When Does Transparency Improve Performance? Evidence from 23,000 Aid Projects in 148 Countries

Daniel 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

28 January 2019

Abstract: Access to information (ATI) policies—policies that formally guarantee a right to request information from public institutions—are often praised for enhancing transparency, accountability, and trust, yet empirical evidence that they lead to better governance outcomes is strikingly mixed. We argue that ATI policies only improve institutional performance when accompanied by enforcement mechanisms that prevent public officials from avoiding compliance with information requests that could reveal misconduct or ineffectiveness. In addition, we expect such mechanisms to deliver stronger performance dividends in the presence of robust "top-down" oversight by political principals and "bottom-up" oversight by citizens, civil society organizations, and other third-party stakeholders. We test our argument using a novel dataset on the performance of more than 23,000 foreign aid projects financed by 12 major bilateral and multilateral donor agencies in 148 countries between 1980 and 2016—the largest such dataset of its kind. We employ a staggered difference-in-differences design that exploits temporal variation in the adoption of ATI policies with and without enforcement mechanisms, confirming the results via an instrumental variables approach. The findings highlight the importance of the interaction between bottom-up and top-down mechanisms of accountability in conditioning how transparency interventions influence governance outcomes.

Keywords: transparency, accountability, freedom of information, institutional performance, public expenditure efficiency, aid effectiveness

Acknowledgments: We owe a debt of gratitude to Muhammad Saad Imtiaz for his outstanding research assistance. We also thank Alice Evans, Osvaldo Feinstein, and Chris Wilson for helpful comments on earlier drafts. The College of William and Mary, Johns Hopkins SAIS, and the William and Flora Hewlett Foundation provided the generous funding that made this study possible.

Introduction

In recent decades, access to information (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 citizens, civil society organizations, and other stakeholders to monitor institutional activities and decisions but also fundamentally alter their political relationships with such institutions. While ATI policies are frequently praised for enhancing transparency, accountability, and trust (Banisar 2006; Birkinshaw 2006; Florini 2007), however, they have a decidedly mixed empirical record. One set of studies demonstrates that they can increase bureaucratic efficiency (Vadlamannati and Cooray 2016a), curb corruption (Cordis and Warren 2014; Djankov et al. 2010; Peisakhin and Pinto 2010), and improve the selection of politicians (Fisman et al. 2017). Another set of studies, however, has found that they actually increase corruption (Escaleras et al. 2009; Costa 2013; Vadlamannati and Cooray 2016b) and erode trust in public institutions (Worthy 2010) while having no discernible positive impact on public understanding or political participation (Worthy 2013). We seek to contribute to this debate by theorizing and empirically evaluating the conditions under which ATI policies lead to better governance outcomes.

We argue that a major impediment to improved outcomes is the failure to properly *enforce* ATI policies: public officials have incentives to avoid complying with information requests that could reveal misconduct or poor performance, and are unlikely to be sanctioned by other bureaucrats—who themselves stand to benefit from nondisclosure—for acting on such incentives. The existence of reliable mechanisms for detecting, exposing, and remedying noncompliance is therefore essential for ensuring that ATI policies generate the new information necessary for bringing about improvements in institutional performance. We do not, however, expect the positive performance effects of enforcement mechanisms to be identical in all settings; rather, we posit that they increase with the robustness of two forms of institutional oversight: "police-patrol" oversight, or top-down, centralized, and active oversight by political principals; and "fire-alarm" oversight, or bottom-up, decentralized, and periodic oversight by citizens, civil society organizations, and other third-party stakeholders (McCubbins and Schwartz 1984).

We test these claims in the context of international development assistance, assessing 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—independently and conditionally influences the performance of aid projects they finance in low- and middle-income countries. Our analysis draws on a novel dataset on the outcome of more than 23,000 projects funded by 12 donors in 148 countries between 1960 and 2016—the largest such dataset of its kind. We employ a difference-in-differences identification strategy 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 if their disclosure requests are rejected.

This research design has a number of attractive features. First, and perhaps most importantly, it yields empirical conclusions with high levels of external validity. A significant reason for the mixed state of the existing literature on the effects of ATI policies is that it largely consists of single-setting studies covering short time periods, whose its findings—while based on compelling micro-level evidence—may not "travel" to different geographical and temporal contexts (Fox 2015: 348). Our wide-ranging collection of project performance data presents a rare opportunity to assess the impact of ATI policies on a large sample of governance outcomes across a diverse range of empirical settings. Second, the adoption of ATI policies by donors can reasonably be viewed as exogenous to the country contexts in which their projects are carried out. This is unusual in purely domestic policy domains: ATI policies are typically adopted as a result of reform efforts that may stem from or be related to factors that themselves influence such outcomes. This does not, of course, mean that ATI policy adoption is unrelated to donor or project characteristics; for example, donors with more successful projects may be more willing to institutionalize transparency. We show, however, that our findings are robust to the use of an instrumental variables (IV) strategy that leverages sources of variation

in the adoption of ATI policies with independent appeals mechanisms are plausibly exogenous to donor and project features as well as recipient country contexts. Third, the staggered adoption of ATI policies and appeals mechanisms, combined with the broad temporal coverage of projects in our dataset, allows us to control for a broad range of potentially confounding country-, donor-, and year-specific variables by comparing average pre- and post-adoption trends in project performance within a difference-in-differences framework. It allows us, in other words, to secure the benefits of strong external validity without sacrificing the benefits of rigorous causal identification.

The study contributes and add nuance to a broader research agenda—in political science and several other disciplines—that seeks to identify the specific circumstances in which transparency influences political outcomes. Our findings support an emerging consensus in this literature that transparency interventions are more effective in environments characterized by high levels of civic engagement, press freedom, and other forms of bottom-up accountability (Buntaine et al. 2018; Calland and Bentley 2013; Ferraz and Finan 2008; Fox 2015; Fung 2013; Grossman and Mitchelitch 2015; Kosack and Fung 2014; Larreguy et al. 2015; Lindstedt and Naurin 2010; Worthy 2015). Even in these favorable settings, however, we find that ATI policies only improve governance outcomes in the presence of robust enforcement mechanisms. In other words, our findings underscore the importance of the *interaction* between bottom-up monitoring and top-down enforcement in conditioning the governance effects of institutionalized transparency.

By highlighting independent appeals processes as an important instrument of enforcement, our study also contributes to a related line of research on the political effects of institutions that receive, monitor, and respond to complaints from stakeholders. A growing body of empirical research has shown that complaints institutions (and other nonelectoral methods of political participation) can enhance government responsiveness to citizen preferences (Deininger and Mpuga 2005; Cleary 2007; Bratton 2012; Trucco 2017; Ba 2017).[1] Our findings suggest an additional channel through which such institutions 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 drawing attention to the role of institutionalized transparency in influencing project outcomes (Denizer et al. 2013; Dreher et al. 2013; Winters 2014; Geli 2014; Buntaine 2016; Honig 2018; Lall 2017).

The Politics of Institutionalizing Transparency: Accountability, Enforcement, and Performance

There are several reasons why the adoption of ATI policies might generally be expected to improve governance outcomes. Increased disclosure by public sector institutions enhances the capacity of citizens, civil society organizations, the media, and other (public and private) stakeholders to monitor their performance (Peisakhin and Pinto 2010; Peisakhin 2012; Distelhorst 2017; Berliner et al. 2018; Buntaine et al. forthcoming). If requested information reveals suboptimal performance, it can be used by political principals—whether on their own initiative or in response to public pressure—to impose penalties on these institutions and demand change (Berliner and Erlich 2015; Nielson et al. 2003; Grigorescu 2010; Lorentzen et al. 2014). A growing body of experimental research suggests that when public officials (or their agents) know that they are being monitored or that their actions may be publicly disclosed, they are less likely to shirk or engage in malfeasant behavior and more likely to discharge their responsibilities efficiently and responsibly (Dal Bó et al. 2018; Carlson and Seim 2018; Jablonski and Seim 2017). Thus, ATI policies should incentivize public institutions to improve their performance both in anticipation of and in response to sanctions.

At the same time, there are reasons to doubt that ATI policies will—on their own—generate sufficient information to bring about improvements in institutional performance. Rather than increasing their effort

[1] As Botero et al. (2013: 959) put it, "[b]etter-educated countries have better governments, an empirical regularity that holds in both dictatorships and democracies" and "[p]ossible reasons for this fact are that educated people are more likely to complain about misconduct by government officials and that more frequent complaints encourage better behavior from officials. … Citizens' complaints might thus be an operative mechanism that explains the link between education and the quality of government."

and efficiency in response to such policies, public officials may pursue the less costly option of avoiding compliance with (legitimate) information requests that could reveal poor performance and thus lead to sanctions. The simplest way to circumvent compliance is to reject—or instruct those responsible for managing the information request process to reject—such requests on narrow procedural or technical grounds or due to alleged resource constraints (Prat 2005; Hood 2007; Holsen and Pasquier 2012; Trapnell 2014; Berliner 2017). These actions are often tolerated—or even encouraged—by senior officials in public sector organizations who themselves stand to directly or indirectly benefit from nondisclosure.

The observable implication is that ATI policies will only yield information that can be used to improve public-sector outcomes when they are accompanied by robust *enforcement* mechanisms—mechanisms that counterbalance incentives for bureaucratic noncompliance (Cordis and Warren 2014; Vadlamannati et al. 2016a).[2] 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, [the] 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."

The key mechanism for enforcing ATI policies at both the domestic and the international level is the existence of a recourse process that enables information seekers to appeal to an independent body—a body composed of individuals who are not part of the institution's staff—if their disclosure requests are rejected. For government agencies, appeals bodies take the form of judicial or quasi-judicial institutions charged with overseeing agency compliance with FOI legislation and sometimes imposing or recommending penalties for noncompliance (such as an information commission or ombudsman) (Holsen and Pasquier 2012). For international organizations, they are typically ad-hoc panels of external ATI experts from civil society, business, or government that are given the authority to uphold or reverse disclosure decisions. Penalties for noncompliance are imposed by the governing or executive bodies of these organizations and not by the appeals bodies themselves.

In addition to ensuring compliance with information requests, independent appeals mechanisms have the benefit of increasing confidence in and hence usage of the information request process and of clarifying 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."

The performance benefits of a properly enforced ATI policy apply no less to donor agencies—bilateral and multilateral public institutions that finance and oversee development projects in low-income and middle-income countries. These agencies are part of the long chain of delegation that characterizes the foreign aid project design and delivery.[3] Agency staff can be held to account for performance of the projects that they

[2] Cordis and Warren (2014) evaluate the impacts of weak and strong FOI laws at the state-level in the United States. They provide evidence that switching from a weak FOIA law to a strong FOI law increases the probability of detecting corrupt acts by public officials and reduces public sector corruption. These treatment effects accrue disproportionately to states where media coverage of public sector corruption is high. Vadlamannati et al. (2016a) report similar results but with cross-country evidence.

