision.pdf).

<sup>&</sup>lt;sup>4</sup> Author interview of 26 March 2019.

<sup>&</sup>lt;sup>5</sup> Author interviews of 26 and 27 March 2019.

<sup>&</sup>lt;sup>6</sup> Author interview of 26 March 2019.

knowledge, but also transform their political relationships to these institutions. While ATI policies have been widely praised for enhancing transparency, accountability, and trust in public institutions (e.g., Banisar 2006; Florini 2007), empirical evidence that they lead to better institutional performance is strikingly mixed.<sup>7</sup> Some studies find that they increase bureaucratic efficiency (Vadlamannati and Cooray 2016a) and reduce levels of corruption (Cordis and Warren 2014; Peisakhin and Pinto 2010). Other studies suggest that they may actually increase corruption (Escaleras et al. 2009; Costa 2013; Vadlamannati and Cooray 2016b), reduce bureaucratic quality (Costa 2013), and fail to improve administrative decisionmaking (Worthy 2010, 2013). We seek to contribute to this important debate by investigating the conditions under which ATI policies improve the performance of public institutions.

The bureaucratic features of ATI policies rarely draw close attention—but perhaps they should. We argue that a critical impediment to improved institutional performance is the failure to properly *enforce* ATI policies. Public officials, whether in government agencies or international organizations, have incentives to avoid complying with legitimate information requests that could reveal poor performance. Noncompliance is not only difficult to detect for actors outside public institutions but also unlikely to be sanctioned by actors within such institutions, who themselves stand to benefit from the option of withholding information. The existence of reliable mechanisms for detecting, exposing, and remedying noncompliance is thus essential for ensuring that ATI policies curtail "agency slack" and generate new information that can be used to improve institutional performance. As one senior World Bank official involved in handling the Bartın request put it, the organization's appeals process created "a tremendous incentive to do a good job and comply with [ATI] requests. If you're shoddy with it, [an appeal] causes twice as much trouble and takes twice as much time."<sup>8</sup>

We empirically evaluate our argument in the context of international development assistance. Specifically, we examine how the adoption of ATI policies by bilateral and multilateral donor agencies, such as the United Kingdom's Department for International Development (DFID) and the World Bank, influences the outcome of projects that they finance in low- and middleincome countries. Our analysis draws on a new dataset on the performance of around 20,000

<sup>&</sup>lt;sup>7</sup> Transparency, accountability, and trust are, of course, intrinsically valuable aside from any effect they may have on performance outcomes.

<sup>&</sup>lt;sup>8</sup> Author interview of 5 April 2019.

projects funded by 12 donor agencies in 183 countries between 1956 and 2016—the largest dataset on foreign aid project outcomes compiled to date.<sup>9</sup> The dataset contains holistic ratings of the extent to which projects achieve their principal objectives in an efficient manner, which are produced by donor staff and independent evaluation experts using a common set of assessment criteria and quality standards. We employ a staggered difference-in-differences design that exploits temporal variation in the adoption of ATI policies with and without a key enforcement mechanism: the existence of a formal recourse process that allows information seekers to appeal to an independent body when their ATI requests are denied.

This research design has three attractive features. First, it yields empirical findings that are likely to enjoy high levels of external validity. We believe that a central reason for the mixed state of the existing literature on the governance effects of ATI policies is that it largely consists of single-setting studies (e.g., one country, one sector, one public institution) that examine short time periods.<sup>10</sup> Our collection of project performance data presents a rare opportunity to assess the impact of ATI policies on a large, geographically diverse, and temporally extensive sample of performance outcomes across a wide range of sectors. Second, the staggered adoption of ATI policies and independent appeals mechanisms, combined with the temporal scope of projects in our dataset, allows us to control for a wide range of potentially confounding country-, donor-, and year-specific factors by comparing pre- and post-adoption trends in project success within a DD framework. Third, and relatedly, the adoption of ATI policies by donor agencies is plausibly exogenous to the country contexts in which their projects are delivered. This feature is unusual in purely domestic settings, where the adoption of ATI policies is usually a product of factors that may themselves influence institutional performance outcomes.

We find that the adoption of ATI policies by donors is not, in general, associated with improvements in the performance of projects that they finance. However, when such policies are accompanied by independent appeals processes, we observe a strong and positive association between these strong ATI policies and performance. These results are robust to a

 <sup>&</sup>lt;sup>9</sup> The same types of projects are undertaken in these countries in the absence of international funding. For a recent analysis of the factors shaping the outcome of domestically-financed development projects, see Williams (2017).
 <sup>10</sup> While many of these studies employ compelling identification strategies, their findings often do not "travel" to different geographical and temporal contexts (Fox 2015, 348).

variety of estimation methods, model specifications, and samples as well as to the use of instrumental variables. This effect, we argue, exists primarily because aid agency staff make decisions in the *shadow of the future*: they take steps to improve the design and supervision of the projects that they manage when they anticipate that ATI appeals might expose problems that could result in sanctions. In addition, we find support for several micro-level observable implications of this argument, including that previous usage of and success with independent appeals mechanisms in a given recipient country—which raises the likelihood that projects within that jurisdiction will be subject to external scrutiny—results in better project outcomes; that these localized "appeals shocks" encourage donor staff to devote greater effort and resources to project design and supervision, a sign of reduced agency slack; and that the main association is stronger when recipient countries have robust civil societies and high levels of press freedom – conditions under which citizens are more likely to make use of ATI policies and independent appeals mechanisms and to put pressure on donors and national governments to address project problems – but weaker when such countries already possess good governance characteristics.

Our findings contribute to a broader research agenda in political science and several other disciplines that seeks to identify the specific circumstances under which transparency improves the performance of public institutions. By highlighting the role of stakeholderempowering enforcement mechanisms in determining whether and when ATI policies enhance performance, they support an emerging consensus in this literature that—on their own information and bottom-up monitoring are "not enough" (Fox 2015, 248; also see Banerjee et al. 2010; Lieberman et al. 2014; Olken 2007). Instead, our findings reveal the importance of the *interaction* between bottom-up monitoring and top-down enforcement for holding public institutions to account. Our conditional findings, moreover, add to a growing body of evidence that transparency interventions are more potent in environments characterized by higher levels of civic engagement, press freedom, and other forms of bottom-up accountability (e.g., Grossman and Mitchelitch 2015; Kosack and Fung 2014; Lindstedt and Naurin 2010).

In drawing attention to independent appeals processes as an instrument of enforcement, the study also contributes to a related body of research on the political effects of institutionalized mechanisms for receiving, monitoring, and responding to complaints from stakeholders. This research has shown that such mechanisms—and other nonelectoral methods of political

participation—have the potential to increase government responsiveness to citizens (Bratton 2012; Cleary 2007). Our findings suggest an additional channel through which they can promote good governance, namely, ensuring the proper enforcement of transparency policies. Finally, our study adds to a burgeoning literature on the performance of foreign aid projects and donor agencies more generally by highlighting the role of institutionalized transparency in influencing the outcome of international development projects (Buntaine 2016; Denizer et al. 2013; Dreher et al. 2013; Honig 2018, 2019; Lall 2017; Winters 2014).

#### Access to Information Policies in Donor Agencies

ATI policies establish a formal process through which any public or private actor can request information held by donor agencies, including information about foreign aid projects that they finance. That is, ATI requests can be submitted both by actors "below" these agencies, such as citizens, civil society organizations (CSOs), academics, and journalists, and by actors "above" them, such as legislators and members of the executive branch (in aid-providing as well as aid-receiving countries).<sup>11</sup> For bilateral donors, which are either government agencies or state-owned development banks, ATI policies assume the form of freedom of information (FOI) laws. Most of these laws are based on the United States' 1966 Freedom of Information Act and were passed during the 1990s and 2000s in response to civil society campaigns for increased public-sector transparency and accountability (Ackerman and Sandoval-Ballasteros 2006; Banisar 2006; Florini 2007).<sup>12</sup>

Multilateral ATI policies take the form of binding rules approved by donors' governing or executive bodies. These policies are generally modeled on FOI laws and have similar features, enshrining the principle that the public has a right to know about the activities and policies of multilateral institutions.<sup>13</sup> Since the World Bank's pioneering 1994 Policy on Disclosure of Information, such rules have spread to a number of multilateral development banks and

http://pubdocs.worldbank.org/en/835741505831037845/pdf/2016-AI-Survey-Report-Final.pdf.

<sup>&</sup>lt;sup>11</sup> See e.g. World Bank Access to Information Survey 2016, available at:

<sup>&</sup>lt;sup>12</sup> Some studies have also highlighted the role of domestic political factors, such as electoral rules and party competition, in influencing the passage of FOI laws (e.g., Berliner 2014).

<sup>&</sup>lt;sup>13</sup> Other features that these policies often share with national FOI laws include a clear definition of types of information that can be shielded from public disclosure; a harm test for nondisclosure decisions (i.e., a determination that the harm caused by disclosure outweighs the benefits); an override provision that mandates disclosure when

Donor agency	Acronym	Donor type	ATI policy (year adopted)	Independent appeals mechanism
Department of Foreign Affairs and Trade, Australia	DFAT	Bilateral	Freedom of Information Act (1982)	Administrative Appeals Tribunal
Asian Development Bank	AsDB	Multilateral	Confidentiality and Disclosure of Information policy (1994)	None
			The Public Communications Policy of the Asian Development Bank (2005)	None
African Development Bank	AfDB	Multilateral	Public Communications Policy (2011) Disclosure of Information Policy (1997)	Independent Appeals Panel None
			The African Development Group Policy on Disclosure of Information (2005)	None
			Disclosure and Access to Information: The Policy (2012)	Appeals Panel
Caribbean Development Bank	CDB	Multilateral	Caribbean Development Bank Information Disclosure Policy (2011)	Appeals Panel
Department for International Development, United Kingdom	DFID	Bilateral	Freedom of Information Act (2000)	Information Commissioner's Office
Deutsche Gesellschaft für Internationale	GiZ	Bilateral	Federal Act Governing Access to Information held by the Federal	Federal Commissioner for Freedom of Information
Zusammenarbeit GmbH, Germany			Government (Freedom of Information Act) (2005)	
Global Environment Facility	GEF	Multilateral	GEF Practices on Disclosure of Information (2011)	None
Global Fund to Fight AIDS, Tuberculosis and Malaria	GFATM	Multilateral	Documents Policy (2007)	None
International Fund for Agricultural Development	IFAD	Multilateral	IFAD Policy on the Disclosure of Documents (1998)	None
			IFAD Policy on the Disclosure of Documents (revised) (2006)	None
			IFAD Policy on the Disclosure of Documents (revised) (2010)	None
Japan International Cooperation Agency	JICA	Bilateral	Act on Access to Information Held by Administrative Organs (1999)	Information Disclosure and Personal Information Protection Review Board
Kreditanstalt Fuer Wiederaufbau, Germany	KfW	Bilateral	Federal Act Governing Access to Information held by the Federal Government (Freedom of Information	Federal Commissioner for Freedom of Information
World Bank	WB	Multilateral	Act) (2005) World Bank Policy on Disclosure of Information (1994)	None
			World Bank Policy on Disclosure of Information (revised) (2002)	None
			World Bank Policy on Access to	Access to Information
			Information (2010)	Committee (first stage);
				Access to Information
				Appeals Board (second stage)

# Table 1. ATI Policies and Appeals Mechanisms Adopted by Donors in Dataset

doing so is in furtherance of a compelling public interest; and a time-bound process for submitting and deciding upon ATI requests.

international development finance institutions—a trend generally attributed to a combination of civil society and norm diffusion pressures (Buntaine 2016; Nelson 2001; Park 2005).

Most projects financed by donor agencies are implemented by public institutions in recipient countries (working in conjunction with donor country offices). ATI requests can thus shed valuable light on the inner workings of bureaucracy in developing countries. Consider again the civil society activist from Bartın, whose information request concerned a World Bank-financed project delivered entirely by the Turkish government: the Turkish Housing Development Administration in Ankara designed the project, identified the target beneficiaries, handled the bidding, awarded the contracts, and disbursed the funds. Nonetheless, the activist submitted her information request to the international financier of the project rather than to Turkey's national government, presumably because the World Bank's ATI policy provided a credible avenue through which she could obtain information that Turkish authorities might be reluctant to disclose.

While the activist's inquiry was unsuccessful, most ATI requests do result in disclosure. For example, the World Bank reports that of the 726 requests that received a decision during the Fiscal Year 2017, 501 led to the solicited information's release.<sup>14</sup> Information generated by successful requests—whether such requests were motivated by private interests or an explicit desire to hold public institutions to account—is publicly available. All disclosures made by the World Bank, for instance, are posted on its official ATI website.<sup>15</sup>

#### Monitoring, Enforcement, and Performance

#### Why the Right to Information Is Not Enough

There are several reasons why ATI policies might be expected to improve institutional performance. Expanded disclosure by public institutions enhances the ability of citizens, CSOs, the media, and other stakeholders to monitor their activities (Anderson et al. 2019; Berliner et al. 2018; Distelhorst 2017; Peisakhin 2012). If ATI requests reveal poor performance,

<sup>&</sup>lt;sup>14</sup> World Bank Access to Information Annual Report: Fiscal Year 2017, p. 7, available at: http://pubdocs.worldbank.org/en/742661529439484831/WBG-AI-2017-annual-report.pdf
<sup>15</sup> https://www.worldbank.org/en/access-to-information.

this information can be used by political principals — whether legislators and the executive branch in the case of government agencies or member states in the case of international organizations — to sanction institutions (Berliner and Erlich 2015; Nielson and Tierney 2003; Grigorescu 2007; Lorentzen et al. 2014). As suggested by the Bartın case, even if no information is actually released, the *threat* of disclosure should motivate public officials to avoid behavior that could subsequently be punished. Under some conditions, when bureaucrats are aware that they are being monitored or that their actions may be publicly disclosed, they may be less likely to shirk or to engage in malfeasant behavior and more likely to carry out their duties efficiently and responsibly (Dal Bó et al. 2018; Carlson and Seim 2018; Jablonski and Seim 2017). In short, ATI policies may limit agency slack by incentivizing public officials to increase their effort and efficiency both in anticipation and as a result of sanctions.

At the same time, there are reasons to doubt that ATI policies will—on their own—be sufficient to improve institutional performance. Rather than boosting their effort and productivity in response to such policies, public officials may choose the less burdensome option of refusing to comply with legitimate ATI requests that could expose underperformance. Perhaps the least costly way to avoid compliance is to reject such requests on procedural, technical, resource-related, or availability grounds (Prat 2005; Hood 2007; Holsen and Pasquier 2012; Trapnell 2014; Berliner 2017). As ATI requesters typically lack the information and expertise to dispute such decisions—and disputes can easily be characterized as differences in the interpretation of ATI policy provisions—this form of noncompliance has the added advantage of being difficult to detect. While illegitimate denials could be discovered by officials in other parts of the institution, these actors also benefit from the ability to conceal information about their performance and thus have weak incentives to actively monitor—let alone to sanction—such practices.

The implication is that ATI policies will only succeed in reducing agency slack when they are accompanied by reliable *enforcement* mechanisms—mechanisms that counterbalance incentives for noncompliance. As Neuman (2006, 10) emphasizes in the domestic context: "Enforcement of the law is critical; if there is widespread belief that [FOI] legislation will not be enforced, this so-called right to information becomes meaningless. If the enforcement mechanisms are weak or ineffective, it can lead to arbitrary denials, or it can foment the 'ostrich effect,' whereby there is no explicit denial but rather the government agencies put their

heads in the sand and pretend that the law does not exist. Thus, some external review mechanism is critical to [a FOI] law's overall effectiveness."

#### Independent Appeals Processes

What kind of "external review mechanism" can create a shared expectation of compliance? The principal mechanism for enforcing ATI policies is the existence of a formal recourse process that enables information seekers to appeal to a quasi-judicial independent body—a body composed of individuals who do not report to any member of the donor's staff—when their requests are denied. In multilateral donors, these bodies are usually panels of external ATI experts from civil society, business, or government who are authorized to uphold or reverse initial disclosure decisions. The World Bank's two-stage appeals process is slightly unusual in this respect, as it involves both an Access to Information Committee, which is composed of 10 director-level staff members and makes the first decision on an appeal; and an Access to Information Appeals Board, which comprises three outside ATI experts and considers appeals against Access to Information Committee decisions.

In domestic settings, independent appeals bodies take the form of judicial institutions that are responsible for overseeing overall government adherence to FOI legislation and, in some cases, for imposing or recommending penalties for noncompliance (Holsen and Pasquier 2012). For example, if a FOI request submitted to the United Kingdom's DFID is rejected, the information seeker can appeal to the Information Commissioner's Office (ICO), a non-departmental public body that reports to the country's parliament and possesses the authority to enforce compliance with such requests. Unlike at the multilateral level, therefore, ATI policy enforcement is embedded in the apparatus of domestic law. As reported in Table 1, nine of the 12 donors in our dataset possessed an ATI policy with an independent appeals mechanism as of 2016.<sup>16</sup>

The presence of an independent appeals mechanism can markedly affect how donors handle initial ATI requests. The ICO is a case in point. In 2008, an ICO ruling forced DFID to disclose the winning tender proposal for a consultancy contract along with the scores awarded to all

<sup>&</sup>lt;sup>16</sup> In our sample, all five bilateral donors simultaneously adopted ATI policies and independent appeals mechanisms, whereas most of the multilateral donors introduced them at different times.

submitted proposals.<sup>17</sup> This ruling set a precedent within DFID for automatic acceptance of FOI requests for tender-related information. In 2015, for instance, DFID accepted a FOI request for the tender documents associated with forensic audits of two local banks in Anguilla submitted by Keith Stone Greaves, a journalist and talk show host for the country's public broadcaster.<sup>18</sup> Greaves requested these documents to facilitate debate and to disseminate information of public interest on his weekly radio program "Talk Caribbean." As he explained to us, "I just wanted to inform the public...The public had a right to know what was happening with their banks."<sup>19</sup>

Therefore, in addition to enhancing compliance with ATI requests, independent appeals mechanisms can boost confidence in and usage of the request process and set precedents that clarify the scope of ATI policy provisions. Hazell and Worthy (2010, 353) describe these advantages in the domestic context: "[A] strong appeals process potentially locks FOI into a positive cycle of use, learning, and improvement, in which the request process and appeal system improve and the exemptions are clarified through interpretation. Such a finding would be a sign of [a FOI law] performing well. Conversely, if FOI is not used or the appeal system is weak, FOI may become locked into a negative cycle of disuse, neglect and stagnation."<sup>20</sup>

#### ATI Policies and Foreign Aid Effectiveness

The performance benefits of a properly enforced ATI policy are no less applicable to public institutions that finance foreign aid projects. Foreign aid is characterized by a lengthy chain of delegation involving politicians in aid-providing countries, donor agencies, project managers, contractors, government agencies in recipient countries, and intended beneficiaries. Information asymmetries within each principal-agent relationship can short-circuit the feedback loop between the two ends of the delegation chain (Easterly 2007). While project staff can be

 $(https://www.whatdotheyknow.com/request/request_for\_tender\_documents\_for\#incoming-737056).$ 

<sup>&</sup>lt;sup>17</sup> Information Commissioner's Office Decision Notice #FS50088016, 27 November 2008.

<sup>&</sup>lt;sup>18</sup> United Kingdom Freedom of Information Request F2015-398, 26 October 2015. The identity of ATI requesters is confidential; however, Greaves posted his request (and DFID's response) on a public website:

<sup>&</sup>lt;sup>19</sup> Author telephone interview with Keith Stone Greaves, 11 March 2019, Washington D.C.

<sup>&</sup>lt;sup>20</sup> As Caroline Anstey, Chair of the Center for Global Development Working Group on Commercial Transparency and a former Managing Director of the World Bank, recently remarked, "I don't ever think you're ever going to get an [ATI] process that is so perfect that you don't need appeals." Response to question by one of the authors at a Center for Global Development public event, 14 March 2019, Washington D.C.

held to account for performance outcomes by supervisors and political principals, they typically have weak incentives to ensure that projects are effectively designed and implemented. Instead, they are often rewarded for *quick* delivery (Phillips 2009). As multilateral donors depend on government clients who are willing to borrow from them, delaying project loan approvals and disbursements due to concerns about future performance can threaten their basic business model (Dollar and Svensson 2000; Kilby 2009). Similarly, bilateral donors receive "use-it-or-lose-it" appropriations from national legislatures, which encourages them to "push money out the door" as rapidly as possible (Drazen 2007, 672).

When project information can be readily and reliably accessed through a well-enforced ATI policy, actors inside and outside the delegation chain can more easily monitor the behavior of public officials, increasing the likelihood that poor performance will be detected and sanctioned (whether through outright dismissal or demotion, failure to receive promotions, or reassignment to less prestigious or consequential work). Independent appeals processes can improve project performance outcomes through two distinct mechanisms. The first is a *project correction effect* whereby an ATI appeal concerning a given project results in performance-enhancing modifications to this same project, whether by generating new information about design or implementation issues (if the appeal is successful) or by prompting officials to pay greater attention to potential performance problems (if the appeal is unsuccessful) (Ensminger and Leder-Luis 2018; Legovini et al. 2015; Reinikka and Svensson 2005). The second is a *shadow of the future* effect whereby officials take steps to strengthen project design and implementation in anticipation of ATI appeals that could reveal problems (Faust 2011; Buntaine 2016). As the Bartın case suggests, even ATI requests that concern completed projects (and are ultimately denied) can increase bureaucrats' awareness that they are being monitored.

While acknowledging that project correction effects can be important—anecdotal evidence indicates that appeals can substantially alter the behavior of staff involved in concerned projects—we expect project improvements to occur primarily through shadow of the future effects. As a proportion of total projects, the number of actual appeals cases is relatively small.<sup>21</sup> Additionally, most appeals are submitted after the projects that they concern have

<sup>&</sup>lt;sup>21</sup> For instance, the World Bank, which possesses the most high-profile appeals mechanism of any multilateral donor, adjudicated just 71 appeals cases between 2010 and mid-2019, a period in which it conducted thousands of projects.

concluded—that is, when new information cannot be used to remedy problems with these projects.<sup>22</sup> Therefore, while shadow of the future effects may be weaker on a per-project basis, they are likely to affect a far higher proportion of projects than correction effects. While empirically distinguishing between these two types of effects is challenging, we will provide some systematic evidence on shadow of the future effects and the specific channels through which they operate.

#### Hypotheses

The preceding discussion can be summarized in the following hypotheses:

H1: The adoption of ATI policies by donor agencies will not, by itself, be associated with an improvement in the performance of projects that they finance; however,

H2: The adoption of ATI policies with independent appeals mechanisms will be associated with an improvement in project performance.

#### Data

#### Outcome Variable

In line with a growing literature on the determinants of foreign aid effectiveness, we measure project performance using holistic, ex-post success ratings produced by donor staff and evaluation experts (Denizer et al. 2013; Dreher et al. 2013; Buntaine 2016; Honig 2018, 2019; Kilby 2015). These ratings are attractive because they provide a consistent and comparable measure of performance across projects, sectors, countries, and time. The specific measurement criteria are based on a series of Organisation for Economic Cooperation and Development (OECD) standards on the evaluation of five distinct dimensions of project performance: (1) relevance (to national development goals); (2) efficiency; (3) achievement of

<sup>&</sup>lt;sup>22</sup> Indeed, only 10 of the World Bank's 71 appeals cases received a final decision before the completion of the project in question. Six of these projects were completed as of mid-2019, with the other four still ongoing. We focus on the World Bank in part due to its transparency regarding appeals cases (discussed further below).



Figure 1. Trends in Project Success for Individual Donors

project objectives; (4) impact; and (5) sustainability (OECD 1991, 2000, 2010).<sup>23</sup> While success ratings do not directly capture the specific outcomes achieved in a given project—which inevitably vary across projects—previous research has demonstrated that they are positively correlated with broader indicators of socioeconomic development (Warner 2010; Metzger and Guenther 2015).

Through a large-scale data collection effort that began in 2012 and involved extensive communications and negotiations with donor staff and evaluation teams, we obtained ratings for approximately 20,000 projects financed by 12 donors between 1956 and 2016 (essentially

<sup>&</sup>lt;sup>23</sup> Although the OECD's standards may be interpreted in different ways by donors, the inclusion of donor fixed effects in our empirical analysis controls for any constant donor-specific, time-invariant heterogeneity in interpretation or scaling.

all projects rated by these donors during this period).<sup>24</sup> These ratings cover projects undertaken in 183 countries across all regions of the world and in virtually every sector of government activity. The majority of projects take place in four regions—East Asia and the Pacific (21.2% of the total), Latin America and the Caribbean (13.6%), South Asia (13.6%), and Sub-Saharan Africa (30.0%)—and are confined to a single country (95.9%). Section 1 of the Appendix provides donor-by-donor descriptive statistics on project location, geographical scope, ratings, timing, and length.

The outcome variable in our analysis, *Project Success*<sub>*r,d,t*</sub>, is the rating for a project financed by donor *d* in recipient country *r* that began in year *t*, which is measured on a Likert-type scale that ranges from 1 for "highly unsatisfactory" to 6 for "highly satisfactory."<sup>25</sup> Figure 1 plots average levels of *Project Success*<sub>*r,d,t*</sub> over time for each donor. No general cross-donor trends are discernible from the graphs. Some donors' ratings exhibit evidence of modest "grade inflation", with average levels rising over time; other donors have seen their ratings decline in more recent years. For a third group of donors, ratings have fluctuated around an approximately constant level throughout the period covered by the data. Donor-specific summary statistics for *Project Success*<sub>*r,d,t*</sub> can be found in Appendix Section 1.