[3] Aid delivery consists of a long chain principal-agent (delegation) relationships: politicians in donor countries delegate implementation to an aid agency (bilateral or multilateral); aid agency administrators delegate to project managers; project managers in turn hire contractors to implement specific tasks; contractors deliver their products and services to recipient country organizations and agencies, who have their own internal delegation chains; and intended beneficiaries are at the very end of this long chain of intermediaries. All of these intermediaries have their own incentives, which often do not align with the preferences of intended beneficiaries (Gibson et al. 2012). Principals also face increasing levels of informational attrition at each additional link in the chain of delegation (Martens et al. 2012). Thus, critics of foreign aid charge that this long chain of delegation short-circuits the feedback loop between donors and their intended beneficiaries (Easterly 2006: 17; Easterly 2007: 330).

manage by supervisors and political principals.[4] However, they generally have weak incentives to ensure that the projects they manage are *effectively* designed and implemented; rather, donors tend to instead reward staff based on their ability to get projects *quickly* designed and implemented (Phillips 2009). Multilateral development banks depend on clients who are willing to borrow from them. Delaying loan approvals and disbursements on performance grounds threatens their basic business model (Dollar and Svensson 2000; Kilby 2009). Likewise, bilateral donor agencies receive "use-it-or-lose-it" appropriations from domestic legislatures, which creates incentives to "push money out the door" as quickly as possible (Lancaster 2002).[5]

The strength of these organizational incentives, however, are likely to vary under conditions of high and low information disclosure. When shareholders, legislators, auditors, evaluators, and third-party monitors have limited information about projects, it is unlikely donors will face countervailing pressures for effective project design and delivery. By contrast, in settings where information can be easily accessed, actors inside and outside the formally delegated authority (principal-agent) relationship can more easily demand accountability.

Pressures for better project performance can work through two types of observer effects: a *correction effect* whereby donors and aid-receiving governments more effectively design and implement existing projects because they know that they are being monitored and they fear possible sanctions (Reinikka and Svensson 2005; Francken et al. 2009; Legovini et al. 2015; Ensminger and Leder-Luis 2018), and a *selection effect* whereby greater transparency leads to better selection of future projects for fear of the consequences of proposing ill-conceived projects, allocating insufficient effort to project implementation, or misappropriating project funds (Faust 2011; Buntaine 2016).

Fire Alarms and Police Patrols at the Recipient Country and the Donor Level

While the adoption and proper enforcement of ATI policies by donors should, on average, increase the supply of information about projects, the willingness and ability of stakeholders to use such information to improve outcomes are likely to vary according to the socioeconomic, political, and institutional context in which projects are designed and delivered. McCubbins and Schwartz (1984) distinguish between two ways in which the activities of public institutions can be monitored: direct observation from the top ("police patrols") and third-party observation from below ("fire alarms"). Examples of the former include audits, investigations, and evaluations that are undertaken by public institutions themselves. Examples of the latter include anonymous hotlines, web platforms, and formal complaint mechanisms that allow members of the general public to report cases of misconduct and poor performance to public sector organizations (or their overseers). Police-patrol and fire-alarm mechanisms of oversight are potentially important moderating variables that could amplify or constrain the effectiveness of an institution's information disclosure regime.

While fire-alarm mechanisms of oversight are attractive due to their low cost, their efficacy requires the presence of civic monitors to sound the alarm when problems arise. Development projects take place in diverse settings: in some recipient countries, civil society organizations are numerous, well-established, and influential and there is political space for these groups to perform monitoring functions without impediment; in other countries, projects are undertaken without the benefit of many local groups that can credibly conduct

[4] In fact, in most donors, ex-post evaluation of project outcomes pre-dates the adoption of public disclosure policies, so the introduction of such policies adds another layer to the existing accountability environment, which could in turn affect the behaviors of those who work inside donors.

[5] Reflecting on his time as a project manager at the World Bank in the 1980s and 1990s, Whittle (2013) writes that "we faced clear incentives—namely, to get as many projects approved by the board as possible, and then to get them implemented in the allotted time frame; within the allotted budget; and without protests by the intended beneficiaries, public relations problems from local or international civil society organizations, or complaints from the government. As long as we responded to these clear incentives, we got promotions and steady pay raises. But if we spent too much money monitoring the projects and slowed implementation down to try to correct for problems, our projects could be downgraded, and this would generally be reflected on our performance evaluations."

fire alarm functions (Human Rights Watch 2015).[6] Thus, a key observable implication is that "[m]onitoring [donor activities] should be more likely when civil society groups have overcome barriers to collective action and established organizational means to collect and disseminate information" (Buntaine 2015: 101).[7]

There are also grounds for expecting well-enforced ATI policies to improve project outcomes in countries with strong police-patrol oversight. Once a donor has negotiated the basic terms and conditions of a given project agreement with a recipient government, management and oversight responsibilities are delegated to a particular government ministry or agency responsible for supervising those actually implementing project activities (e.g., private contractors, civil society organizations, public officials with frontline service delivery responsibilities). From a principal-agent perspective, therefore, recipient governments are one of the principals—in a long chain of delegation—that stand to benefit from more information about agents involved in designing and implementing projects. While a small number of recipient government officials, such as those in the Ministry of Finance's aid management unit, may be aware of the details of a project when it is initially negotiated, public disclosure of such information by donors can increase the amount of information available to a broader set of police patrol (accountability) institutions in the recipient county, including the legislature, auditor general offices, anti-corruption agencies, accountants and inspectors in line ministries, and law enforcement officials.

Incentives for effective project design and implementation *within* donors may also be strengthened when those responsible for monitoring the activities of donor staff and contractors gain access to new information generated by ATI policies with independent appeals mechanisms. Principals can better control wayward agents when they possess more information about the latter's behavior (Epstein and O'Halloran 1994; Hawkins et al. 2006b; Grigorescu 2010; Conceição-Heldt 2013). Thus, police-patrol mechanisms within donors, such as anti-corruption units and independent evaluation units, should perform more effectively when they can easily access information about the projects that they are responsible for monitoring.[8] Additionally, in settings where information can be easily accessed, it should be easier for those outside the principal-agent relationship—including citizens, civil society organizations, and independent media—to sound the alarm when donor projects are poorly designed or fail to meet their objectives during implementation (Buntaine 2015; Graham and Zvobgo 2018).

Donors can also create opportunities or impose constraints on potential monitors by offering or denying them access to accountability mechanisms. Theory and evidence suggest that transparency with accountability is a more potent treatment than transparency alone (Lindstedt and Naurin 2010; Lieberman et al. 2014; Lourenço 2015; Cuccinielo et al 2017; Seligsohn et al. 2018). Of particular theoretical relevance is whether and when donors adopt accountability mechanisms that allow the intended beneficiaries of their projects to lodge complaints, secure public acknowledgements of non-compliance and harm, and receive monetary and nonmonetary forms of compensation (Graham and Zvobgo 2018). These types of accountability institutions are consequential for two reasons. First, they create incentives for both the complainant and the target of the complaint to resolve the matter informally and expeditiously due to the high transaction costs of entering a formal grievance redressal process (Gauri 2013: 112).[9] Second, the quasi-judicial bodies that oversee accountability mechanisms take decisions that establish precedents, which can in turn trigger Bayesian updating among donor staff and produce positive selection effects in which future projects are brought

[6] In some countries, some governments have even gone so far as to criminalize communications between their citizens and international organizations (World Bank 2017: 8).
[7] Conversely, "[u]nder threat of political repression, civil society groups may not provide monitoring, choosing instead less confrontational approaches to dealing with [donors]" (Buntaine 2015: 101).
[8] Indeed, police-patrol staff within donors often complain that they are not able to effectively discharge their duties when project management staff (who have incentives to conceal instances of poor performance or non-compliance) slowly or selectively share information (e.g., Wessal et al. 2015: 7-8).
[9] The targets of the complaints can also suffer major reputational and career advancement losses. For example, World Bank staff have been fired after their projects received an unfavorable ruling by the Inspection Panel (Buntaine 2015: 103).

before their grievance redressal institutions.[10] The World Bank is a case in point. In 1994, it established an independent "Inspection Panel" that investigates project-related complaints lodged by individuals or groups within borrower countries, and there is empirical evidence that its track record of rulings on poorly performing projects has prompted World Bank staff to "panel-proof" future projects (Buntaine 2015).[11]

There are also reasons to believe that these types of accountability institutions can work synergistically with properly enforced ATI policies. A civil society group's ability to monitor donor projects and make effective use of grievance redressal mechanisms is fundamentally contingent upon its ability to access information about donor projects. In order to establish that a failure of compliance or performance has occurred, complainants need to meet a high evidentiary burden, and doing so is difficult in the absence of information about the decisions, actions, and motivations[12] of donor staff during project design and implementation.[13]

Hypotheses

The preceding discussion can be summarized in the following hypotheses:

H1: The adoption of ATI policies by donors will not, by itself, be associated with an improvement in project performance; however,

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

With respect to the factors that moderate these relationships, we expect that:

H3: The adoption of ATI policies with independent appeals mechanisms will have a stronger effect on project performance in the presence of greater oversight of recipient country governments by legislatures and other public bodies and of donors by independent evaluation units.

H4: The adoption of ATI policies with independent appeals mechanisms will have a stronger effect on project performance when recipient countries are characterized by greater bottom-up accountability (e.g., higher levels of civic engagement and media freedom) and when donors have accountability mechanisms; by contrast,

H5: The effect of the adoption of ATI policies on project performance does not vary with the robustness of top-down or bottom-up project oversight at either the recipient country or donor level.

Data

Outcome Variable

[10] For this reason, Gauri (2013: 111) distinguishes between "wholesale" and "retail" redress: "if an identifiable formal policy or informal practice leads to a larger numbers of service recipients being improperly denied a benefit, it will be less expensive for administrators to change the policy or practice than to deal with the complaints serially."

[11] This was the intended effect of the Inspection Panel. Clark (2003:2, emphasis added) writes that "[t]he creation of the Inspection Panel in 1993-1994 was essentially a publicly demanded response to a major credibility gap at the World Bank, whereby the words on paper and the fundamental objectives of the policies did not match the reality of implementation at the project level. It was hoped that this citizen-driven process would provide some means of holding the bank accountable to the people affected by its lending decisions, *and that having such a mechanism in place would lead to the avoidance of further disastrous projects*."

[12] For example, a complainant may seek to establish that a donor project manager "concealed difficulties so as to get a loan approved by the Board in order to meet his lending targets" (Wade 2009: 40).

[13] There are also some theoretical reasons to expect that civil society will be able to more effectively monitor specific types of donor activities. Buntaine (2015: 101) writes that "[c]ivil society groups have a superior ability to monitor policies and programs that directly impact identifiable and small groups of people. Under these conditions, information collected by civil society groups is likely to be credible, since outcomes of interest for oversight can be directly linked to operational actions by IOs." Conversely, Buntaine (2015: 101) argues that "[f]or accountability mechanisms like anti-fraud offices and staff misconduct hotlines, civil society groups may not be in a position to observe the majority of infractions."