#### Treatment Variables

We merge the project performance ratings with a second set of original data on donor-level ATI policies, which cover the same agencies and time period. Our first treatment variable, *ATI Policy*<sub>*d*,*t*</sub>, is a dummy for whether donor *d* possesses an ATI policy in year *t*. Our main source of information on FOI laws is the Right to Information Rating database compiled by Access Info Europe and the Centre for Law and Democracy. We access multilateral ATI policies from donors' websites.<sup>26</sup> Our second treatment, *Appeals Mechanism*<sub>*d*,*t*</sub>, is a dummy for whether donor *d* possesses an ATI policy mether donor *d* possesses an ATI policies from donors and the Centre for Law and Democracy. We access multilateral ATI policies from donors' websites.<sup>26</sup> Our second treatment, *Appeals Mechanism*<sub>*d*,*t*</sub>, is a dummy for whether donor *d* possesses an ATI policy with an independent appeals mechanism for denied

<sup>&</sup>lt;sup>24</sup> An earlier version of the database was introduced by [author] (2018). The current version adds several donors and almost doubles the number of projects.

<sup>&</sup>lt;sup>25</sup> These classifications follow the World Bank's rating system, which is the best known. While some donors employ alternative scales (e.g., from 1 to 4), we transform them to a common six-point scale for ease of substantive interpretation. As discussed below, our findings are not contingent on this transformation.

<sup>&</sup>lt;sup>26</sup> The Right to Information Rating database can be accessed at: https://www.rti-rating.org. To access superseded multilateral ATI policies, we use the Internet Archive's Wayback Machine (https://archive.org/web).

Variable	Observations	Mean	Std. Dev	Min	Max
Project Success <sub>r,d,t</sub>	20686	4.20	1.17	1	6
ATI Policy <sub>d,t-1</sub>	21301	0.48	0.50	0	1
Appeals Mechanism <sub>d,t-1</sub>	21301	0.15	0.36	0	1
Recipient GDP per Capita Growth <sub>r,t-1</sub>	23963	2.80	6.01	-65.00	140.50
Recipient Log GDP per Capita,,t-1	23886	7.57	1.31	4.75	11.88
Recipient Aid/GNI <sub>r,t-1</sub>	20932	7.10	10.76	-2.63	242.29

**Table 2.** Summary Statistics for Baseline Variables

information requests in year *t*. Wherever possible, our coding decisions follow the Right to Information Rating database and Publish What You Fund's Aid Transparency Index, the two existing comparative assessments of ATI appeals mechanisms.<sup>27</sup>

## **Control Variables**

We control for three recipient country-level variables that commonly feature in analyses of project performance: the annual growth rate of a recipient country's gross domestic product (GDP) per capita (*Recipient GDP per Capita Growth*<sub>*r*,t</sub>); the log of a recipient country's GDP per capita (*Recipient Log GDP per Capita*<sub>*r*,t</sub>); and the net official development assistance provided to a recipient country as a percentage of its gross national income (*Recipient Aid/GNI*<sub>*r*,t</sub>). All variables are measured as of year *t* in current United States dollars using data from the World Bank's World Development Indicators database.<sup>28</sup>

Table 2 provides summary statistics for all variables in the dataset; more detailed descriptions of each variable are provided in Appendix Section 1.

<sup>&</sup>lt;sup>27</sup> All but three donors in our dataset—CDB, GEF, and IFAD—are included in at least one of the two assessments. Our coding diverges from these assessments in only one case: unlike Publish What You Fund, we code the World Bank as possessing an independent appeals mechanism. Publish What You Fund acknowledges the existence of such a mechanism but suggests that "it is limited and there is no right to appeal certain information items" (http://www.publishwhatyoufund.org/the-index/2016/donor/world-bank-ida/). As discussed below, our findings are not contingent on this coding decision.

<sup>&</sup>lt;sup>28</sup> http://wdi.worldbank.org/tables.

#### **Empirical Analysis**

We employ a difference-in-differences design that estimates the difference between the change in project success before and after the two treatments in groups of observations that do and do not "receive" these treatments. The unit of analysis is a donor-recipient-project-year (there are 12 donors, 183 recipient countries, and 60 years). To assess the relationship between ATI policies and project success (H1), we estimate the following three-way fixed effects model with ordinary least squares (OLS):

$$Project \ Success_{r,d,t} = \alpha + \phi_d + \gamma_r + \psi_t + \beta_1 ATI \ Policy_{d,t-1} + \beta_2 Controls_{r/d,t-1} + \varepsilon_{r,d,t}$$
(1)

where  $\phi_d$ ,  $\gamma_r$ , and  $\psi_t$  are fixed effects for donors, recipient countries, and years, respectively. All covariates are lagged by one year in part to avoid simultaneity issues and in part because they are unlikely to have an instantaneous impact on project performance. To address the possibility of serial correlation in the outcome variable, we multiway cluster robust standard errors both by recipient country and by donor (Bertrand, Duflo, and Mullainathan 2004).

The three sets of fixed effects control for potential confounders that are specific to donors but invariant across recipient countries and years; specific to recipient countries but invariant across donors and years; and specific to years but invariant across donors and recipient countries. The resulting difference-in-differences estimator exploits variation in project success *within* groups of donor-recipient observations over time, addressing many possible sources of endogeneity while avoiding direct inter-donor comparisons, which could be problematic due to the partly subjective nature of performance evaluation (Honig 2019, 184-185). More specifically, since the timing of the two treatments varies across donors, the estimator represents a weighted average of all possible two-period difference-in-differences estimators comparing donor-recipient groups that are either treated or not treated in a given year (Goodman-Bacon 2018).<sup>29</sup>

<sup>&</sup>lt;sup>29</sup> Every project is part of the control group in some of these estimators. Weights come from the size of each group and the variance of the treatment.

We assess H2 in two ways. First, we substitute *Appeals Mechanism*<sub>d,t-1</sub> for *ATI Policy*<sub>d,t-1</sub> in Equation 1:

$$Project \ Success_{r,d,t} = \alpha + \phi_d + \gamma_r + \psi_t + \beta_1 A ppeals \ Mechanism_{d,t-1} + \beta_2 Controls_{r/d,t-1} + \varepsilon_{r,d,t}$$
(2)

Second, we add *ATI Policy*<sub>*d,t*-1</sub> to Equation 2:

$$Project \ Success_{r,d,t} = \alpha + \phi_d + \gamma_r + \psi_t + \beta_1 ATI \ Policy_{d,t-1} + \beta_2 Appeals \ Mechanism_{d,t-1} +$$
(3)  
$$\beta_2 Controls_{r/d,t-1} + \varepsilon_{r,d,t}$$

In Equation 2, the coefficient on *Appeals Mechanism*<sub>*d,t-1*</sub> captures the difference between the change in project success in observations that are subject to an ATI policy with an independent appeals mechanism versus observations that are subject either to no ATI policy or to an ATI policy without an independent appeals mechanism. In Equation 3, the control group shrinks to observations that are subject to an ATI policy without an independent appeals mechanism, allowing us to isolate the "added value" of properly enforcing ATI policies.

#### Results

The results of Equations 1-3, reported in Table 3, are consistent with H1 and H2. In Equation 1, the estimated coefficient on *ATI Policy*<sub>*d*,t-1</sub> is positive but small and not statistically distinguishable from zero (Model 1). In substantive terms, the possession of an ATI policy – with or without an independent appeals mechanism – is associated with an average increase in *Project Success*<sub>*r*,*d*,*t*</sub> of 0.02 (on a 1-6 scale). In percentile terms, this increase would not alter the rank of a project at the median level of *Project Success*<sub>*r*,*d*,*t*</sub> by a single percentage point.

In Equation 2, by contrast, the coefficient estimate on *Appeals Mechanism*<sub>*d,t*-1</sub> is positive, substantially larger, and statistically significant at the 1% level (Model 2). On average, *Project Success*<sub>*r,d,t*</sub> is 0.29 higher in the presence of an ATI policy with an independent appeals mechanism—several times the increase associated with the presence of an ATI policy in general. Substantively, this effect size is equivalent to moving from the 50<sup>th</sup> percentile of *Project Success*<sub>*r,d,t*</sub> to almost the 60<sup>th</sup> percentile.

Outcome: Project Success <sub>t</sub>	(1)	(2)	(3)
ATI Policy <sub>t-1</sub>	0.020		-0.067
	(0.097)		(0.066)
Appeals Mechanism <sub>t-1</sub>		0.290***	0.320***
		(0.081)	(0.084)
Recipient GDP per Capita Growth <sub>t-1</sub>	0.006***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)
Recipient Log GDP per Capitat-1	-0.187**	-0.184**	-0.184**
	(0.072)	(0.071)	(0.072)
Recipient Aid/GNI <sub>t-1</sub>	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)
Observations	17,929	17,929	17,929
R-squared	0.131	0.133	0.133
Recipient Country Fixed Effects	Yes	Yes	Yes
Donor Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes

Table 3. Relationship between ATI Policies, Appeals Mechanisms, and Project Success

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses.

<sup>\*</sup>p<0.01, <sup>\*\*</sup>p<0.05, <sup>\*</sup>p<0.1

Including both treatments as regressors in Equation 3 does not materially alter the size or significance level of the coefficient on *Appeals Mechanism*<sub>*d,t*-1</sub> (Model 3). When ATI policies are accompanied by an independent appeals mechanism, *Project Success*<sub>*r,d,t*</sub> is 0.32 higher than when they lack such a mechanism. Interestingly, the coefficient on *ATI Policy*<sub>*d,t*-1</sub> turns negative, indicating that the presence of an ATI policy without an independent appeals mechanism is associated with lower levels of project performance than the absence of an ATI policy. As in Model 1, however, the size of the coefficient is substantively small: *Project Success*<sub>*r,d,t*</sub> declines by an average of just 0.07 as *ATI Policy*<sub>*d,t*-1</sub> shifts from 0 to 1.

#### Robustness

Parallel Trends Assumption

The key identifying assumption of the difference-in-differences estimator is that trends in the outcome variable would have been the same in treated and control units in the absence of the treatment, conditional on covariates. We probe this assumption using two common strategies (Angrist and Pischke 2008, Ch. 5). First, we include 1-8 year leads and lags as well as a contemporaneous version of *Appeals Mechanism*<sub>d</sub> in Equation 2, with the expectation that the estimated coefficients on the leads will be statistically zero. While year-by-year effects are relatively noisy, none of the lead coefficients are significant at the 5% level.<sup>30</sup> Second, we show that the results are robust to the inclusion of recipient country-specific time trends in the model i.e., interactions between a dummy for each recipient country and a linear time trend.<sup>31</sup> These terms help to control for systematic differences in the pretreatment trajectory of the outcome variable between the treated and control groups. Section 2 of the Appendix provides both sets of estimates.

#### Validating Outcome Measure

Although project outcomes are evaluated according to a common set of criteria and quality standards, it is possible that staff are able to "game" ratings for their own ends (such as improving perceptions of their performance) (Buntaine et al. 2017; Denizer et al. 2013). While this possibility does not provide an obvious explanation for the baseline results—it is not clear why staff would only seek to game ratings *after* the adoption of ATI policies with independent appeals mechanisms—it may nonetheless raise concerns about the general reliability of our outcome measure.

We seek to address such concerns using two strategies. First, we re-estimate the baseline models with an independent measure of project success constructed by Malik and Stone (2018), which covers 4,206 World Bank projects—approximately 3,300 of which are in our dataset—conducted between 1994 and 2013. This measure is based on the extraction of more granular information about projects from Implementation Completion and Results (ICR) reports

<sup>&</sup>lt;sup>30</sup> In contrast, the coefficient on the 2-year lag is positive and significant, while those on the 5- and 7-year lags are positive and very close to significant.

<sup>&</sup>lt;sup>31</sup> We cannot control for donor-specific time trends because the set of interactions between donor dummies and a linear time trend would fully absorb the treatment effects (which are, in effect, donor-specific year fixed effects).

produced by World Bank project teams.<sup>32</sup> Specifically, Malik and Stone identify all project subobjectives listed in a given ICR and, drawing on qualitative and quantitative information in this report, generate a progress score for each one (on an ordinal 0-4 scale).<sup>33</sup> They then average scores across sub-objectives into an overall performance rating for the project. These ratings should be less susceptible to bias both because they were not produced by World Bank staff and because they guarantee that progress on every project sub-objective is assessed separately and equally weighted. The correlation between our outcome variable, *Project Success*<sub>r,d,t</sub>, and Malik and Stone's performance index is positive but far from perfect (r = 0.43).

Second, we re-estimate the baseline models restricting the sample to projects that were rated when donors possessed an independent unit (e.g., division, department, office) whose primary or exclusive task is to evaluate their performance. Examples include the World Bank's Independent Evaluation Group (IEG), the Asian Development Bank's Independent Evaluation Department, and the United Kingdom's Independent Commission for Aid Impact. This strategy has the advantage of allowing us to include an additional nine donors in the analysis, expanding the sample from around 3,300 projects to almost 12,000 projects. As shown in Appendix Section 3, both sets of results are consistent with the main findings.

#### Assessing Inferential Leverage

Given the size and heterogeneity of our dataset, it is important to understand the sources of inferential leverage in the analysis. We investigate this issue using two recently developed strategies. First, we calculate the "effective sample" in Equations 1 and 2 – the sample that is actually used to generate the estimates – using the multiple regression weights approach proposed by Aronow and Samii (2016).<sup>34</sup> Summary statistics for this sample are very similar to those for the nominal sample, suggesting that it is broadly representative of the dataset (and thus that the findings have strong internal validity). Second, we perform a Bacon

<sup>&</sup>lt;sup>32</sup> The World Bank's Independent Evaluation Group (IEG) audits a subset of ICRs (and the scores therein). Due to resource constraints, field-based audits ("Project Performance Assessment Reports" or PPARs) are conducted for 20-30% of ICRs and desk-based audits ("ICR Reviews" or ICRRs) are conducted for 70-80% of ICRs.