In line with a growing literature on the determinants of aid effectiveness, we measure project performance using holistic, ex-post success ratings produced by donor staff, specialized evaluation departments, and external evaluators (such as professional consultancies) (Buntaine 2016; Denizer et al. 2013; Dreher et al. 2013; Geli 2014; Honig 2019). Specifically, the outcome variable in our analysis, *Project Success*$_{r,d,t}$, is the success rating for a project financed by donor *d* in recipient country *r* in year *t*, which is measured on a Likert-type scale that ranges from ranges from 1 for "highly unsatisfactory" to 6 for "highly satisfactory."[14] Through personal communications with donor staff and FOI requests (where possible), we were able to obtain ratings on more than 23,000 projects financed by 12 major donors between 1960 and 2016. Seven donors are multilateral organizations: African Development Bank (AfDB), the Asian Development Bank (AsDB), the Caribbean Development Bank (CDB), the Global Environment Facility (GEF), the Global Fund to Fight AIDS, Tuberculosis, and Malaria (GFATM), the International Fund for Agricultural Development (IFAD), and the World Bank. The remaining five are national government agencies: Australia's Department of Foreign Affairs and Trade (DFAT) (which incorporates Australian Aid), the UK's Department for International Development (DFID), Germany's Technical Cooperation Agency (GiZ) and German Development Bank (KfW), and the Japanese International Cooperation Agency (JICA). This collection of performance ratings represents the most comprehensive source of information on development project outcomes ever compiled.[15]

The underlying criteria used to measure project success is relatively consistent across donors, drawing on an Organisation for Economic Cooperation and Development (OECD) standard according to which ratings should capture a project's relevance, effectiveness, efficiency, sustainability, and impact (OECD 1991, 2000). While the OECD standard may, of course, be understood in different ways by donors, the inclusion of donor fixed effects in the below analyses controls for any constant donor-level variation in interpretive standards (as well as any constant rescaling of ratings and hence possible "grade inflation").

Treatment Variables

We merge the project success ratings with a second set of original data on the existence and characteristics of donor ATI policies, which covers the same agencies and time period. Our first treatment variable, *ATI Policy*$_{d,t}$, is a dummy for whether donor *d* possesses an ATI policy (as defined earlier) in year *t*. For bilateral donors, as noted earlier, ATI policies assume the form of national FOI legislation, which can be accessed on online legal repositories; for multilateral donors, they take the form of binding policy decisions taken by organizational governing or executive bodies, which are posted on their websites.[16] In some instances, these decisions were preceded by the adoption of informal guidelines or recommendations on disclosure, such as the World Bank's 1985 "Directive on Disclosure of Information" to staff (see World Bank 2009, Annex A). As these rules are typically nonbinding and not made public, we restrict our focus to formal ATI policies.

Our second treatment, *Appeals Mechanism*$_{d,t}$, is a dummy for whether donor *d*'s ATI policy provides for an independent appeals mechanism for rejected information request in year *t*. In line with existing comparative assessments of ATI policies—most notably Publish What You Fund's Aid Transparency Index and Access Info Europe and the Centre for Law and Democracy's Right to Information Rating—we exclude independent appeal mechanisms that are limited to a subset of information categories covered by a given

[14] These classifications are drawn from 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 consistent six-point scale to facilitate substantive interpretation.
[15] The database incorporates and substantially expands the Project Performance Database constructed by Honig (2018), adding four donors and approximately 9,000 projects.
[16] Superseded policies can be found on archived web pages. We use the Internet Archive's Wayback Machine to access such policies (https://archive.org/web/).

policy.[17] For the nine donors in our dataset that have been assessed in the Aid Transparency Index or the Right to Information Rating databases—namely, all those excluding the CDB, GEF, and IFAD—our coding decisions follow these assessments. A full list of ATI policies and accompanying independent appeals mechanisms for the 12 donors in our dataset are provided in the Appendix.

Moderating Variables

To test our hypotheses regarding the factors that moderate the relationship between ATI policies with appeals mechanisms and project performance, we supplement the outcome and treatment data with a variety of proxies for the strength of police patrol and fire alarm mechanisms of project oversight at both the recipient country and the donor level.

We employ three proxies for police-patrol oversight by recipient country governments, all of which are based on surveys of country experts conducted for Version 8 of the Varieties of Democracy (V-Dem) Dataset (Coppedge et al. 2018a).[18] $Executive\ Oversight_{r,t}$ is a standardized interval scale measuring the response provided for recipient country *r* to the survey question: "If executive branch officials were engaged in unconstitutional, illegal, or unethical activity, how likely is it that a body *other than the legislature*, such as a comptroller general, general prosecutor, or ombudsman, would question or investigate them and issue an unfavorable decision or report?" Responses were originally recorded on an ordinal scale ranging from 0 for "Extremely unlikely" to 4 for "Certain or near certain." $Legislature\ Investigates_{r,t}$ is an identical scale based on the question: "If the executive were engaged in unconstitutional, illegal, or unethical activity, how likely is it that a legislative body (perhaps a whole chamber, perhaps a committee, whether aligned with government or opposition) would conduct an investigation that would result in a decision or report that is unfavorable to the executive?" $Legislature\ Questions_{r,t}$ is a dummy variable transformed into a standardized interval scale that measures the response to the question: "In practice, does the legislature routinely question executive branch officials?" The original response options were 0 for "No—never or very rarely" or 1 for "Yes—routinely."

The main mechanism of police-patrol oversight at the donor level is the presence of an independent unit (e.g., office, department, or division)—a unit that does not report to any member of the donor's staff—whose exclusive or primary task is to evaluate donor policies and activities. Examples include the AsDB's Independent Evaluation Department, DFAT's Independent Evaluation Committee, the IMF's Independent Evaluation Office, and the KfW's Evaluation Department. $Evaluation\ Unit_{d,t}$ is a dummy for whether donor *d* possesses an independent evaluation unit in year *t*. Data on this variable were gathered from donor websites (current and archived), policy documents, academic studies, personal correspondence with donor staff, and OECD Development Assistance Committee (DAC) reporting.

We measure the robustness of fire-alarm oversight at the recipient country level using two survey-based indicators in the V-DEM Dataset. $CSO\ Participatory\ Environment_{r,t}$ is a standardized interval scale measuring the extent of popular involvement in civil society organizations (CSOs). The original scale ranges from 0 for "Most associations are state-sponsored, and although a large number of people may be active in them, their participation is not purely voluntary" to 3 for "There are many diverse CSOs and it is considered normal for people to be at least occasionally active in at least one of them." $Alternative\ Sources\ of\ Information_{r,t}$ is a 0-1 index constructed via Bayesian factor analysis of three standardized ordinal scales: (1) a measure of media bias against opposition parties or candidates, with an original range of 0 for "The print and broadcast media cover only the official party or candidates, or have no political coverage, or there are no opposition parties or

[17] An interesting example is the World Bank's appeal process, which, the 2016 ATI notes, is "limited and [contains] no right to appeal certain information items" (http://www.publishwhatyoufund.org/the-index/2016/donor/world-bank-ida/; accessed 13 September 2018).

[18] According to the codebook, "A Country Expert is typically a scholar or professional with deep knowledge of a country and of a particular political institution. Generally, that person is a citizen or resident of the country being coded. Multiple experts (usually five or more) code each variable" (Coppedge et al. 2018b, 29).

candidates to cover" to 4 for "The print and broadcast media cover all newsworthy parties and candidates more or less impartially and in proportion to their newsworthiness"; (2) a measure of how many major media outlets routinely criticize the government, with an original range of 0 for "None" to 4 for "All major media outlets criticize the government at least occasionally"; and (3) a measure of the diversity of political perspectives represented in the media, with an original range of 0 for "The major media represent only the government's perspective" to 3 for "All perspectives that are important in this society are represented in at least one of the major media."

At the donor level, we measure two related but distinct mechanisms of fire-alarm oversight. *Misconduct Mechanism*$_{d,t}$ is a dummy for whether donor *d* possesses an independent mechanism for receiving and addressing allegations of project-related misconduct from (internal and external) stakeholders in year *t*. Such mechanisms usually focus on fraud and corruption but may also encompass related forms of misconduct such as abuse of power, coercion, collusion, conflict of interest, and maladministration. *Complaints Mechanism*$_{d,t}$ is a dummy for the existence of an independent mechanism for receiving and addressing complaints from external stakeholders who have been (or are likely to be) adversely affected by projects. Most complaints mechanisms are restricted to adverse effects stemming from violations of donors' own social and environmental policies—the earliest and most famous example being the World Bank's Inspection Panel—though some consider complaints about issues such as human rights violations (e.g., GFATM's Human Rights Complaint procedure), a lack of appropriate consultation with stakeholders (e.g., GEF's Conflict Resolution Commissioner), and unfair or unreasonable treatment of stakeholders (e.g., Australia's Commonwealth Ombudsman). Data sources for these variables are similar to those for *Evaluation Mechanism*$_{d,t}$.

Control Variables

We control for three recipient country-level variables that commonly feature in empirical analyses of project performance: the annual growth rate of a recipient country's gross domestic product (GDP) per capita (*Recipient GDP per Capita Growth*$_{r,t}$); the log of a recipient country's GDP per capita (*Recipient Log GDP per Capita*$_{r,t}$); and the net value of official development assistance (ODA) provided to a recipient country as a percentage of its gross national income (GNI) (*Recipient Aid/GNI*$_{r,t}$). All variables are measured in current United States dollars using data from the World Bank's World Development Indicators database (World Bank 2018).

<< Table 1 around here >>

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

Baseline Analysis

Our baseline analysis employs a staggered difference-in-differences design that exploits temporal variation in the adoption of ATI policies and accompanying appeals mechanisms. To assess the unconditional effect of each treatment variable on project success, we begin by estimating the following equations with ordinary least squares (OLS):

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

$$Project\ Success_{r,d,t} = \alpha + \gamma_r + \phi_d + \psi_t + \beta_1 Appeals\ Mechanism_{d,t-3} + \beta_2 Controls_{r/d,t-3} + \varepsilon_{r,d,t} \quad (2)$$

where γ_r, ϕ_d, and ψ_t are fixed effects for recipient countries, donors, and years, respectively. We lag all treatment and control variables by three years (as indicated by the temporal subscripts) on the grounds that they are unlikely to influence the overall success of existing or future projects—which are generally conducted over several years—within a shorter space of time. To address the possibility of serial correlation in the outcome variable, we cluster robust standard errors at both the recipient country and the donor level.

The fixed effects control for unobserved heterogeneity that is both specific to individual donors and recipient countries but invariant over time and specific to individual years but invariant across donors and recipient countries. In doing so, they yield an estimate of the average treatment effect *within* donors and recipient countries over time. The causal parameter of interest, β_1, therefore represents the difference between the average change over time in project success for donors that "receive" the treatment (i.e., for which *ATI Policy*$_{d,t\text{-}3}$ = 1 in Equation 1 and *Appeals Mechanism*$_{d,t\text{-}3}$ = 1 in Equation 2) and the same difference for donors that do not "receive" the treatment (i.e., for which *ATI Policy*$_{d,t\text{-}3}$ = 0 in Equation 1 and *Appeals Mechanism*$_{d,t\text{-}3}$ = 0 in Equation 2). In addition to addressing several sources of potential confounding, this identification strategy has the advantage of avoiding direct inter-donor comparisons of project performance, which may be problematic due to the variable's partially subjective nature (Honig 2019).

To test our conditional hypotheses, we add interaction terms between each treatment variable *Appeals Mechanism*$_{d,t\text{-}3}$ *Appeals Mechanism*$_{d,t\text{-}3}$ and the various recipient country- and donor-level proxies for police-patrol and fire-alarm oversight:

$$Project\ Success_{r,d,t} = \alpha + \gamma_r + \phi_d + \psi_t + \beta_1 Treatment_{d,t\text{-}3} + \beta_2 Police\ Patrol_{r/d,t\text{-}3} + \beta_3 Treatment_{d,t\text{-}3} \times Police\ Patrol_{r/d,t\text{-}3} + \beta_4 Controls_{r,t\text{-}3} + \varepsilon_{r,d,t} \quad (3)$$

$$Project\ Success_{r,d,t} = \alpha + \gamma_r + \phi_d + \psi_t + \beta_1 Treatment_{d,t\text{-}3} + \beta_2 Police\ Patrol_{r/d,t\text{-}3} + \beta_3 Treatment_{d,t\text{-}3} \times Police\ Patrol_{r/d,t\text{-}3} + \beta_4 Controls_{r,t\text{-}3} + \varepsilon_{r,d,t} \quad (4)$$

Where *Treatment*$_{d,t\text{-}3}$ is *Appeals Mechanism*$_{d,t\text{-}3}$ or *ATI Policy*$_{d,t\text{-}3}$.

Unconditional Models

Table 2 reports the results for four variants of Equation 1. In Model 1, all control variables and fixed effects are omitted. The coefficient on *ATI Policy*$_{d,t\text{-}3}$ is positive but small and fails to reach statistical significance at any level. When the three controls are added in Model 3, the coefficient's sign, size, and significance level are essentially unchanged. Model 2 includes only the three sets of fixed effects. The coefficient turns negative and becomes large and significant at the five percent level. The addition of the controls in Model 4 (the full Equation 1) causes it to once again become small and non-significant, though does not alter its sign. The results are thus consistent with H1, with the slightly surprising caveat that the existence of an ATI policy is (weakly) associated with *worse* project performance.

<< Table 2 around here >>

Estimates for the same four variants of Equation 2, displayed in Table 3, provided similarly strong support for H2. In the pared-down Model 1, the coefficient on $Appeals\ Mechanism_{d,t\text{-}3}$ is large, positive, and significant at the one percent level, a result that is largely unaltered by the addition of the controls in Model 3. In a similar pattern to Equation 1, the coefficient remains positive but shrinks, acquires a larger standard error, and loses significance in Model 2. Incorporating the controls and fixed effects in Model 4 yields results consistent with Models 1 and 3 . This estimated treatment effect is sizable from a substantive as well as a statistical perspective: the existence of an ATI policy with an independent appeals mechanism is associated, on average, with a 0.36-point increase in $Project\ Success_{r,d,t}$ (on a 1-6 scale). This is a striking contrast to the 0.03-point *decline* associated with general adoption of an ATI policy (indicated by the results of Model 4, Table 2).

<< Table 3 around here >>

Conditional Models

Figures 1 and 2 plot the estimated marginal effect of $Appeals\ Mechanism_{d,t\text{-}3}$ on $Project\ Success_{r,d,t}$ at different levels of the four police-patrol and fire-alarm proxies, respectively; underlying regression results for Equations 3 and 4 are reported in the Appendix (all models include the full set of controls and fixed effects). In Figure 1, contrary to H3, the effect estimates slightly decline with the three recipient country-level proxies for police patrol oversight and (essentially) do not vary with the donor-level proxy. The near-horizontal slopes of the estimates ensure that they remain positive and significant at the 95 percent level at most or all values of the moderating variables. In other words, the size of the treatment effect is similar when police-patrol oversight is weak and strong. For example, $Appeals\ Mechanism_{d,t\text{-}3}$ is associated with an increase in $Project\ Success_{r,d,t}$ of 0.41 points when $Legislature\ Investigates_{r,t\text{-}3}$ is at its minimum value (i.e., it is "extremely unlikely" that a legislative body would conduct an investigation of an executive engaged in unconstitutional, illegal, or unethical activity that would result in an unfavorable outcome) and of 0.32 points—only 0.09 points less—when the variable is at its maximum value (i.e., it is "certain or nearly certain" that such an investigation would be carried out).

<< Figure 1 around here >>

The results for the fire-alarm proxies, by contrast, are broadly consistent with H4. The effect estimates are positive but small and statistically indistinguishable from zero at low levels of recipient country-level moderators. At high levels, however, they become sizable and significant at the 95 percent level. These differences are substantively large, with average effect estimates almost tripling in size as the moderators move from their lowest to their highest values. For instance, the existence of an independent appeals mechanism raises $Project\ Success_{r,d,t}$ by 0.13 points when $CSO\ Participatory\ Environment_{r,t\text{-}3}$ is at its minimum value (i.e., most associations are state-sponsored and participation is not purely voluntary) and 0.48 points—almost four times larger—when the variable is at its maximum value (i.e., there are many diverse CSOs and at least occasional participation is considered normal). At the donor level, effect estimates are positive and significant at all both values of each moderator. However, whereas they increase sharply with the presence of a misconduct mechanism, they are essentially unchanged by the presence of a complaints mechanism. Specifically, $Appeals\ Mechanism_{d,t\text{-}3}$ is associated with a rise in $Project\ Success_{r,d,t}$ of 0.18 points when $Misconduct\ Mechanism_{d,t\text{-}3} = 0$ and of 0.55 points when $Misconduct\ Mechanism_{d,t\text{-}3} = 1$; the equivalent figures for $Complaints\ Mechanism_{d,t\text{-}3}$ are 0.28 points and 0.3 points, respectively.[19]

[19] Interestingly, as shown in the Appendix, three of the four coefficients on the constituent (non-interacted) proxies, which represent the marginal effects of these variables when $Appeals\ Mechanism_{d,t\text{-}3} = 0$, are negative. In other words, most mechanisms of fire-alarm oversight are associated with better project performance when donors possess an ATI policy with an independent appeals mechanism but *worse* performance when they lack such a mechanism.

<< Figure 2 around here >>

Figures 3 and 4 display the equivalent of the previous set of marginal effects plots for $ATI\ Policy_{d,t-3}$. In line with H5, effect estimates are close to zero and statistically indistinguishable from zero at all levels of every police-patrol proxy and three of the four fire-alarm proxies (it is marginally significant when $Complaints\ Mechanism_{d,t-3}$ has a value of 1). Nor do they vary in a consistent direction, increasing with two of the police-patrol proxies and one of the fire-alarm proxies but declining with remaining five moderators.

<< Figure 3 around here >>

To summarize, the existence of an ATI policy with an independent appeals mechanism only has a strong positive impact on project performance in presence of robust decentralized, indirect oversight of recipient country governments by civil society and the media and of donors by internal and external stakeholders. Surprisingly, however, this effect slightly *declines* with the strength of centralized, direct oversight of recipient country governments by legislatures (and other public bodies) and of donors by independent evaluation units. As expected, the relationship between the existence of an ATI policy alone and project performance does not vary with the robustness of either police-patrol or fire-alarm mechanisms of oversight.

<< Figure 4 around here >>

Instrumental Variables Analysis

While adoption of ATI policies with independent appeals mechanisms can reasonably be viewed as exogenous to the country settings in which projects are implemented, it could nevertheless could be influenced by unobserved factors related to project success—or by project success *itself* (for instance, if donors with better-performing projects are more willing to disclose information to stakeholders). In other words, our empirical strategy could remain vulnerable to endogeneity in treatment assignment. In this section, we seek to address this possibility by using an IV strategy that leverage sources of plausibly exogenous variation in the adoption of ATI policies with independent appeals mechanisms.

As the determinants of appeals mechanism adoption differ for bilateral and multilateral donors, we employ separate instruments for each group. Building on a widely used spatial IV strategy in the political economy literature, we construct two instruments for bilateral donors: (1) the proportion of a donor's geographical neighbors that possess a domestic FOI law with an independent appeals mechanism;[20] and (2) the proportion of a donor's five largest trading partners that possess such a law. We create three variants of this variable that are measured as of year $t-4$, $t-5$, and $t-6$, that is, with lags of one, two, and three years relative to $Appeals\ Mechanism_{d,t-3}$, respectively. The logic behind these instruments is that the transparency policies of a bilateral donor's neighboring countries and trading partners are likely to influence its own transparency policies but do directly affect the performance of its development projects (rendering the exclusion restriction plausible).

[20] A geographical neighbor is defined as a sovereign state with which the donor country shares a land or maritime border.

For multilateral donors, we use the proportion of a donor's five largest shareholder countries that possess a domestic FOI law with an independent appeals mechanism as our instrument. Similarly to before, the motivation for this choice is that the transparency policies of a multilateral donor's major shareholders—which are typically its most powerful and influential members—should predict its own transparency policies but not exert a direct impact on the performance of its projects. As with the bilateral instruments, we construct three variants of this variable that are lagged one, two, and three years relative to *Appeals Mechanism*$_{d,t\text{-}3}$. We then merge each instrument with its equivalent variant of the two bilateral instruments, creating a total of six combined instruments (two for each of the three lag structures).

We implement the IV analysis using a two-stage least squares (2SLS) regression model. In the first stage, predicted values of *Appeals Mechanism*$_{d,t\text{-}3}$ are generated by regressing this variable on the combined instrument and all control variables and fixed effects in the full baseline models:

$$Appeals\ Mechanism_{d,t\text{-}3} = \alpha + \gamma_r + \phi_d + \psi_t + \beta_1 Combined\ Instrument_{d,t\text{-}3} + \beta_2 Controls_{r,t\text{-}3} + \varepsilon_{r,d,t} \quad (5)$$

In the second stage, the *Project Success*$_{r,d,t}$ is regressed on the predicted values of *Appeals Mechanism*$_{d,t\text{-}3}$ from the first stage as well as the control variables and fixed effects:

$$Project\ Success_{r,d,t} = \alpha + \gamma_r + \phi_d + \psi_t + \beta_1 \widehat{Appeals\ Mechanism}_{d,t\text{-}3} + \beta_2 Controls_{r,t\text{-}3} + + \varepsilon_{r,d,t} \quad (6)$$

Table 4 presents the second-stage results (Equation 6) for the six combined instruments. In the first-stage model (Equation 5), as reported in the bottom row, all instruments have a Cragg-Donald F-Statistic far exceeding 10, the standard threshold for distinguishing "weak" from "strong" instruments. Note, however, smaller lags relative to *Appeals Mechanism*$_{d,t\text{-}3}$ are associated with substantively higher F-Statistics, indicating that more recent values of the instruments better predict the treatment in year *t*.