 <sup>&</sup>lt;sup>33</sup> A score of 0 indicates zero or negative progress; 1 indicates that up to one third of the sub-objective was achieved;
 2 indicates that between one third and two thirds of the sub-objective was achieved;
 3 indicates that at least two thirds but less than 100% of the sub-objective was achieved; and 4 indicates 100% achievement or overachievement.
 <sup>34</sup> Multiple regression weights can only be calculated with one treatment at a time, hence the exclusion of Equation 3.

decomposition (Goodman-Bacon 2018) on the estimated coefficient on *Appeals Mechanism*<sub>d,t-1</sub> in Equation 2, which disaggregates it into four separate two-period difference-in-differences estimates that compare (1) projects treated in different years, (2) projects treated in a given year with always-treated projects; (3) projects treated in a given year with never-treated projects, and (4) always-treated projects with never-treated projects.<sup>35</sup> All four estimates are large and positive. The first and third estimates have a collective weight of 90% in the aggregate estimate and are almost identical in size to it. The results of both tests are reported in Section 4 of the Appendix.

#### Alternative Samples

We also examine the sensitivity of the baseline results to three alternative sample restrictions: (1) including only those projects that began during a five-year window around the date of treatment adoption *j* (i.e., the period from j - 2 to j + 2), which helps to control for potential confounders that vary between pre- and post-treatment periods; (2) including only those projects that began either before or immediately after year *j*, which mitigates the possibility that donors select "easier" projects after treatment adoption;<sup>36</sup> (3) excluding all projects conducted after (a) 1990, (b) 1995, and (c) 2000 on the grounds that older projects may have been rated according to different criteria and quality standards. All three sets of results are similar to the baseline estimates (see Appendix Section 5).

#### Instrumental Variables Analysis

While the adoption of ATI policies can reasonably be viewed as exogenous to recipient country-specific factors, it could nevertheless be affected by omitted variables related to

<sup>&</sup>lt;sup>35</sup> This is implemented with the bacondecomp command in Stata, which requires collapsing the data into a donor-year panel (the overall coefficient estimate does not substantively change). The results also include a "within" residual component capturing variation in the evolution of covariates across projects sharing the same treatment. <sup>36</sup> Specifically, we estimate a series of models in which the sample is restricted to projects that begin before years *j* + 1, *j* + 2, ..., *j* + 5. Note that many of the likely determinants of project "difficulty," such as recipient country income and location (Briggs 2019), are also captured by our controls and fixed effects.

#### Table 4. Instrumental Variables Estimates (Second Stage)

Outcome: Project Successt	(1)	(2)	(3)	(4)	(5)	(6)
ATI Policy <sub>t-1</sub>	0.013		-0.265**	0.231		-0.255
	(0.149)		(0.092)	(0.194)		(0.229)
Appeals Mechanism <sub>t-1</sub>		0.679**	0.938***		0.590***	0.746***
		(0.235)	(0.238)		(0.183)	(0.189)
Observations	16 943	16 943	16 943	16 943	16 943	16 943
Recipient Country Donor &				,		
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Instrument reference group	Neighbors	Neighbors	Neighbors	Trading	Trading	Trading
				partners	partners	partners
Cragg-Donald F-Statistic (first stage)	7913	4742	2314	3820	6542	2466

Second-stage two-stage least squares estimates with robust standard errors, clustered by donor and recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub> and Recipient Aid/GNI<sub>t-1</sub>.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

project success—or by project success itself (for instance, if donors with more successful projects are more willing to disclose information about their activities).<sup>37</sup>

To address the possibility of endogeneity in treatment assignment, we employ an instrumental variables (IV) approach that leverages sources of plausibly exogenous variation in the adoption of ATI policies. As the determinants of adoption differ for bilateral and multilateral donors, we construct separate instruments for each group. Building on evidence of the diffusion of FOI laws across countries (Berliner 2014) and a common spatial IV strategy in the political economy literature, we construct two instruments for bilateral donors: (1) the lagged proportion of a donor's (a) geographical neighbors and (b) five largest trading partners that possess a FOI law with an independent appeals mechanism (for *Appeals Mechanism<sub>d,t-1</sub>*) or without an

<sup>&</sup>lt;sup>37</sup> The latter scenario is clearly less likely in the case of bilateral donors, which, as discussed earlier, have typically adopted FOI laws in response to civil society pressures for a general strengthening of governmental transparency and accountability—pressures that have little connection to foreign aid outcomes.

independent appeals mechanism (for *ATI Policy*<sub>*d*,*t*-1</sub>).<sup>38</sup> Our instrument for multilateral donors is the lagged proportion of a donor's five largest shareholder countries that possess an ATI policy with an independent appeals mechanism (for *Appeals Mechanism*<sub>*d*,*t*-1</sub>) or without an independent appeals mechanism (for *ATI Policy*<sub>*d*,*t*-1</sub>). We merge this variable with each bilateral instrument, producing two combined instruments for the treatments. The logic behind these instruments is that the adoption of an ATI policy by a donor's neighbors, major trading partners, and principal shareholders—or variables that predict this event—are likely to influence its own likelihood of adopting such a policy but do not directly affect the outcome of foreign aid projects that it finances (rendering the exclusion restriction plausible).<sup>39</sup>

We implement the IV analysis using a two-stage least squares estimator. In the first stage, we generate predicted values for each treatment by regressing it on the combined instrument and all controls and fixed effects in the baseline models:

$$Treatment_{d,t-1} = \alpha + \gamma_r + \phi_d + \psi_t + \beta_1 Combined Instrument_{d,t-1} + \beta_2 Controls_{r,t-1} + \varepsilon_{r,d,t}$$
(4)

In the second stage, *Project Success*<sub>*r,d,t*</sub> is regressed on the predicted values of the treatment from the first stage as well as all controls and fixed effects:

$$Project \ Success_{r,d,t} = \alpha + \gamma_r + \phi_d + \psi_t + \beta_1 Treatment_{d,t-1} + \beta_2 Controls_{r,t-1} + \varepsilon_{r,d,t}$$
(5)

Table 4 presents the second-stage results for the baseline equations. In the first stage, as reported in the bottom row, the combined instrument has a high Cragg-Donald F-Statistic in every model, ruling out possible bias from a "weak" instrument. All four estimated coefficients on the instrumented measures of *Appeals Mechanism*<sub>d,t-1</sub> are positive, sizable, and significant at the 1% or 5% level. These coefficients substantially exceed the corresponding baseline estimates: on average, the presence of an ATI policy with an independent appeals mechanism

<sup>&</sup>lt;sup>38</sup> A geographical neighbor is defined as a sovereign state with which the donor country shares a land or maritime border.

<sup>&</sup>lt;sup>39</sup> It is possible, of course, that these instruments are capturing omitted long-run or structural forces that push toward ATI policy adoption as well as project success. It is not obvious, however, what variables would affect both the disclosure decisions of a high-income country or major multilateral organization and the on-the-ground delivery of a public project in a low- or middle-income nation. Note, moreover, that the inclusion of donor and recipient country fixed effects helps to mitigate confounding by slow-moving omitted variables.

is associated with an increase in *Project Success*<sub>*r,d,t*</sub> of 0.74 — equivalent to moving from the 50<sup>th</sup> percentile of *Project Success*<sub>*r,d,t*</sub> to between the 61<sup>st</sup> and 90<sup>th</sup> percentiles. This increase is slightly greater for the neighbor-based instrument (0.81) than for the trading partner-based instrument (0.67), though in both cases it is more than twice as large as the average baseline figure (0.31). Similarly to before, the estimated coefficients on the instrumented measures of *ATI Policy*<sub>*d,t-1*</sub> are positive in the Equation 1 models, negative in the Equation 3 models, and mostly nonsignificant (the only significant estimate has a negative sign). The results thus provide additional support for H1 and H2 while suggesting that any potential endogeneity in treatment assignment in the baseline analysis worked *against* rather than for our argument.<sup>40</sup>

#### Placebo Test

Another possible concern is that *Appeals Mechanism<sub>d,t-1</sub>* is merely serving as a proxy for the overall stringency or quality of a donor's disclosure regime. We therefore conduct a placebo test in which the treatment is a dummy for whether an ATI policy codifies the "presumption of disclosure" principle, that is, a provision that establishes disclosure as the general rule and hence requires a compelling reason for nondisclosure. Generally regarded as a hallmark of a robust disclosure regime (Mendel 1999), this principle is one of the chief indicators of ATI policy stringency in both the Right to Information Rating Database and the Aid Transparency Index. The estimated coefficient on the placebo treatment is small, negative, and nonsignificant (see Appendix Section 6).

#### Additional Robustness Checks

The baseline results are robust to several additional checks, further information on which is provided in Section 7 of the Appendix. First, we include a number of additional control variables, some of which feature in previous analyses of project performance: (1) project size;<sup>41</sup> (2) dummies for project sector; and (3) a dummy for recipient country membership of the United Nations Security Council, which has been shown to influence project performance

<sup>&</sup>lt;sup>40</sup> The implication is that the adoption of ATI policies with independent appeals mechanism is more likely to stem from *weak* rather than strong project performance. This is consistent with some accounts of the creation of the World Bank's Policy on Disclosure of Information in 1994 (e.g., Fox and Brown 1998; Nielson and Tierney 2003).
<sup>41</sup> This is measured in terms of loan amount, loan commitment, or project expenditures, data on which are expressed in varying currencies across donors (donor fixed effects control for such time-invariant, donor-specific factors).

ratings (Kilby and Michaelowa forthcoming). Second, we examine whether our findings systematically vary between bilateral and multilateral donors by disaggregating the sample by donor.<sup>42</sup> Third, following a thread in the literature on project performance ratings (discussed in more detail in the Appendix), we collapse *Project Success*<sub>*r,d,t*</sub> into a binary variable based on (1) the mean rating in the sample, (2) the median rating in the sample, and (3) the maximum possible rating. Fourth, we employ longer lags (ranging from 2 to 5 years) for the treatment variables. Fifth, rather than converting them to a common scale—and hence adding noise—we leave ratings in their raw form (inverting some scales to ensure consistency in direction). Sixth, we experiment with an alternative coding of *Appeals Mechanism*<sub>*d,t-1*</sub> for the World Bank (see fn. 37). Finally, we estimate standard errors using three alternative techniques: (1) nonparametric bootstrapping; (2) clustering by donor only; and (3) clustering by donor × recipient country.

#### Exploring Observable Implications: An Analysis of World Bank Projects

In this section, we explore several observable implications of our argument. If the logic of the argument is correct, the likelihood that an ATI policy with an independent appeals mechanism will lead to increased scrutiny of projects should be higher when stakeholders in recipient countries have previously used this mechanism—and even higher when they have used it to *successfully* challenge an ATI denial (see Figure 2). Thus, the filing of an appeals in a given recipient country should be associated with an improvement in the performance of projects in this country, while the filing of a successful appeal should be associated with an even greater improvement. In addition, we should expect these localized appeals shocks to not only enhance project performance but also to trigger a set of intermediate behavioral changes by officials responsible for delivering projects. In particular, as the likelihood of greater scrutiny increases, officials should devote greater effort and resources to the preparation and supervision of projects, signaling a reduction in agency slack.

We assess these implications by analyzing different characteristics of several thousand World Bank projects from the past three decades. We focus on these projects for three reasons. First,

<sup>&</sup>lt;sup>42</sup> While the treatment effects naturally vary across donors, there is no consistent difference in their size and strength between the two groups. Note that some donors in our sample (including all bilateral donors) adopt ATI policies and independent appeals mechanisms simultaneously; thus, we only report the results for Equation 2 (omitting donor and year fixed effects).





unlike other donors in our dataset, the World Bank publishes a comprehensive online list of its ATI appeals cases that includes information on the projects they concern and their outcome. Second, through an ATI request, we were able to obtain micro-level data on budgetary allocations made by local World Bank staff for project design and supervision activities.<sup>43</sup> We use these data to construct parallel measures of design and supervision effort: *Preparation Cost Ratio<sub>r,t</sub>*, expenditures on project preparation activities as a percentage of the total project budget; and *Supervision Cost Ratio<sub>r,t</sub>*, expenditures on project supervision activities as a percentage of the total project budget.<sup>44</sup> Third, the World Bank's IEG has evaluated more than 9,000 projects on the quality of their design (since 1991) and their supervision (since 1994).<sup>45</sup> Design quality is measured with the indicator *Quality at Entry<sub>r,t</sub>*, which captures "[t]he extent to which the Bank identified, facilitated preparation of, and appraised the operation such that it was most likely to achieve planned development outcomes and was consistent with the Bank's fiduciary role.<sup>#46</sup> Supervision quality is measured with *Quality of Supervision<sub>r,t</sub>*, which captures "the extent to which the Bank proactively identified and resolved threats to the achievement of relevant development outcomes and the Bank's fiduciary role [during project]

<sup>&</sup>lt;sup>43</sup> The data are now publicly available at the following link:

http://pubdocs.worldbank.org/en/224581443472828614/FY2000-FY2015-Completion-and-Supervision-Cost.