<< Table 4 around here >>

All coefficients on the instrumented measures of *Appeals Mechanism*$_{d,t\text{-}3}$ are positive, large, and significant at the five percent or the one percent level. On average, the estimates are almost one one-fifth larger than in the full baseline model (Model 4, Table 2): the adoption of ATI policy with an independent appeals mechanism is attended by a rise in *Project Success*$_{r,d,t}$ of 0.43. This figure is marginally larger for the combined instruments based on the neighbor reference group for bilateral donors (0.46) than those based on the trading partner reference group (0.4), though even the latter substantially exceeds the baseline marginal effect estimate. The findings thus provide compelling evidence that the earlier results were not driven or inflated by endogeneity in treatment assignment; to the contrary, they suggest that these results may have slightly *underestimated* the true treatment effect.[21]

Additional Robustness Checks

We subject the results in Table 3 to a series of additional robustness checks; estimates are reported in the Appendix. First, to probe plausibility of the assumption that outcome trends in treated and control units would have been the same in the absence of the treatment—the key identifying assumption of the different-

[21] As shown in the Appendix, these results are robust to the employment of five- and six-year lags on the instruments (i.e., two- and three-year lags relative to the *Appeals Mechanism*$_{d,t\text{-}3}$).

in-differences estimator—we employ two standard strategies: (1) adding to the regression model 1-10 year leads and lags as well as a contemporaneous version of $Appeals\ Mechanism_{d,t-3}$, with the expectation that the coefficients on the leads will be zero); and (2) controlling for donor- and recipient country-specific linear time trends, which allows treated and control groups to follow different outcome trajectories.[22] Second, we investigate whether $Appeals\ Mechanism_{d,t-3}$ is merely serving as a proxy for differences in the overall stringency of an ATI policy by conducting a placebo test in which the variable replaced by a dummy for whether donor *d*'s ATI policy codifies the "presumption of disclosure principle"—a provision establishing disclosure as the general rule and hence requiring a compelling reason for nondisclosure—in year $t - 3$. Third, we experiment with four alternative lag structures for $Appeals\ Mechanism_{d,t-3}$, namely, measuring the variable as of year $t - 1$, $t - 2$, $t - 4$, and $t - 5$, respectively. Fourth, we reparametrize $Project\ Success_{r,d,t}$ as three separate binary variables, fitting a logistic version of Equation 2 in each case: (1) a dummy for whether a given project's rating exceeds the sample mean; (2) a dummy for whether a given project's rating exceeds the sample median; and (3) a dummy for whether a given project's rating is equal to the scale maximum (i.e., the highest rating assigned by the donor).

The results are robust to each of these tests as well several additional specifications, including limiting the sample to projects evaluated in years where donors possess an independent evaluation function (as defined earlier) and hence might be expected to produce more "objective" ratings; restricting the analysis to the period after (1) 1990 and (2) 1995 to address the possibility that evaluation methods and standards have substantially changed over time; and employing an alternative coding of $Appeals\ Mechanism_{d,t-3}$ for the World Bank—the donor with the largest number of projects in the dataset—as the original assignment of "0" might be seen as controversial by some aid practitioners. These tests are described in greater detail in the Appendix.

Discussion

Our empirical analysis offers a window into an important but understudied question in the study of transparency and governance: Under what conditions do ATI policies improve the performance of public institutions? Foreign aid projects are just one instantiation of socially, economically, and politically consequential outputs delivered by such institutions, yet they provide an ideal set of empirical cases for examining the effects of institutionalized transparency on governance. Unlike in a purely domestic setting, the adoption of ATI policies by donor agencies is plausibly exogenous to the country contexts in which projects are implemented. The staggered nature of ATI policy adoption across donors, combined with the wide temporal variation of projects in our dataset, enables identification of the causal effect of such policies by comparing average pre- and post-adoption trends in project success within a difference-in-differences framework.

Foreign aid projects are certainly atypical in some respects. Intended beneficiaries are not taxed for the goods and services that they receive. Nor do they typically have voice, vote, or jurisdictional exit options when projects are poorly implemented. As a result, few mechanisms exist for holding donors accountable if projects harm local communities and ecosystems, underperform vis-à-vis performance objectives, or run afoul of host government rules, regulations, and laws.[24] These unfavorable conditions cause many projects to fail or falter during implementation (Easterly 2006: 17; Winters 2014; Findley et al. 2017a; Ensminger and Leder-Luis 2018).

Yet our findings suggest that, even under these unfavorable conditions, ATI policies can help to repair the broken feedback loop between public institutions and their intended beneficiaries by reducing information asymmetries at multiple links in a lengthy chain of principal-agent relationships. Critically, however, this fix

[22] Visual inspection of pretreatment outcome trends in treated and control groups—another common method for assessing the counterfactual trends assumption—is challenging when there are multiple treated groups and periods and differences in treatment timing.

[24] This is a key difference between foreign aid programs and programs that governments implement within their own jurisdictional boundaries ((Martens et al. 2002; Gibson et al. 2002; Whittle 2013).

requires more than the mere formalization of a right to request information from such institutions: we find no evidence that the adoption of an ATI policy—on its own—results in 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 to an independent body when their disclosure requests are rejected—a body that enforces compliance with such policies and thus ensures that they generate information that can be used by stakeholders to improve project design and implementation.

Furthermore, the adoption of ATI policies with independent appeals mechanisms yields stronger project performance dividends when recipient countries are characterized by higher levels of civic engagement and press freedom and when donors provide institutionalized channels through which stakeholders can draw attention to project-related misconduct. The moderating effects of these fire-alarm mechanisms of project oversight are likely to stem from the increased public pressures they create on recipient country governments, donors, and their contractors to utilize information generated by ATI policies to improve the delivery of existing projects. They could also, of course, prompt these actors to strengthen the design of future projects to pre-empt potential sanctions for ineffective performance. What *is* clear is that citizens, civil society organizations, the media, and other stakeholders in recipient countries play a critical role in ensuring that expanded information disclosure translates into better project outcomes.

Contrary to our expectations, we do not find that police-patrol mechanisms of project oversight—such as scrutiny of recipient country governments by legislatures and other public bodies and the existence of an independent evaluation function at the donor level—moderate the positive relationship between properly enforced ATI policies and project success. What might explain this finding? One possibility is that donor staff and their recipient country counterparts have limited capacity to use information revealed by ATI policies to force project implementers make course corrections after scopes of work have been finalized and contracts have been signed. It is also possible that ATI policies help auditors and evaluators identify problems but only at later stages of project implementation or even after project closure, limiting opportunities for remedial action. Another potential explanation is that, since recipient country governments generally do not use independent appeals mechanisms themselves, donor agencies do not alter their behavior in anticipation of these mechanisms being used. Determining why mechanisms of police-patrol oversight do not amplify the impact of ATI policies on project outcomes is an important avenue for future research.

Neither fire-alarm nor police-patrol oversight moderates the relationship between the adoption ATI policies in general and project performance. In other words, ATI policies alone yield no benefits for project performance even in the presence of robust bottom-up and top-down accountability at the recipient country and the donor level. This result further underscores the importance of proper enforcement in ensuring that ATI policies produce novel information that can be used to improve project performance.

What policy lessons can be drawn from these findings? Perhaps the most important is that the *design* of transparency interventions matters; in the absence of reliable mechanisms for deterring or preventing bureaucratic noncompliance, efforts to expand disclosure by public institutions are unlikely to enhance their performance. In particular, the findings show that mechanisms for collecting, evaluating, and addressing complaints from stakeholders can serve as a potent instrument for enforcement—an instrument that harnesses the benefits of both bottom-up and top-down accountability. Our findings regarding the conditioning effects of fire-alarm oversight are harder to translate into specific policy actions, given the close relationship between levels of civic engagement and press freedom and characteristics such as democracy, respect for human rights, and the rule of law. They do, however, suggest that policymakers can increase the chances that well-enforced transparency policies will result in better governance outcomes by providing institutionalized channels through which stakeholders—even in settings with limited civic engagement and freedom of expression—can draw attention to institutional misconduct and other performance problems.

References

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. 2015. *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.

Bennear, Lori S., and Sheila M. Olmstead 2008. The impacts of the right to know: information disclosure and the violation of drinking water standards. *Journal of Environmental Economics and Management* 56(2): 117–130.

Birkinshaw, Patrick. 2006. "Freedom of Information and Openness: Fundamental Human Rights?" Administrative Law Review 58 (1): 177-218.

Björkman, Martina, and Jakob Svensson. 2009. Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda. *Quarterly Journal of Economics* 124 (2): 735–69.

Bjorkman Nyqvist, Martina, Damien de Walque and Jakob Svensson. 2017. Experimental Evidence on the Long-Run Impact of Community-Based Monitoring. *American Economic Journal: Applied Economics* 9(1): 33–69.

Bobonis, Gustavo J., Luis R. Cámara Fuertes, and Rainer Schwabe. Forthcoming. Monitoring Corruptible Politicians. American Economic Review 106 (8): 2371-2405.

Botero, J, A Ponce and A Shleifer. 2013. Education, Complaints, and Accountability. *Journal of Law and Economics* 56(4): 959–996.

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

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, UK: Oxford University Press.

Buntaine, Mark T., Daniel L. Nielson, and Jacob T. Skaggs. 2017. Escaping the Valley of Disengagement: Two Field Experiments on Motivating Citizens to Monitor Public Goods. AidData Working Paper #41. Williamsburg, VA: AidData. http://docs.aiddata.org/ad4/pdfs/wps41_escapingthevalley_combined.pdf.

Buntaine, Mark, Sarah Bush, Ryan Jablonski, Daniel Nielson, and Paula Pickering. 2018. SMS texts on corruption help Ugandan voters hold elected councillors accountable at the polls. *Proceedings of the National Academy of Sciences* 115(26): 6668-6673.

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

Calland, Richard, and Kristina Bentley. 2013. The Impact and Effectiveness of Transparency and Accountability Initiatives: Freedom of Information. *Development Policy Review* 31 (S1): S69-S87.

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.

Chong, Alberto., Ana L. De La O, Dean Karlan, and Leonard Wantchekon. 2015. Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice, and Party Identification. *The Journal of Politics* 77 (1): 55–71.

Clark, Dana. 2003. "Understanding the World Bank Inspection Panel." In *Demanding Accountability: Civil Society Claims and the World Bank Inspection Panel*, eds. Dana Clark, Jonathan A. Fox and Kay Treakle, pp. 1-24. Lanham: Rowman & Littlefield Publishers.

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

Conceição-Heldt, Eugénia da. 2013. Do agents 'run amok'? Agency slack in the EU and US Trade Policy in the Doha round. *Journal of Comparative Policy Analysis* 15(1): 21–36.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, …, and Daniel Ziblatt. 2018a. "*V-Dem [Country-Year/Country-Date] Dataset v8*." Varieties of Democracy (V-Dem) Project. https://doi.org/10.23696/vdemcy18.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, …, and Daniel Ziblatt. 2018a. 2018b. "V-Dem Codebook 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.

Cucciniello, M., Porumbescu, G. A. and Grimmelikhuijsen, S. 2017. 25 Years of Transparency Research: Evidence and Future Directions. *Public Administration Review* 77: 32–44.

Dal Bó, Ernesto, Frederico Finan, Nicholas Y. Li, Laura Schechter. 2018. Government Decentralization Under Changing State Capacity: Experimental Evidence from Paraguay. Working Paper.

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.

Deininger, Klaus, and Paul Mpuga. 2005. Does Greater Accountability Improve the Quality of Public Service Delivery? Evidence from Uganda. *World Development* 33(1): 171–191.

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

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2010. *Disclosure by Politicians. American Economic Journal: Applied Economics* 2 (2): 179-209.

Dollar, David and Jakob Svensson. 2000. What Explains the Success or Failure of Structural Adjustment Programmes? *Economic Journal* 110 (466): 894-917.

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.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2013. Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India. *The Quarterly Journal of Economics* 128 (4): 1499–1545.

Easterly, William. 2006. *The White Man's Burden*. NY: Penguin Press.

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

Easterly, William, and Tobias Pfutze. 2008. Where does all the money go? Best and worst practices in foreign aid. *Journal of Economic Perspectives* 22 (2): 29-52.

Ensminger, Jean, and Jetson Leder-Luis. 2018. Measuring Strategic Data Manipulation: Evidence from a World Bank Project. Working Paper.

Epstein, David and Sharyn O'Halloran. 1994. Administrative Procedures, Information, and Agency Discretion. *American Journal of Political Science* 38 (3): 697-722.

Escaleras, Monica, Shu Lin, and Charles Register. 2009. Freedom of Information Acts and Public Sector Corruption. *Public Choice* 145 (3–4): 435–60.

Etzioni, Amitai. 2010. Is Transparency the Best Disinfectant? *Journal of Political Philosophy* 18(4): 389–404.

Etzioni, Amitai. 2014. The Limits of Transparency. *Public Administration Review* 74(6): 687–88.

Faust, Jörg. 2011. Donor transparency and aid allocation. DIE Discussion Paper 12/2011. Bonn: German Development Institute (DIE).

Ferraz, Claudio, and Frederico Finan. 2008. Exposing Corrupt Politicians: The Effect of Brazil's Publicly Released Audits on Electoral Outcomes. *Quarterly Journal of Economics* 123 (2): 703–45.

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

Findley, Michael G., Helen V. Milner and Daniel L. Nielson. 2017b. The choice among aid donors: The effects of multilateral vs. bilateral aid on recipient behavioral support. *Review of International Organizations* 12(2): 307–334.

Fisman, Raymond, Florian Schulz, and Vikrant Vig. 2017. Financial disclosure and political selection: Evidence from India. Working Paper. https://pdfs.semanticscholar.org/5ae3/cc20d675a46aeaec69b59f9f0be2e2eeed81.pdf.

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

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

Francken, Nathalie, Bart Minten, and Johan F. M. Swinnen. 2009. Media, monitoring, and capture of public funds: Evidence from Madagascar. *World Development* 37 (1): 242-255.

Fukuyama, Francis. 2015. Why Is Democracy Performing So Poorly? *Journal of Democracy* 26 (1): 11–20.

Fung, Archon. 2013. Infotopia: Unleashing the Democratic Power of Transparency. *Politics & Society* 41 (2): 183-212.

Gauri, Varun. 2013. Redressing Grievances and Complaints Regarding Basic Service Delivery. *World Development* 41 (C): 109-119.

Geli, Patricia; Kraay, Aart; Nobakht, Hoveida. 2014. Predicting World Bank Project Outcome Ratings. Policy Research Working Paper No. 7001. World Bank Group, Washington, DC. http://documents.worldbank.org/curated/en/878191468163759728/pdf/WPS7001.pdf.

Gibson, Clark C., Krister Andersson, Elinor Ostrom, and Sujai Shivakumar. 2002. *The Samaritan's Dilemma: The Political Economy of Development Aid*. Oxford, UK: Oxford University Press.

Golden, Miriam, and Brian Min. 2013. Distributive Politics around the World. *Annual Review of Political Science* 16: 73–99.

Gottlieb, Jessica. 2016. Greater Expectations: A Field Experiment to Improve Accountability in Mali. *American Journal of Political Science* 60 (1): 143–57.

Gould, Erica R. 2017. What Impedes IO Accountability Mechanisms? The Case of Multilateral Development Bank Internal Accountability Offices. Working Paper.

Graham, Benjamin A.T. and Kelebogile Zvobgo. 2018. The World Bank as an Enforcer of Human Rights. Working Paper. http://wp.peio.me/wp-content/uploads/2018/01/PEIO11_paper_52-1.pdf.

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

Grigorescu, Alexandru. 2010. The Spread of Bureaucratic Oversight Mechanisms across Intergovernmental Organizations. *International Studies Quarterly* 54: 871-886.

Grossman, Guy, Macartan Humphreys, and Gabriella Sacramone- Lutz. 2014. "I wld like u WMP to extent electricity 2 our village": On Information Technology and Interest Articulation. *American Political Science Review* 108 (3): 688-705.

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.

Hawkins, Darren, David Lake, Daniel Nielson, and Michael Tierney (eds.). 2006a. *Delegation and Agency in International Organizations*. New York: Cambridge University Press.

Hawkins, Darren, David Lake, Daniel Nielson, and Michael Tierney. 2006b. "Delegation under anarchy: States, international organizations, and principal-agent theory." In *Delegation and Agency in International Organizations*, eds. Darren Hawkins, David Lake, Daniel Nielson, and Michael Tierney. Cambridge: Cambridge University Press, pp. 3–38.

Hawkins, Darren and Wade Jacoby. 2006. "How agents matters." In *Delegation and Agency in International Organizations*, eds. Darren Hawkins, David Lake, Daniel Nielson, and Michael Tierney. Cambridge: Cambridge University Press, pp. 199-228.

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.

Hood, Christopher. 2007. What Happens When Transparency Meets Blame Avoidance? *Public Management Review* 9 (2): 191–210.

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*, forthcoming.

Human Rights Watch . 2015. At Your Own Risk: Reprisals against Critics of World Bank Group Projects. New York: Human Right Watch.

International Fund for Agricultural Development (IFAD). 2004. Supervision Modalities in IFAD Supported Projects. Corporate-level Evaluation Report No. 1540. Rome: IFAD. Accessed at https://www.ifad.org/documents/10180/6df90a53-514f-49ac-925a-9f7a02d880b9.

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. http://docs.aiddata.org/ad4/pdfs/WPS45-How_Transparency_Affects_Distributional_Politics.pdf

Keefer, Philip, and Stuti Khemani. 2016. The Government Response to Informed Citizens: New Evidence on Media Access and the Distribution of Public Health Benefits in Africa. *World Bank Economic Review* 30 (2): 233–267.

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

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.

Lancaster, Carol. 2002. The Devil Is in the Details: From the Millennium Challenge Account to the Millennium Challenge Corporation. Washington DC: Center for Global Development.

Larreguy, Horacio A., John Marshall, and James M. Snyder, Jr. 2014. Revealing Malfeasance: How Local Media Facilitates Electoral Sanctioning of Mayors in Mexico. NBER Working Paper No. 20697. https://www.nber.org/papers/w20697.pdf.

Legovini, Arianna, Vincenzo Di Maro, and Caio Piza. 2015. Impact Evaluation Helps Deliver Development Projects. World Bank Policy Research Working Paper 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.

Lindberg, Staffan I. 2013. Mapping accountability: core concept and subtypes. *International Review of Administrative Sciences* 79(2): 202–226.

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.

Lourenço, Rui Pedro. 2015. An analysis of open government portals: A perspective of transparency for accountability. *Government Information Quarterly* 32 (3): 323-332.

Martens, B., Mummert, U., Murrell, P., and Seabright, P. 2002. *The institutional economics of foreign aid.* New York: Cambridge University Press.

McCubbins, Mathew, and Thomas Schwartz. 1984. Congressional Oversight Overlooked: Police Patrols versus Fire Alarms. *American Journal of Political Science* 28 (1): 165–179.

McCubbins, Mathew D., Roger G. Noll, and Barry R. Weingast. 1987. Administrative Procedures as Instruments of Political Control. *Journal of Law, Economics, and Organization* 3 (2): 243-279.

Michener, Gregory. 2011. FOI Laws around the World. Journal of Democracy 22 (2): 145–159.

Michener, Greg. 2015. How Cabinet Size and Legislative Control Shape the Strength of Transparency Laws. *Governance* 28(1): 77–94.

Mitchell, Ronald B. 1998. Sources of Transparency: Information Systems in International Regimes. *International Studies Quarterly* 42(1): 109–30.

Moon, Samuel. And Tim Williamson. 2010. Greater aid transparency: crucial for aid effectiveness. Overseas Development Institute (ODI) Briefing Paper No. 35. London, UK: ODI

Nelson, Paul. 2001. Transparency Mechanisms in the Multilateral Development Banks. *World Development* 19: 1835-1847.

Neuman, Laura. 2006. Access to Information Laws: Pieces of the Puzzle – An Analysis of the International Norms. White Paper. Atlanta: GA: The Carter Center.

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*.

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

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

Pandey, Priyanka, Sangeeta Goyal, and Venkatesh Sundararaman. 2009. Community Participation in Public Schools: Impact of Information Campaigns in Three Indian States. *Education Economics* 17 (3): 355–375.

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

Peisakhin, Leonid, and Paul Pinto. 2010. Is Transparency an Effective Anti-Corruption Strategy? Evidence from a Field Experiment in India. *Regulation and Governance* 4 (3): 261–280.

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

Porumbescu, Gregory. 2017. Linking Transparency to Trust in Government and Voice. *American Review of Public Administration* 47 (5).

Prat, Andrea. 2005. The Wrong Kind of Transparency. *American Economic Review* 95(3): 862–77.

Pritchett, Lant. 2002. It Pays to Be Ignorant: A Simple Political Economy of Rigorous Program Evaluation. *Journal of Economic Policy Reform* 5(2): 251–88.

Rasul, Imran, and Daniel Rogger. 2018. Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service. *Economic Journal* 128 (608): 413-446.

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.

Reinikka, Ritva, and Jakob Svensson. 2011. The Power of Information in Public Services: Evidence from Education in Uganda. *Journal of Public Economics* 95 (7): 956–66.

Roberts, Alasdair. 2010. A Great and Revolutionary Law? The First Four Years of India's Right to Information Act. *Public Administration Review* 70 (6): 925-933.

Rose-Ackerman, Susan. 2004. The Challenge of Poor Governance and Corruption. Copenhagen Consensus Challenge Paper.

Roseth, Benjamin and Vivek Srivastava. 2015. Engaging for Results in Civil Service Reforms: Early Lessons from a Problem-Driven Engagement in Sierra Leone. *Journal of Management and Strategy* 6 (2): 53-69.

Schedler, Andreas. 1999. "Conceptualizing Accountability." In *The Self-Restraining State: Power and Accountability in New Democracies*, eds. Andreas Schedler, Larry Diamond, and Marc F. Plattner. Boulder, CO: Lynne Rienner Publishers. Pp 13-29.

Seligsohn, Deborah, Mengdi Liu, and Bing Zhang. 2018. The sound of one hand clapping: transparency without accountability. *Environmental Politics* 27(5): 804-829.

Sjoberg, Fredrik M., Jonathan Mellon, and Tiago Peixoto. 2017. The Effect of Bureaucratic Responsiveness on Citizen Participation. *Public Administration Review* 77 (3): 340–351.

Snell, Rick. 2001. Freedom of Information: The Experience of the Australian States - an Epiphany? *Federal Law Review* 29(3): 343–358.