<sup>&</sup>lt;sup>44</sup> Costs are measured in thousands or millions of United States dollars.

<sup>&</sup>lt;sup>45</sup> See the World Bank Project Performance Ratings Dataset, available at:

https://datacatalog.worldbank.org/dataset/ieg-world-bank-project-performance-ratings.

<sup>&</sup>lt;sup>46</sup> https://ieg.worldbankgroup.org/sites/default/files/Data/reports/ieg-wb-project-performance-ratingscodebook\_092015.pdf, p. 17.

implementation]."47 Both indicators have the same 1-6 ordinal scale as the World Bank's project performance ratings.

To test the above observable implications, we make two sets of changes to Equation 2. First, we assign our measures of project design and supervision effort and quality – *Preparation Cost Ratio<sub>r,t</sub>*, *Supervision Cost Ratio<sub>r,t</sub>*, *Quality at Entry<sub>r,t</sub>*, and *Quality of Supervision<sub>r,t</sub>* – as separate outcome variables alongside *Project Success<sub>r,d,t</sub>*.<sup>48</sup> Second, we replace *Appeals Mechanism<sub>t-1</sub>* with three alternative treatments: (1) *Appeals Shock<sub>r,t-1</sub>*, a dummy for whether any projects in recipient country *r* have previously been the subject of an ATI appeal as of year *t* – 1; (2) *Successful Appeals Shock<sub>r,t-1</sub>*, a dummy for whether any such projects have been the subject a successful appeal; and (3) *Unsuccessful Appeals Shock<sub>r,t-1</sub>*, a dummy for whether any such projects have been the subject of an unsuccessful appeal.<sup>49</sup> There are thus five outcome variables and three treatments, which combine to produce 15 different models:

$$\begin{cases}
Project Success_{r,t} \\
Preparation Cost Ratio_{r,t} \\
Supervision Cost Ratio_{r,t} = \alpha + \gamma_r + \psi_t + \beta_1 \\
Quality at Entry_{r,t} \\
Quality of Supervision_{r,t}
\end{cases}
\begin{cases}
Appeals Shock_{r,t} \\
Successful Appeals Shock_{r,t} + \\
Unsuccessful Appeals Shock_{r,t}
\end{cases}$$
(6)

 $\beta_2 Controls_{r,t} + \varepsilon_{r,d,t}$ 

As reported in Table 5, the results are generally consistent with each implication. The estimated coefficient on the treatment is positive in all 15 models, statistically significant in 11, and close to significance in the remaining four. All coefficients on *Successful Appeals Shock*<sub>*r*,*t*-1</sub> are significant at the 1% level and larger than those on *Appeals Shock*<sub>*r*,*t*-1</sub>. In terms of size, both sets of estimates exceed the main treatment effect: the previous filing of a local appeal is associated with an average increase in *Project Success*<sub>*r*,*d*,*t*</sub> of 0.54, while the previous filing of a

<sup>&</sup>lt;sup>47</sup> https://ieg.worldbankgroup.org/sites/default/files/Data/reports/ieg-wb-project-performance-ratingscodebook\_092015.pdf, p. 18.

<sup>&</sup>lt;sup>48</sup> In all analyses in this subsection, donor and year fixed effects are dropped (since there is only one donor and the treatment is year-specific) and standard errors are clustered by recipient country only.

<sup>&</sup>lt;sup>49</sup> Of the 71 ATI appeals cases in the World Bank's online registry, 42 cases concern single-country projects and two cases concern multicountry (regional) projects. These projects were conducted in 23 unique countries. If a given project is the subject of multiple appeals in different years, we code *Appeals Shock*<sub>t-1</sub> using the earliest date. In Appendix Section 8, we show that the results of the below tests are robust to recoding the treatment as a count variable that measures the cumulative number of appeals/successful appeals/unsuccessful appeals in *r* as of *t* – 1.

Table 5. Analysis of World Bank Project Design, Supervision, and Success

O	utcome:	Project	Preparation	Supervision	Quality at	Quality of
		Success	Cost Ratio	Cost Ratio	Entry	Supervision
		(1)	(2)	(3)	(4)	(5)
Appeals Shock <sub>r,t-1</sub>		0.544***	0.004	0.010**	0.303***	0.193
		(0.117)	(0.003)	(0.004)	(0.113)	(0.158)
Observations		8,816	2,736	2,640	6,271	6,830
		(6)	(7)	(8)	(9)	(10)
Successful Appeals Shock <sub>r,t-1</sub>		1.170***	0.010**	0.022***	0.891***	0.841***
		(0.188)	(0.004)	(0.007)	(0.313)	(0.114)
Observations		8,816	2,736	2,640	6,271	6,830
		(11)	(12)	(13)	(14)	(15)
Unsuccessful Appeals Shock <sub>r,t-1</sub>		0.493***	0.004	0.010**	0.225**	0.150
		(0.111)	(0.003)	(0.004)	(0.095)	(0.160)
Observations		8,816	2,736	2,640	6,271	6,830
Recipient Country Fixed Effects		Yes	Yes	Yes	Yes	Yes
Controls		Yes	Yes	Yes	Yes	Yes

OLS estimates with robust standard errors, clustered by recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub> and Recipient Aid/GNI<sub>t-1</sub>. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

successful local appeal is associated with an increase of 1.17. These effects correspond to shifts from the 50<sup>th</sup> percentile of *Project Success*<sub>*r,d,t*</sub> to between the 61<sup>st</sup> and the 90<sup>th</sup> percentiles.

Taken together, these findings support the logic of Figure 3: local appeals strengthen the positive association between ATI policies with independent appeals mechanisms and project performance, and successful local appeals strengthen this relationship still further. Put differently, the treatment effect accrues disproportionately to projects that are likely to be subject to increased external scrutiny, suggesting that shadow-of-the-future effects are a key channel through which well-enforced ATI policies improve project outcomes.

#### **Does Recipient Country Context Matter?**

Finally, this section briefly examines three additional implications of the argument that concern the broader societal and institutional context in which projects are undertaken. First, well-

enforced ATI policies should have a stronger association with project success when recipient countries are characterized by higher levels of bottom-up accountability, particularly in the form of civic engagement and media freedom. Informed and engaged citizens should be more likely to learn about and utilize ATI policies; activate independent appeals mechanisms when ATI requests are denied; and leverage information from successful requests or appeals to pressure donors and recipient governments to address project problems.<sup>50</sup> Second, the bigger association between strong appeals policies and higher levels of project performance should be stronger when the size of the problem better information might be used to address (e.g. malfeasance in the execution of projects) is more frequent, and when countries' alternative channels of oversight are weaker (increasing the returns to be had from a strong ATI's additional channel). Thus we anticipate the effect will be stronger in countries where corruption is more widespread and those where the rule of law is weaker. Third, inasmuch as we hypothesize – and the primary findings support – that ATI policies without appeals mechanisms are of limited effectiveness, we expect the logic above to apply to ATI policies with appeals, but not ATI policies more generally.

To test these implications, we include in Equations 1 and 2 interactions between the treatment and four recipient country-level variables: a measure of popular involvement in CSOs from the Varieties of Democracy dataset (Coppedge et al. 2018); a composite index of media freedom from the Freedom of the Press dataset (Freedom House 2017); and indices of the rule of law and of control of corruption from the Worldwide Governance Indicators database.<sup>51</sup> The results are consistent with all three implications. As shown in Figure 3 (where the top row depicts findings for all ATI policies, the bottom row ATI policies with appeals mechanisms), the estimated marginal effect of *ATI Policy*<sub>*d*,*t*-1</sub> on *Project Success*<sub>*r*,*d*,*t*</sub> remains small and statistically indistinguishable from 0 across all levels of the four moderators.<sup>52</sup> The effect of *Appeals Mechanism*<sub>*d*,*t*-1</sub>, in contrast, increases with the two measures of bottom-up accountability, attaining significance only at high values of these variables, and decreases with the two measures of governance quality, attaining significance only at low values of these variables. On

<sup>&</sup>lt;sup>50</sup> As Buntaine (2015, 101) argues, "Monitoring should be more likely when civil society groups have overcome barriers to collective action and established organizational means to collect and disseminate information...Under threat of repression, civil society groups may not provide monitoring, choosing instead less confrontational approaches to dealing with [donors]."

<sup>-</sup> https://info.worldbank.org/governance/wgi/. The four variables are described in more detail in Appendix 1.

<sup>&</sup>lt;sup>52</sup> Underlying regression results for the figure are reported in Section 9 of the Appendix.

**Figure 3.** Marginal Effects of Treatment Variables on *Project Success* at Varying Levels of Bottom-Up Accountability and Governance Quality



Dashed lines represent 95% confidence intervals.

average, *Appeals Mechanism*<sub>d,t-1</sub> is associated with a rise in *Project Success*<sub>r,d,t</sub> of 0.18 at the minimum values of the two former measures and of 0.39 at the maximum values; and a rise of 0.33 at the minimum values of the two latter variables and of almost exactly 0 at the maximum values. In sum, there is evidence that the performance effects of ATI policies with appeals mechanisms – but not of ATI policies in general – are moderated in nuanced and significant ways by local contextual factors.

#### **Discussion and Conclusion**

Our empirical analysis offers a window into understanding whether and under what conditions ATI policies improve the performance of public institutions. Foreign aid projects, which are typically delivered by recipient country governments, are just one instantiation of socioeconomically consequential outputs produced by such institutions. However, they offer an attractive set of cases for examining the effects of transparency on performance outcomes. The introduction of ATI policies by donor agencies is plausibly exogenous to the country contexts in which projects take place. Moreover, the staggered nature of adoption across donors, together with the temporal breadth of projects in our dataset, makes it possible to control for a wide range of potentially confounding factors by comparing pre- and post-adoption trends in project performance.

Foreign aid projects are certainly atypical in some respects. The intended beneficiaries are not taxed for the goods and services they receive; nor do they typically have voice, vote, or jurisdictional exit options when projects adversely affect them. Indeed, few mechanisms are available for holding donors accountable if projects harm local communities and ecosystems, fail to achieve stated objectives, or violate host government rules and regulations (Whittle 2013). These unfavorable conditions cause many projects to falter during their design or implementation phase (Easterly 2007; Ensminger and Leder-Luis 2018; Findley et al. 2017; Winters 2014). Yet our findings suggest that even in these circumstances, ATI policies can help to repair the broken feedback loop between public institutions and intended beneficiaries by reducing information asymmetries within the multiple principal-agent chains that connect them. Critically, however, this fix requires more than the mere right to request information from public institutions: we find no evidence that the adoption of ATI policies alone leads to better project outcomes. Instead, we only identify a positive treatment effect when ATI policies are accompanied by recourse mechanisms that allow information seekers to appeal rejected requests through an independently managed process—a process that, in effect, prevents bureaucrats from avoiding compliance with valid ATI requests that could expose ineffective performance.

Micro-level evidence of the consequences of ATI appeals at the recipient country level is consistent with the logic of our shadow of the future argument: the filing of local appeals is associated with a sizable improvement in project success as well as design and supervision, while successful appeals cases are associated with an even larger improvement. These findings suggest that as the level of expected scrutiny received by projects increases, so too does the effort and resources that public officials put into designing and implementing them. We also find that ATI policies with appeals mechanisms have a greater impact on project performance when recipient countries have higher levels of civic engagement and press freedom. The more robust these forms of bottom-up accountability, the more likely citizens are to take advantage of ATI policies and appeals processes and to pressure donors and recipient governments to respond to evidence of poor performance. These policies also have a stronger impact when recipient countries have weaker preexisting rule of law and greater corruption, which increase the added value of a new institutional avenue through which to acquire project information.

Perhaps the most significant policy implication of our analysis is that the design and context of transparency interventions matter. In the absence of reliable mechanisms for preventing bureaucratic noncompliance, efforts to institutionalize transparency in public institutions are unlikely to meaningfully alter their performance. In particular, the findings show that mechanisms for collecting, evaluating, and addressing complaints from stakeholders can serve as a potent instrument for deterring noncompliance - an instrument that harnesses the benefits of both bottom-up monitoring and top-down enforcement. Our findings regarding the moderating effects of civil society robustness and press freedom are harder to translate into specific policy actions, given the close connection between these variables and more fundamental country characteristics such as democracy, respect for human rights, and the rule of law. They do, however, imply that policymakers can help to "clear the pathway" between well-enforced transparency policies and better performance outcomes by providing channels through which stakeholders-even in settings with limited civic engagement and freedom of expression - can draw attention to performance problems. In other words, they suggest that there are systematic features of design and context that can help practitioners pave a middle ground between treating all ATI policies the same and treating each case as sui generis. ATI policies can and do matter in improving real-world outcomes; that impact is conditional on the institutional features of the policy itself and the pre-existing broader information environment to which they are added.

#### References

Ackerman, John and Irma Sandoval-Ballesteros. 2006. "The Global Explosion of Freedom of Information Laws." *Administrative Law Review* 58 (1): 85–130.

Anderson, Sarah, Mark Buntaine, Mengdi Liu, and Bing Zhang. 2019. "Non-Governmental Monitoring of Local Governments Increases Compliance with Central Mandates: A National-Scale Field Experiment in China." *American Journal of Political Science* 63 (3): 626-643.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Aronow, Peter M., and Cyrus Samii. 2016. "Does Regression Produce Representative Estimates of Causal Effects?" *American Journal of Political Science* 60 (1): 250-267.

Banerjee, Abhijit V., Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani. 2010. "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy* 2 (1): 1-30.

Banisar, David. 2006. "Freedom of Information around the World 2006: A Global Survey of Access to Government Information Laws." Privacy International. http://www.freedominfo.org/documents/global\_survey2006.pdf.

Berliner, Daniel, and Aaron Erlich. 2015. "Competing for Transparency: Political Competition and Institutional Reform in Mexican States." *American Political Science Review* 109 (1): 110-128.

Berliner, Daniel. 2017. "Sunlight or Window Dressing? Local Government Compliance with South Africa's Promotion of Access to Information Act." *Governance* 30 (4): 641-661.

Berliner, Daniel, Benjamin E. Bagozzi, and Brian Palmer-Rubin. 2018. "What Information Do Citizens Want? Evidence from One Million Information Requests in Mexico." *World Development* 109: 222-235.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics* 119 (1): 249-275.

Bratton, Michael. 2012. "Citizen Perceptions of Local Government Responsiveness in Sub-Saharan Africa." *World Development* 40 (3): 516-527.

Briggs, Ryan. 2019. "Why Does Aid Not Target The Poorest?" AidData Working Paper #85. Williamsburg, VA: AidData at William & Mary. https://www.aiddata.org/publications/why-doesaid-not-target-the-poorest.

Buntaine, Mark T. 2015. "Accountability in Global Governance: Civil Society Claims for Environmental Performance at the World Bank." *International Studies Quarterly* 59 (1): 99-111.

Buntaine, Mark. 2016. *Giving Aid Effectively: The Politics of Environmental Performance and Selectivity at Multilateral Development Banks*. Oxford: Oxford University Press.

Buntaine, Mark T., Bradley C. Parks, Benjamin P. Buch. 2017. "Aiming at the Wrong Targets: The Difficulty of Improving Domestic Institutions with International Aid." *International Studies Quarterly* 61 (2): 471-488.

Carlson, Elizabeth, and Brigitte Seim. 2018. "Honor among Chiefs: An Experiment on Monitoring and Diversion Among Traditional Leaders in Malawi." AidData Working Paper #62 Williamsburg, VA: AidData at William & Mary. http://docs.aiddata.org/ad4/pdfs/WPS62\_Honor\_Among\_Chiefs.pdf.

Cleary, Matthew R. 2007. "Electoral Competition, Participation, and Government Responsiveness in Mexico." *American Journal of Political Science* 51 (2): 283-299.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, ..., and Daniel Ziblatt. 2018. "V-Dem [Country-Year/Country-Date] Dataset v8." Varieties of Democracy (V-Dem) Project. https://doi.org/10.23696/vdemcy18. Cordis, Adriana S. and Patrick L. Warren. 2014. "Sunshine as Disinfectant: The Effect of State Freedom of Information Act Laws on Public Corruption." *Journal of Public Economics* 115: 18-36.

Costa, Samia. 2013. "Do Freedom of Information Laws Decrease Corruption?" *The Journal of Law, Economics, and Organization* 29(6); 1317–1343.

Dal Bó, Ernesto, Frederico Finan, Nicholas Y. Li, Laura Schechter. 2018. "Government Decentralization Under Changing State Capacity: Experimental Evidence from Paraguay." NBER Working Paper No. 24879. https://www.nber.org/papers/w24879.

Denizer, Cevdet, Daniel Kaufmann, and Aart Kraay. 2013. "Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance." *Journal of Development Economics* 105: 288–302.

Distelhorst, Greg. 2017. "The Power of Empty Promises: Quasi-Democratic Institutions and Activism in China." *Comparative Political Studies* 50 (4): 464-498.

Dreher, Axel, Stephan Klasen, James Raymond Vreeland, and Eric Werker. 2013. "The Costs of Favoritism: Is Politically Driven Aid Less Effective?" *Economic Development and Cultural Change* 62 (1): 157-191.

Easterly, William. 2007. "Was Development Assistance a Mistake?" *American Economic Review* 97 (2): 328-332.

Ensminger, Jean, and Jetson Leder-Luis. 2018. "Measuring Strategic Data Manipulation: Evidence from a World Bank Project." Working Paper. http://jee.caltech.edu/documents/1255/Ensminger\_and\_Leder-Luis.pdf.

Escaleras, Monica, Shu Lin, and Charles Register. 2009. "Freedom of Information Acts and Public Sector Corruption." *Public Choice* 145 (3–4): 435-460.

Faust, Jörg. 2011. "Donor transparency and aid allocation." DIE Discussion Paper 12/2011. Bonn: German Development Institute (DIE). http://edoc.vifapol.de/opus/volltexte/2012/3552/pdf/DP\_12.2011.pdf

Findley, Michael G., Adam S. Harris, Helen V. Milner and Daniel L. Nielson. 2017. "Who Controls Foreign Aid? Elite versus Public Perceptions of Donor Influence in Aid-Dependent Uganda." *International Organization* 71 (4): 633-663.

Florini, Ann. 2007. *The Right to Know: Transparency for an Open World*. New York, NY: Columbia University Press.

Fox, Jonathan A. 2015. "Social Accountability: What Does the Evidence Really Say?" *World Development* 72: 346-361.

Fox, Jonathan A., and Lloyd David Brown, eds. 1998. *The Struggle for accountability: The World Bank, NGOs, and Grassroots Movements*. Cambridge: MIT Press.

Freedom House. 2017. *Freedom of the Press 2017*. London: Freedom House. https://freedomhouse.org/sites/default/files/FOTP\_2017\_booklet\_FINAL\_April28.pdf

Goodman-Bacon, Andrew. *Difference-in-Differences with Variation in Treatment Timing*. 2018. NBER Working Paper No. 25018. https://www.nber.org/papers/w25018.

Grigorescu, Alexandru. 2007. "Transparency of Intergovernmental Organizations: The Roles of Member States, International Bureaucracies and Nongovernmental Organizations." *International Studies Quarterly* 51 (3): 625-648.

Grossman, Guy, and Kristin Michelitch. 2018. "Information Dissemination, Competitive Pressure, and Politician Performance between Elections: A Field experiment in Uganda." *American Political Science Review* 112 (2): 280-301.

Hazell, Robert and Ben Worthy. 2010. "Assessing the Performance of FOI in Different Countries." *Government Information Quarterly* 27 (4) 352-359.

Holsen, Sarah, and Martial Pasquier. 2012. "Insight on Oversight: The Role of Information Commissioners in the Implementation of Access to Information Policies." *Journal of Information Policy* 2: 214–241.

Honig, Dan. 2019. "When Reporting Undermines Performance: The Costs of Politically Constrained Organizational Autonomy in Foreign Aid Implementation." *International Organization* 73 (1): 171-201.

Honig, Dan. 2018. Navigation by Judgment. New York: Oxford University Press.

Honig, Dan and Catherine Weaver. 2019. "A Race to the Top? The Aid Transparency Index and the Social Power of Global Performance Indicators." *International Organization* 73 (3): 579-610.

Jablonski, Ryan, and Brigitte Seim. 2017. "How Transparency Affects Distributional Politics: A Field Experiment among Elected Incumbents in Malawi." AidData Working Paper #45. Williamsburg, VA: AidData at William & Mary. http://docs.aiddata.org/ad4/pdfs/WPS45-How\_Transparency\_Affects\_Distributional\_Politics.pdf

Kilby, Christopher. 2009. "The Political Economy of Conditionality: An Empirical Analysis of World Bank Loan Disbursements." *Journal of Development Economics* 89 (1): 51-61.

Kilby, Christopher, and Katharina Michaelowa. Forthcoming. "What influences World Bank project evaluations?" In: Nabamita Duua and Claudia Williamson (eds.), *Foreign Aid and Economic Development: Micro and Macro Perspectives*. Palgrave MacMillan.

Kosack, Stephen and Archon Fung. "Does Transparency Improve Governance?" *Annual Review of Political Science* 17 (2014): 65-87.

Lall, Ranjit. 2017. "Beyond Institutional Design: Explaining the Performance of International Organizations." *International Organization* 71 (2): 245-280.

Legovini, Arianna, Vincenzo Di Maro, and Caio Piza. 2015. "Impact Evaluation Helps Deliver Development Projects." World Bank Policy Research Working Paper no. 7157. http://documents.worldbank.org/curated/en/676351468320935363/pdf/WPS7157.pdf.

Lieberman, Evan S., Daniel N. Posner, and Lily L. Tsai. 2014. "Does Information Lead to More Active Citizenship? Evidence from an Education Intervention in Rural Kenya." *World Development* 60: 69–83.

Lindstedt, Catharina, and Daniel Naurin. 2010. "Transparency Is Not Enough: Making Transparency Effective in Reducing Corruption." *International Political Science Review* 31 (3): 301–22.

Lorentzen, Peter, Pierre Landry, and John Yasuda. 2014. "Undermining Authoritarian Innovation: The Power of China's Industrial Giants." *Journal of Politics* 76 (1): 182–194.

Malik, Rabia, and Randall W. Stone. 2018. "Corporate Influence in World Bank Lending." *Journal of Politics*, 80 (1): 103-118.

Mendel, Toby. 1999. The public's right to know: Principles on freedom of information legislation. London, UK: Article 19. Accessed at https://www.article19.org/data/files/pdfs/standards/righttoknow.pdf

Metzger, Laura and Isabel Guenther. 2015. "How to Assess the Effectiveness of Development Aid Projects: Evaluation Ratings versus Project Indicators." *Journal of International Development* 27 (8): 1496-1520.

Nelson, Paul J. 2001. "Transparency Mechanisms at the Multilateral Development Banks." *World Development* 29 (11): 1835-1847.

Neuman, Laura. 2006. "Access to Information Laws: Pieces of the Puzzle – An Analysis of the International Norms." White Paper, The Carter Center. https://www.cartercenter.org/resources/pdfs/peace/americas/ati\_pieces\_of\_puzzle.pdf

Nielson, Daniel and Michael Tierney. 2003. "Delegation to International Organizations: Agency Theory and World Bank Environmental Reform." *International Organization* 57 (2): 241-276.

OECD. 1991. *The DAC Principles for the Evaluation of Development Assistance*. Paris, France: OECD.

OECD. 2000. DAC Criteria for Evaluating Development Assistance Factsheet: 2. Paris, France: OECD.

OECD. 2010. *Quality Standards for Development Evaluation: DAC Guidelines and Reference Series*. Paris, France: OECD.

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

Park, Susan. 2005. "Norm diffusion within international organizations: a case study of the World Bank." *Journal of International Relations and Development* 8 (2): 111-141.

Peisakhin, Leonid. 2012. "Transparency and Corruption: Evidence from India." *Journal of Law and Economics* 55 (1): 129–49.

Phillips, David A. 2009. *Reforming the World Bank: Twenty Years of Trial—and Error*. Cambridge: Cambridge University Press.

Reinikka, Ritva, and Jakob Svensson. 2005. "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda." *Journal of the European Economic Association* 3 (2-3): 259–67.

Vadlamannati, Krishna Chaitanya and Arusha Cooray. 2016a. "Do Freedom of Information Laws Improve Bureaucratic Efficiency? An Empirical Investigation." *Oxford Economic Papers* 64 (4): 968-993. Vadlamannati, Krishna Chaitanya and Arusha Cooray. 2016b. "Transparency Pays: Evaluating the Effects of Freedom of Information Laws on Government Corruption." *Journal of Development Studies* 53 (1): 116-137.

Warner, Andrew. 2010. "Cost-Benefit Analysis in World Bank Projects." Washington, D.C.: World Bank.

Whittle, Dennis. 2013. "How Feedback Loops Can Improve Aid (and Maybe Governance)." Center for Global Development Essay. Washington D.C.: Center for Global Development.

Williams, Martin. 2017. "The Political Economy of Unfinished Development Projects:
Corruption, Clientelism, or Collective Choice?" *American Political Science Review* 111 (4): 705-723.

Winters, Matthew. 2014. "Targeting, Accountability and Capture in Development Projects." *International Studies Quarterly* 58 (2): 393-404.

Worthy, Ben. 2010. "More Open but Not More Trusted? The Effect of the Freedom of Information Act 2000 on the United Kingdom Central Government." *Governance* 23 (4): 561-582.

Worthy, Benjamin. 2013. "'Some are More Open than Others': Comparing the Impact of the Freedom of Information Act 2000 on Local and Central Government in the UK." *Journal of Comparative Policy Analysis* 15 (5): 395-414.