Trapnell, Stephanie. (ed.) 2014. *Right to Information: Case Studies on Implementation*. Washington, DC: World Bank.

Trucco, Laura. 2017. Broken Cities: The Effect of Government Responsiveness on Citizens' Participation. Working Paper. https://scholar.harvard.edu/files/lauratrucco/files/trucco_jmp_20170111.pdf.

Tsai, Lily L., Nina McMurry and Swetha Rajeswaran. 2018. The effect of civic leadership training on citizen engagement and government responsiveness: experimental evidence from the Philippines. Making All Voices Count Research Report. Brighton, UK: IDS.

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.

Wade, Robert H. 2009. Accountability Gone Wrong: The World Bank, Nongovernmental Organisations and the US Government in a Fight over China. *New Political Economy* 14 (1): 25-48.

Wantchekon, Leonard. 2003. Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin. *World Politics* 55: 399–422.

Wessal, Arianne, Nidhi Khattri, and Dawn Roberts. 2015. Managing Evaluations: A How-To Guide For Managers and Commissioners of Evaluations. Washington DC: Independent Evaluation Group of the World Bank.

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

Williams, Martin J. 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.

World Bank. 2018. World Development Indicators. Washington, D.C.: The World Bank (producer and distributor). http://data.worldbank.org/data-catalog/world-development-indicators.

World Bank. 2009. Toward Greater Transparency through Access to Information. Operations Policy and Country Services. Washington, D.C.: The World Bank.

World Bank. 2017. The World Bank Group Fiscal Year 2017 Annual Update Integrity Vice Presidency [INT]. Washington DC: World Bank.

Worthy, Ben (2010) 'More Open But Not More Trusted? The Effect of FOI on the UK 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.

Worthy, Ben. 2015. The Impact of Open Data in the UK: Complex, Unpredictable, and Political. *Public Administration* 93 (3): 788-805.

Tables and Figures

Table 1. Summary Statistics for Variables in Dataset

Variable	Observations	Mean	Std. Dev	Min	Max
Project Success	23,204	4.21	1.16	1.00	6.00
ATI Policy	23,866	0.48	0.50	0.00	1.00
Appeals Mechanism	23,866	0.15	0.35	0.00	1.00
Executive Oversight	24,347	0.16	1.26	-2.83	3.23
Legislature Investigates	24,377	0.03	1.36	-3.00	3.68
Evaluation Unit	23,866	0.70	0.46	0.00	1.00
Legislature Questions	24,379	0.33	1.27	-2.11	2.46
CSO Participatory Environment	26,788	0.43	1.26	-3.36	3.16
Alternative Sources of Information	26,795	0.56	0.31	0.01	0.98
Misconduct Mechanism	23,866	0.36	0.48	0.00	1.00
Complaints Mechanism	23,866	0.45	0.50	0.00	1.00
Recipient GDP per Capita Growth Rate	26,298	2.79	6.08	-65.00	140.50
Recipient Log GDP per Capita	26,224	7.54	1.28	4.75	11.88
Recipient Aid/GNI	23,164	7.20	10.74	-2.63	242.29

Table 2. Relationship between Existence of ATI Policy and Project Success

Outcome: Project Success$_t$, 1960-2016	(1)	(2)	(3)	(4)
ATI Policy$_{t-3}$	0.034	-0.338**	0.028	-0.028
	(0.086)	(0.118)	(0.091)	(0.098)
Recipient GDP per Capita Growth$_{t-3}$			0.016***	0.002
			(0.004)	(0.002)
Recipient Log GDP per Capita$_{t-3}$			0.018	-0.251***
			(0.019)	(0.050)
Recipient Aid/GNI$_{t-3}$			-0.004*	-0.003
			(0.002)	(0.002)
Constant	4.188***		4.047***	
	(0.073)		(0.183)	
Observations	19,554	19,554	19,554	19,554
R-squared	0.000	0.027	0.009	0.120
Recipient Country Fixed Effects	No	Yes	No	Yes
Donor Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses. Constant omitted in Models 2 and 4 due to collinearity with fixed effects.

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

Table 3. Relationship between Presence of ATI Policy Appeals Mechanism and Project Success

Outcome: Project Success$_t$, 1960-2016	(1)	(2)	(3)	(4)
Appeals Mechanism$_{t-3}$	0.392***	0.356	0.407***	0.358***
	(0.139)	(0.227)	(0.146)	(0.097)
Recipient GDP per Capita Growth$_{t-3}$			0.015***	0.003
			(0.004)	(0.001)
Recipient Log GDP per Capita$_{t-3}$			0.017	-0.242***
			(0.017)	(0.050)
Recipient Aid/GNI$_{t-3}$			-0.006**	-0.003
			(0.003)	(0.002)
Constant	4.158***		4.037***	
	(0.077)		(0.162)	
Observations	19,554	19,554	19,554	19,554
R-squared	0.011	0.026	0.020	0.122
Recipient Country Fixed Effects	No	Yes	No	Yes
Donor Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses. Constant omitted in Models 2 and 4 due to collinearity with fixed effects.

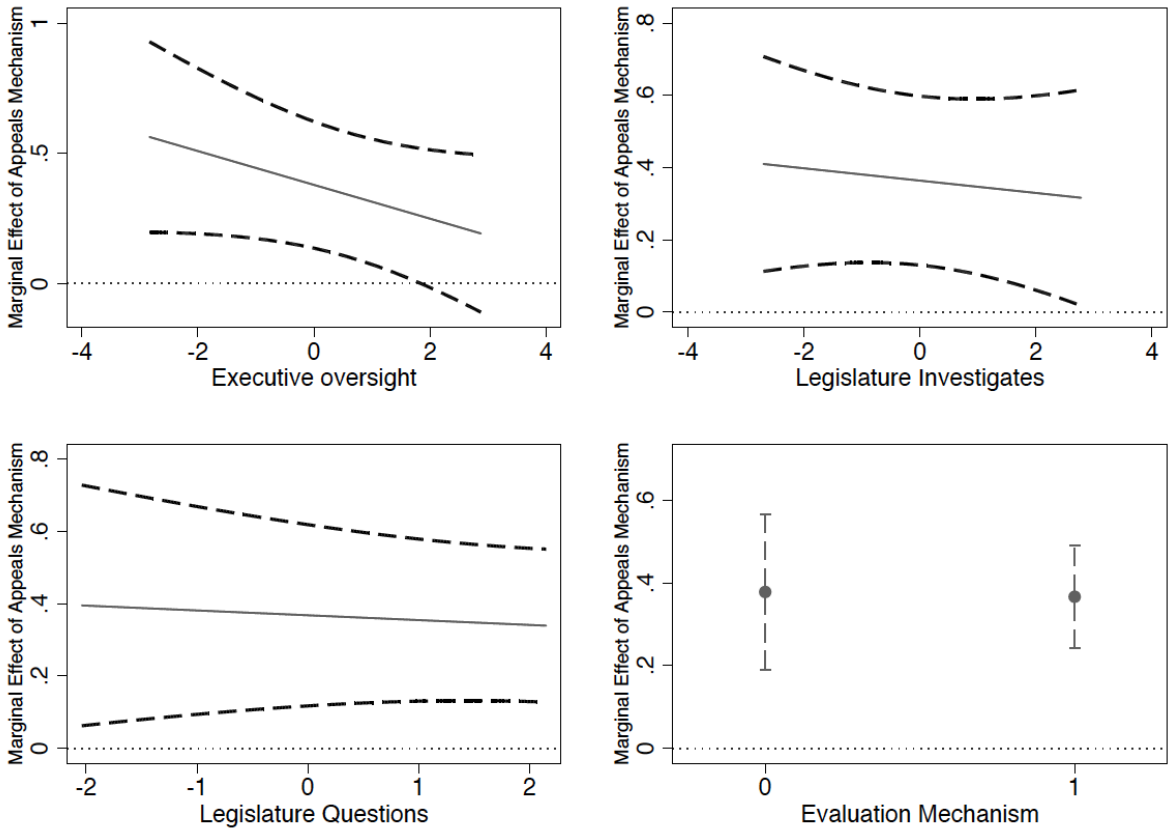
*** p<0.01, ** p<0.05, * p<0.1

Table 4. Instrumental Variables Estimates (Second Stage)

Outcome: Project Success$_p$, 1960-2016	(1)	(2)	(3)	(4)	(5)	(6)
Appeals Mechanism$_{t-3}$	0.436***	0.403***	0.471***	0.405***	0.468**	0.387**
	(0.125)	(0.115)	(0.125)	(0.122)	(0.173)	(0.154)
Observations	19,554	19,554	19,554	19,554	19,554	19,554
Recipient country, donor, and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Instrument reference group	Neighbors	Trading partners	Neighbors	Trading partners	Neighbors	Trading partners
Lags relative to treatment	1	1	2	2	3	3
Cragg-Donald F-Statistic (first stage)	11,263	13,929	6,964	7,615	3,770	3,569

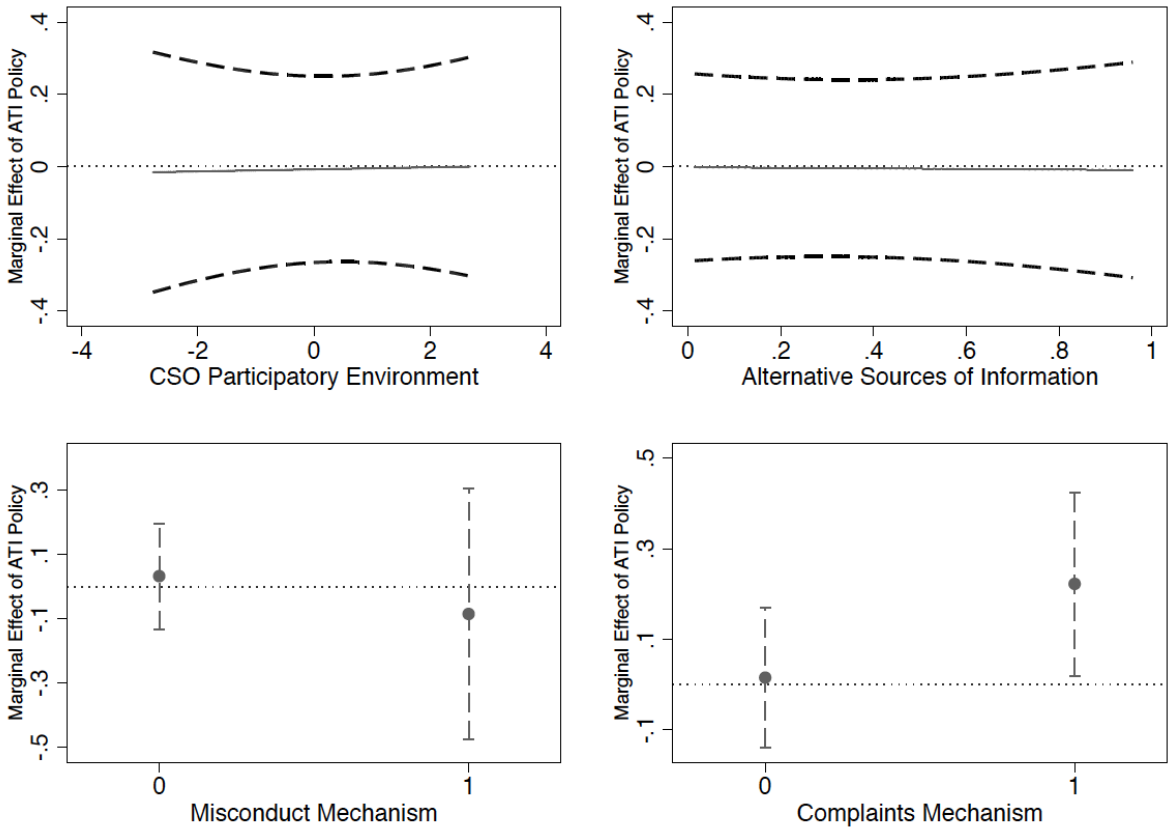
Second-stage 2SLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Figure 1. Marginal Effects of ATI Policy Appeals Mechanism on Project Success at Varying Levels of Police-Patrol Oversight