# Appendix for "When Does Transparency Improve Institutional Performance? Evidence from 20,000 Aid Projects in 183 Countries"

# **1. Additional Information on Dataset**

TADIE AL, I UN DESCRIDIUM UN VARIADIES IN DALASE
--

Variable name	Description	Scale	Source
Project Success	Holistic, ex-post rating of project success produced by donor staff, specialized evaluation units, and external evaluators	Ordinal	Authors' original dataset
ATI Policy	Measure of whether a donor possesses an ATI policy	Binary	Authors' original dataset
Appeals Mechanism	Measure of whether a donor possesses an ATI policy with an independent appeals process for denied information requests	Binary	Authors' original dataset
Recipient GDP per Capita Growth	Recipient country's GDP per capita growth rate	Continuous (percentage)	World Development Indicators dataset
Recipient Log GDP per Capita	Log of recipient country's GDP per capita (in millions of US dollars)	Continuous (logarithmic)	World Development Indicators dataset
Recipient Aid/GNI	Net official development assistance provided to a recipient country as a percentage of its gross national income	Continuous (ratio)	World Development Indicators dataset
CSO Participatory Environment	Recipient country-level measure of popular involvement in civil society organizations	Ordinal, converted to interval by the measurement model	Varieties of Democracy dataset
Total Press Freedom	Recipient country-level measure of press freedom based on laws and regulations, political pressures, economic factors, and repressive actions (e.g., violence, harassment, censorship) that influence media content.	Continuous	Freedom of the Press dataset
Control of Corruption	Recipient country-level measure of control of corruption, capturing perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of	Continuous	Worldwide Governance Indicators dataset

	corruption, as well as "capture" of the state by elites and private interests		
Rule of Law	Recipient country-level measure of the rule of law, capturing perceptions of confidence in and compliance with the rules of society, in particular the quality of contract enforcement, property rights, the police, and the courts	Continuous	Worldwide Governance Indicators dataset

Donor	Observations	Mean	Std. Dev.	Max	Min	Range
AfDB	684	4.14	0.71	6.00	1.95	4.05
AsDB	1,572	3.89	1.05	6.00	1.50	4.50
CDB	21	3.20	0.79	4.80	1.20	3.60
DFAT	610	4.27	0.73	6.00	2.00	4.00
DFID	1,917	4.62	0.97	6.00	1.20	4.80
GEF	1,169	4.32	0.96	6.00	1.00	5.00
GFATM	1,286	4.75	1.25	6.00	1.50	4.50
GiZ	130	4.47	0.87	6.00	2.00	4.00
IFAD	286	4.00	0.80	6.00	2.00	4.00
JICA	716	4.99	1.20	6.00	1.50	4.50
KfW	2,021	4.16	1.11	6.00	1.00	5.00
WB	10,274	4.05	1.23	6.00	1.00	5.00
Overall	20,686	4.20	1.17	6.00	1.00	5.00

# Table A2. Project Success (Outcome Variable) by Donor

# Table A3. Descriptive Statistics on Rated Projects by Donor

Donor	No.	No. multi-	No.	Mean	Mean	Start	End
	projects	country	unique	project	project	year	year
		projects	countries	rating	length		
AfDB	684	55	52	4.14		1988	AfDB
AsDB	1572	0	41	3.89	7.97	1968	AsD B
CDB	21	2	11	3.20		1997	CDB
DFAT	610	154	23	4.27	6.05	1988	DFA T
DFID	1917	0	99	4.62		1987	DFID
GEF	1169	290	132	4.32		1992	GEF
GFATM	1286	45	128	4.75	4.56	2003	GFA TM
GiZ	130	0	67	4.47	7.74	1989	GiZ

IFAD	286	0	86	4.00		1988	IFAD
JICA	716	0	86	4.99	5.78	1981	JICA
KfW	2021	0	108	4.16	7.95	1963	KfW
WB	10274	195	162	4.05	5.99	1956	WB

# 2. Testing Parallel Trends Assumption

**Figure A1.** Results of Equation 2 with Leads, Lags, and Contemporaneous Measure of *Appeals Mechanism* 



Table A4. Baseline Results Controlling for Recipient Country-Specific Trends

Outcome: Project Success <sub>t</sub>	(1)	(2)	(3)
ATI Policy <sub>t-1</sub>	-0.039		-0.148**
	(0.094)		(0.065)
Appeals Mechanism $t_{-1}$		0.305***	0.380***
		(0.087)	(0.093)
Observations	18,489	18,489	18,489
R-squared	0.318	0.319	0.320
Recipient Country-Specific Trends	Yes	Yes	Yes

Recipient Country Fixed Effects	Yes	Yes	Yes
Donor Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
OLS estimates with robust standard errors, clu	ustered by donor	and recipie	nt
country, in parentheses.			
*** p<0.01, ** p<0.05, * p<0.1			

## 3. Validating Outcome Measure

Table A5. Baseline Results with Malik and Stone Measure of Project Performance

Outcome: Project Success <sub>t</sub>	(1)
Appeals Mechanism <sub>t-1</sub>	0.290***
	(0.093)
Observations	0.000
Observations	3,296
R-squared	0.088
Controls	Yes
Recipient Country Fixed Effects	Yes
OLS estimates with robust standard errors, clustered by do	nor and
recipient country, in parentheses. Controls are Recipient GI	DP per
Capita Growth <sub>t-1</sub> , Recipient Log GDP per Capita <sub>t-1</sub> , and Rec	ipient
Aid/GNI <sub>t-1</sub> .	-
*** p<0.01, ** p<0.05, * p<0.1	

Table A6.         Baseline Results with	Sample Restricted to Years w	vith Independent Evaluation Unit
---	------------------------------	----------------------------------

Outcome: Project Success <sub>t</sub>	(1)	(2)	(3)
ATI Policy <sub>t-1</sub>	0.099		0.072
	(0.157)		(0.084)
Appeals Mechanism <sub>t-1</sub>		0.369**	0.361**
		(0.148)	(0.124)
Observations	11 771	11 771	11 771
Observations	11,771	11,771	11,771
R-squared	0.151	0.153	0.153
Controls	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub>, and Recipient Aid/GNI<sub>t-1</sub>.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

#### 4. Assessing Inferential Leverage

Table A7. Bacon Decomposition of Coefficient on Appeals Mechanism<sub>t-1</sub> in Equation 2

Two-period difference-in-differences estimator	Beta	Total Weight
Treated in different years	0.332	0.414
Treated in given year vs. always treated	0.57	0.085
Treated in given year vs. never treated	0.471	0.481
Always treated vs. never treated	4.175	0.001
Within variation (due to controls)	-1.683	0.013

Output of bacondecomp command in Stata. The command requires collapsing the data into a donor-year panel (the aggregate coefficient estimate does not substantively change).



Figure A2. Scatterplot of Bacon Decomposition Results

Always vs never treated = 4.1747313 (weight = .01255245) Within component = -1.682611 (weight = .01255245)

Output of bacondecomp command in Stata.

Table A8. Effective Samples for Equations 1 and 2

	Nor	ninal sar	nple	Effe E	Effective sample, Equation 1			Effective sample, Equation 2		
			Std			Std			Std	
	Obs	Mean	Dev.	Obs	Mean	Dev.	Obs	Mean	Dev.	
Project Success <sub>r,d,t</sub>	20686	4.20	1.17	17,929	4.46	1.11	17,929	4.38	1.14	
ATI Policy <sub>d,t-1</sub>	21301	0.48	0.50	18,339	0.71	0.45	18,339	0.48	0.50	
Appeals Mechanism <sub>d,t-1</sub> Becipient GDP per	21301	0.15	0.36	18,339	0.48	0.50	18,339	0.17	0.38	
Capita Growth <sub>r,t-1</sub>	23963	2.80	6.01	18,330	3.28	4.58	18,330	3.16	5.41	
GDP per Capita <sub>r,t-1</sub>	23886	7.57	1.31	18,330	7.20	0.97	18,330	7.14	0.95	
Aid/GNI <sub>r,t-1</sub>	20932	7.10	10.76	18,202	7.29	10.49	18,202	7.35	9.96	

The effectives samples are calculated using Aronow and Samii's (2016) multiple regression weights procedure.

# 5. Baseline Results with Alternative Sample Restrictions

**Table A9.** Equation 2 Results with Sample Restricted to Five-Year Window around Treatment

 Adoption

Outcome: Project Success <sub>t</sub>	(1)
Appeals Mechanism <sub>t-1</sub>	0.293*** (0.054)
Observations	1,734
R-squared	0.264
Controls	Yes
Recipient Country Fixed Effects	Yes
Donor Fixed Effects	Yes
Year Fixed Effects	Yes
OLS estimates with robust standard errors, clustered by a recipient country, in parentheses. Controls are Recipient Capita Growth <sub>t-1</sub> , Recipient Log GDP per Capita <sub>t-1</sub> , and Re Aid/GNI <sub>t-1</sub> .	donor and GDP per ecipient
*** p<0.01, ** p<0.05, * p<0.1	

**Table A10.** Equation 2 Results with Sample Restricted to Projects Starting Close to and BeforeTreatment Adoption

Outcome: Project Success <sub>t</sub>	(1)	(2)	(3)	(4)	(5)
Appeals Mechanism <sub>t-1</sub>	0.208	0.269	0.277*	0.301**	0.340**

	(0.134)	(0.150)	(0.136)	(0.121)	(0.118)
Observations	13,731	14,015	14,251	14,480	14,720
R-squared	0.111	0.116	0.118	0.119	0.123
Recipient Country Fixed Effects	Yes	Yes	Yes	Yes	Yes
Donor Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Proiects Started Before Year <i>i</i> +	1	2	3	4	5

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses. Year *j* is the date of treatment adoption. Controls are Recipient GDP per Capita Growth<sub>*t*-1</sub>, Recipient Log GDP per Capita<sub>*t*-1</sub>, and Recipient Aid/GNI<sub>*t*-1</sub>. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

#### Table A11. Equation 2 Results with Sample Restricted to Recent Years

Outcome: Project Success <sub>t</sub>	(1)	(2)	(3)
Appeals Mechanism <sub>t-1</sub>	0.260***	0.272***	0.111***
	(0.074)	(0.080)	(0.032)
Observations	12,852	10,865	8,093
R-squared	0.153	0.159	0.193
Controls	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes
Sample period	1990-2016	1995-2016	2000-2016

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub>, and Recipient Aid/GNI<sub>t-1</sub>.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

#### 6. Placebo Test

It is possible that the main results are driven not by the presence not of an independent appeals mechanism per se, but of other ATI policy features that encourage disclosure. To explore this possibility, we re-estimate Equation 2 substituting *Appeals Mechanism*<sub>t-1</sub> for a dummy for whether an ATI policy contains a presumption of disclosure provision, that is, a provision stating that institutional information will be disclosed absent a compelling reason not to do so. As reported in Table A10, the presence of such a provision is neither positively nor strongly associated with project performance, suggesting that *Appeals Mechanism*<sub>t-1</sub> is not merely as a proxy for the general stringency of a given ATI policy.

**Table A12.** Baseline Results with Placebo Treatment (ATI Policy with Presumption of Disclosure Provision)

Outcome: Project Success <sub>t</sub>	(1)	(2)
ATI Policy <sub>t-1</sub>		0.511***
		(0.102)
Presumption of Disclosure <sub>t-1</sub>	-0.080	-0.549***
	(0.086)	(0.099)
Observations	17 020	17 929
R-squared	0.131	0.133
Recipient Country Fixed Effects	Yes	Yes
Donor Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Controls	Yes	Yes

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub>, and Recipient Aid/GNI<sub>t-1</sub>.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

# 7. Additional Robustness Checks

#### Additional Controls

#### Table A13. Baseline Results Controlling for Project Size

Outcome: Project Success <sub>t</sub>	(1)	(2)	(3)
ATI Policy <sub>t-1</sub>	0.025		-0.003
	(0.118)		(0.072)
Appeals Mechanism <sub>t-1</sub>		0.383**	0.384**
		(0.141)	(0.135)
	44.047	44.047	44.047
Observations	11,94 <i>1</i>	11,94 <i>1</i>	11,947
R-squared	0.161	0.164	0.164
Recipient Country Fixed Effects	Yes	Yes	Yes
Donor Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls (Including Project Size)	Yes	Yes	Yes

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub>, and Recipient Aid/GNI<sub>t-1</sub>.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

#### Table A14. Baseline Results Controlling for Sector Fixed Effects

Outcome: Project Success <sub>t</sub>	(1)	(2)	(3)			
ATI Policy <sub>t-1</sub>	-0.045		-0.076			
	(0.095)		(0.057)			
Appeals Mechanism		0.327**	0.342*			
		(0.147)	(0.159)			
		( )	( )			
Observations	15,760	15,760	15,760			
R-squared	0.157	0.158	0.158			
Recipient Country Fixed Effects	Yes	Yes	Yes			
Donor Fixed Effects	Yes	Yes	Yes			
Year Fixed Effects	Yes	Yes	Yes			
Sector Fixed Effects	Yes	Yes	Yes			
Controls	Yes	Yes	Yes			
OLS astimates with reduct standard errors, elustered by denor and recipient country, in						

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub>, and Recipient Aid/GNI<sub>t-1</sub>.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

# **Table A15.** Baseline Results Controlling for Recipient Country UN Security Council Membership

Outcome: Project Success <sub>t</sub>	(1)	(2)	(3)
ATI Policy <sub>t-1</sub>	0.005		-0.075
	(0.088)		(0.059)
Appeals Mechanism <sub>t-1</sub>		0.278***	0.310***
		(0.073)	(0.080)
Observations	10,000	10,000	10,000
Observations	16,922	16,922	16,922
R-squared	0.124	0.126	0.126
Recipient Country Fixed Effects	Yes	Yes	Yes
Donor Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls (Including UNSC Membership)	Yes	Yes	Yes

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub>, and Recipient Aid/GNI<sub>t-1</sub>.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

# Disaggregation by Donor

Table A16. Equation 2 Results with Outcome Variable Disaggregated by Don
--

Outcome: Project													
Success <sub>t</sub>	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Appeals Mechanism <sub>t-1</sub>	0.290***	0.413**				0.126**			-0.052		0.700***	0.219	0.290***
••	(0.081)	(0.171)				(0.063)			(0.153)		(0.147)	(0.209)	(0.081)
	· · · ·	, ,				· · ·			· ,		, ,	, ,	( )
Observations	17,929	597	1,426	13	447	1,652	772	1,160	93	256	668	1,844	17,929
R-squared	0.133	0.286	0.122	0.801	0.078	0.104	0.143	0.239	0.567	0.400	0.214	0.156	0.133
Donor	All	AfDB	AsDB	CDB	DFAT	DFID	GEF	GFATM	GiZ	IFAD	JICA	KfW	WB
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Recipient Country Fixed													
Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Donor Fixed Effects	Yes	No	No	No	No	No	No	No	No	No	No	No	No
Year Fixed Effects	Yes	No	No	No	No	No	No	No	No	No	No	No	No

OLS estimates with robust standard errors, clustered by recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub>, and Recipient Aid/GNI<sub>t-1</sub>. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

# Reparametrizing Outcome Variable

In the main analysis, we use the full richness of project success ratings, following Honig's (2018, 2019) approach. However, another strand of the literature on aid effectiveness operationalizes these ratings as a binary variable (e.g., Denizer et al. 2013; Dollar and Svensson 2000; Dreher et al. 2013; Kilby 2009). The most common method is to classify projects as "successful" if their rating is greater than or equal to the sample median and "unsuccessful" if it is not. We assess the sensitivity of Equation 2's results to three separate classification thresholds: the sample's mean rating (Model 1), median rating (Model 2), and maximum rating (Model 3). Table A14 estimates each model logistic regression. The results are robust to all three reparameterizations with both estimation methods.

**Table A17.** Equation 2 Results with Binary Outcome Measures (Estimated with Logistic Regression)

Outcome: Binary measure of Project Success <sub>t</sub>	(1)	(2)	(3)
Appeals Mechanism <sub>t-1</sub>	0.776***	0.738***	0.932***
	(0.109)	(0.115)	(0.159)
Constant	1.835*	-0.724	-4.709**
	(1.022)	(0.796)	(1.878)
Observations	17,889	17,867	16,803
Controls	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes
Project Success <sub>t</sub> = 1 if value $\geq$	Mean	Median	Maximum

Logistic regression with robust standard errors, clustered by recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub>, and Recipient Aid/GNI<sub>t-1</sub>. \*\*\* p < 0.01, \*\* p < 0.05, \* p < 0.1

# Varying Lag Structures

We also experiment with different lag structures for *Appeals Mechanism*<sub>t-1</sub>. The results of Equation 2 are not sensitive to our choice of a 1-year lag; indeed, as shown in Table A14, the results are even *stronger* with 2-, 3-, 4-year, and 5-year lags.

# Table A18. Equation 2 Results with Varying Treatment Lags

Outcome: Project Success <sub>t</sub>	(1)	(2)	(3)	(4)
	0 000+++			
Appeals Mechanism <sub>t-2</sub>	0.289^^^			

	(0.072)			
Appeals Mechanism <sub>t-3</sub>		0.264*** (0.083)		
Appeals Mechanism <sub>t-4</sub>			0.291*** (0.072)	
Appeals Mechanism <sub>t-5</sub>				0.281*** (0.069)
Observations	17,698	17,439	17,137	16,817
R-squared	0.132	0.132	0.133	0.133
Controls	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub>, and Recipient Aid/GNI<sub>t-1</sub>.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

## Retaining Original Project Ratings for Project Success

 Table A19.
 Baseline Equations with Original Project Ratings (Not Converted to Common Scale)

Outcome: Project Success <sub>t</sub>	(1)
Annaela Machaniam	0.040***
Appeals Mechanism <sub>t-1</sub>	0.242 (0.058)
	,
Observations	20,363
R-squared	0.239
Recipient Country Fixed Effects	Yes
Donor Fixed Effects	Yes
Year Fixed Effects	Yes
Controls	Yes
OLS estimates with robust standard errors, clustered by d	onor and
recipient country, in parentheses. Controls are Recipient C	GDP per
Capita Growth <sub>t-1</sub> , Recipient Log GDP per Capita <sub>t-1</sub> , and Re	cipient
Aid/GNI <sub>t-1</sub> .	
*** p<0.01, ** p<0.05, * p<0.1	

#### Alternative Coding of Independent Appeals Mechanism

As mentioned in fn. 37 of the main text, the Aid Transparency Index codes the World Bank as possessing an ATI policy *without* an independent appeals mechanism due to limitations in the latter's coverage. While this decision runs counter to the conventional wisdom amongst aid

practitioners, who generally consider the World Bank's appeals mechanism as a robust one, Table A18 indicates that our results are robust to it.

**Table A20.** Results of Equation 2 with Alternative Coding of Appeals Mechanism for WorldBank

Outcome: Project Success <sub>t</sub>	(1)
Appeals Mechanism <sub>t-1</sub>	0.414***
	(0.100)
Observations	17,929
R-squared	0.134
Controls	Yes
Fixed Effects	Yes
OLS estimates with robust standard errors, clustered by	donor and
recipient country, in parentheses. Controls are Recipient	GDP per
Capita Growth <sub>t-1</sub> , Recipient Log GDP per Capita <sub>t-1</sub> , and Re	ecipient
Aid/GNI <sub>t-1</sub> .	
*** p<0.01, ** p<0.05, * p<0.1	

Alternative Estimation of Standard Errors

# Table A21. Results of Equations 1-3 with Bootstrapped Standard Errors

Outcome: Project Success <sub>t</sub>	(1)
Appeals Mechanism <sub>t-1</sub>	0.290***
	(0.034)
Observations	17,929
R-squared	0.133
Controls	Yes
Recipient Country Fixed Effects	Yes
Donor Fixed Effects	Yes
Year Fixed Effects	Yes
OLS estimates with bootstrapped standard errors in pare	ntheses.
Controls are Recipient GDP per Capita Growtht-1, Recipie	nt Log GDP
per Capita <sub>t-1</sub> , and Recipient Aid/GNI <sub>t-1</sub> .	

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

#### Table A22. Results of Equations 1-3 with Standard Errors Clustered on Donor

Outcome: Project Success <sub>t</sub>	(1)	(2)	(3)			
ATI Policy <sub>t-1</sub>	0.020		-0.067			
	(0.102)		(0.072)			
Appeals Mechanism <sub>t-1</sub>		0.290***	0.320***			
		(0.083)	(0.086)			
Observations	17,929	17,929	17,929			
R-squared	0.131	0.133	0.133			
Recipient Country Fixed Effects	Yes	Yes	Yes			
Donor Fixed Effects	Yes	Yes	Yes			
Year Fixed Effects	Yes	Yes	Yes			
Controls	Yes	Yes	Yes			
OLS estimates with robust standard errors, clustered b	y donor, in pa	arentheses. C	ontrols are			
OLS estimates with robust standard errors, clustered by donor, in parentneses. Controls are						

Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub>, and Recipient Aid/GNI<sub>t-1</sub>. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table A23.** Results of Equations 1-3 with Standard Errors Clustered on Donor × Recipient

 Country

Outcome: Project Success <sub>t</sub>	(1)	(2)	(3)
ATI Policy <sub>t-1</sub>	0.020		-0.067
	(0.049)		(0.050)
Appeals Mechanism <sub>t-1</sub>		0.290***	0.320***
		(0.055)	(0.059)
Observations	17,929	17,929	17,929
R-squared	0.131	0.133	0.133
Recipient Country Fixed Effects	Yes	Yes	Yes
Donor Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes

OLS estimates with robust standard errors, clustered by donor\*recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub>, and Recipient Aid/GNI<sub>t-1</sub>.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

# 8. World Bank Sub-Analysis with Alternative Treatment

**Table A24.** Analysis of World Bank Project Design, Supervision, and Success with Treatments as Count Variables (see fn. 57)

Outcome:	Project Success	Preparation Cost Ratio	Supervision Cost Ratio	Quality at Entry	Quality of Supervision
	(1)	(2)	(3)	(4)	(5)
Cumulative Prior No. Appeals <sub>r,t-1</sub>	0.318***	0.003*	0.006***	0.156***	0.135*
	(0.063)	(0.001)	(0.002)	(0.054)	(0.076)
Observations	8,816	2,736	2,640	6,271	6,830
	(6)	(7)	(8)	(9)	(10)
Cumulative Prior No. Successful					
Appeals <sub>r,t-1</sub>	1.170***	0.010**	0.022***	0.891***	0.833***
	(0.188)	(0.004)	(0.007)	(0.313)	(0.131)
Observations	8,816	2,736	2,640	6,271	6,830
	(11)	(12)	(13)	(14)	(15)
Cumulative Prior No.					
Unsuccessful Appeals <sub>r,t-1</sub>	0.341***	0.003	0.007**	0.144**	0.161
	(0.077)	(0.002)	(0.003)	(0.065)	(0.113)
Observations	8,816	2,736	2,640	6,271	6,830
Recipient Country Fixed Effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

OLS estimates with robust standard errors, clustered by recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub> and Recipient Aid/GNI<sub>t-1</sub>. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

#### 9. Regression Results Underlying Figure 3

**Table A25.** Relationship Between ATI Policy and Project Success as Moderated by Bottom-Up

 Accountability and Governance Quality

Outcome: Project Success <sub>t</sub>	(1)	(2)	(3)	(4)
ATI Policy <sub>t-1</sub>	0.033 (0.096)	-0.000 (0.187)	0.015 (0.116)	-0.033 (0.126)
CSO Participatory Environment <sub>t-1</sub>	0.036* (0.017)			

ATI Policy <sub>t-1</sub> × CSO Participatory Environment <sub>t-1</sub>	-0.011 (0.033)			
Total Press Score <sub>t-1</sub>		-0.003 (0.002)		
ATI Policy <sub>t-1</sub> × Total Press Score <sub>t-1</sub>		0.003 (0.002)		
Control of Corruption <sub>t-1</sub>			-0.188 (0.141)	
ATI Policy <sub>t-1</sub> × Control of Corruption <sub>t-1</sub>			0.047 (0.074)	
Rule of Law <sub>t-1</sub>				-0.071 (0.105)
ATI Policy <sub>t-1</sub> × Rule of Law <sub>t-1</sub>				-0.042 (0.061)
Observations	17,688	10,732	8,022	8,034
R-squared	0.130	0.157	0.175	0.175
Controls	Yes	Yes	Yes	Yes
Recipient Country Fixed Effects	Yes	Yes	Yes	Yes
Donor Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
OLS actimates with rebust standard arrars, clustered b	v donor and r	opiniont op	intry in nor	onthooo

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub> and Recipient Aid/GNI<sub>t-1</sub>.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table A26.** Relationship Between Appeals Mechanism and Project Success as Moderated by

 Bottom-Up Accountability and Governance Quality

Outcome: Project Success (t)	(1)	(2)	(3)	(4)
Appeals Mechanism <sub>t-1</sub>	0.284***	0.134	0.120	0.124
	(0.087)	(0.132)	(0.078)	(0.085)
CSO Participatory Environment <sub>t-1</sub>	0.035			
	(0.020)			
Appeals Mechanism <sub>t-1</sub> × CSO Participatory Environment <sub>t-1</sub>	0.020			
	(0.041)			
Total Press Score <sub>t-1</sub>		-0.001		
		(0.002)		
Appeals Mechanism <sub>t-1</sub> × Total Press Score <sub>t-1</sub>		0.003		

		(0.002)		
Control of Corruption <sub>t-1</sub>			-0.118	
			(0.113)	
Appeals Mechanism <sub>t-1</sub> × Control of Corruption <sub>t-1</sub>			-0.092	
			(0.066)	
Rule of Law <sub>t-1</sub>				-0.079
				(0.096)
Appeals Mechanism <sub>t-1</sub> × Rule of Law <sub>t-1</sub>				-0.097
				(0.067)
Observations	17,688	10,732	8,022	8,034
R-squared	0.132	0.159	0.176	0.176
Controls	Yes	Yes	Yes	Yes
Recipient Country Fixed Effects	Yes	Yes	Yes	Yes
Donor Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses. Controls are Recipient GDP per Capita Growth<sub>t-1</sub>, Recipient Log GDP per Capita<sub>t-1</sub> and Recipient Aid/GNI<sub>t-1</sub>.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1