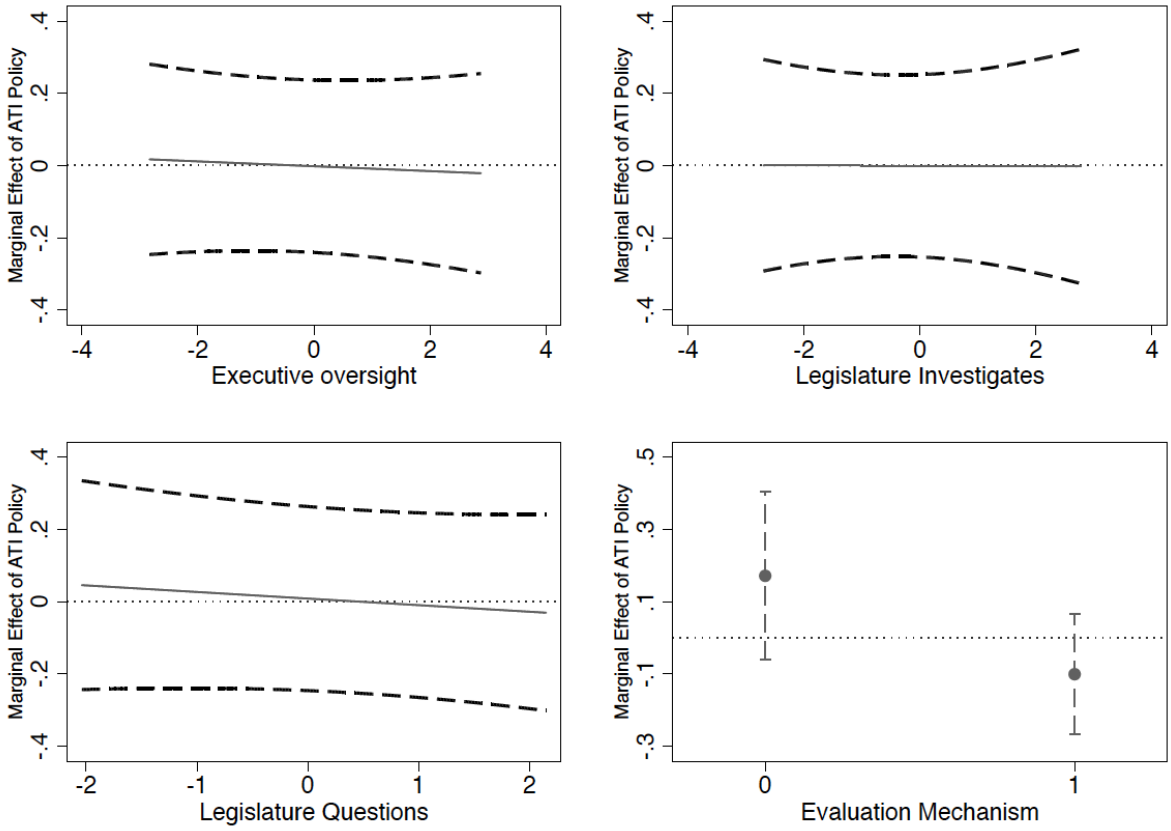
Dotted lines represent 95% confidence intervals.

Figure 2. Marginal Effects of ATI Policy Appeals Mechanism on Project Success at Varying Levels of Fire-Alarm Oversight



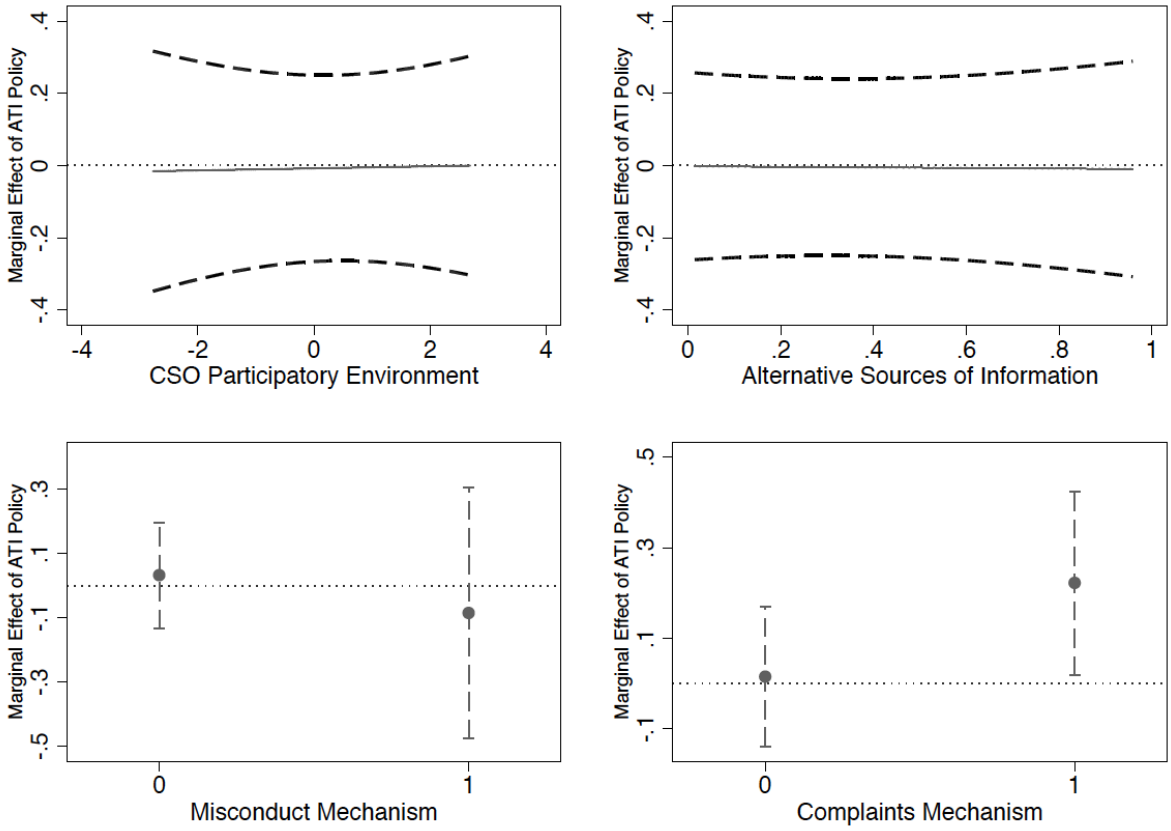
Dotted lines represent 95% confidence intervals.

Figure 3. Marginal Effects of ATI Policy on Project Success at Varying Levels of Fire-Alarm Oversight



Dotted lines represent 95% confidence intervals.

Figure 4. Marginal Effects of ATI Policy on Project Success at Varying Levels of Fire-Alarm Oversight



Dotted lines represent 95% confidence intervals.

Appendix

Further Information on Treatment Variables

Table A1. List of ATI Policies and Independent Appeals Mechanisms

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
			Public Communications Policy	2011	Independent Appeals Panel
African Development Bank	AfDB	Multilateral	Disclosure of Information Policy	1997	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 Zusammenarbeit GmbH, Germany	GIZ	Bilateral	Federal Act Governing Access to Information held by the Federal Government (Freedom of Information Act)	2005	Federal Commissioner for Freedom of Information
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 Act)	2005	Federal Commissioner for Freedom of Information
World Bank	WB	Multilateral	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 Information	2010	None

Detailed Information on Variables in Dataset

Table A2. Full Description of Variables in Dataset

Variable name	Name in dataset	Description	Scale
Project Success	six_overall_rating	Project Success Rating	Continuous, from 1-6
ATI Policy	info_pol	Donor possesses a formal information disclosure policy	Binary
Appeals Mechanism	info_app	There is an independent appeals process for all rejected information requests and the appeals body includes individuals that are independent from the organization.	Binary
Executive oversight	exec_over	From VDEM: If executive branch officials were engaged in unconstitutional, illegal, or unethical activity, how likely is it that a body other than the legislature, such as a comptroller general, general prosecutor, or ombudsman, would question or investigate them and issue an unfavorable decision or report?	Five-point ordinal, converted to standardized interval
Legislature Investigates	legis_inv	From VDEM: If the executive were engaged in unconstitutional, illegal, or unethical activity, how likely is it that a legislative body (perhaps a whole chamber, perhaps a committee, whether aligned with government or opposition) would conduct an investigation that would result in a decision or report that is unfavorable to the executive?	Six-point ordinal, converted to standardized interval
Diagonal Accountability	diag_acc	From VDEM: To what extent is the ideal of diagonal government accountability achieved?	Interval
Civil Society Participation Index	civsoc_par	From VDEM: Are major CSOs routinely consulted by policymaker; how large is the involvement of people in CSOs; are women prevented from participating; and is legislative candidate nomination within party organization highly decentralized or made through party primaries?	Ordinal
Media Self-censorship	media_cens	From VDEM: Is there self-censorship among journalists when reporting on issues that the government considers politically sensitive?	Four-point ordinal, converted to standardized interval
Engaged Society	engag_soc	From VDEM: When important policy changes are being considered, how wide and how independent are public deliberations?	Six-point ordinal, converted to standardized interval
Alternative Sources of Information	alt_info	From VDEM: To what extent is the media (a) un-biased in their coverage (or lack of coverage) of the opposition, (b) allowed to be critical of the regime, and (c) representative of a wide array of political perspectives?	Ordinal

Range of Consultation	range_cons	From VDEM: When important policy changes are being considered, how wide is the range of consultation at elite levels?	Six-point ordinal, converted to standardized interval
Complaints Mechanism	comp_mech	Donor possesses mechanism for receiving and investigating allegations of misconduct (e.g., fraud, corruption, theft)	Binary
Evaluation Mechanism	eval_ind	Donor possesses independent evaluation unit, i.e., a unit responsible for evaluating its performance that does not report to any member of the secretariat	Binary
Misconduct Mechanism	corr_unit	Donor possesses unit primarily or exclusively responsible for investigating, addressing, and preventing misconduct (e.g., fraud, corruption, theft)	Binary
Recipient GDP per Capita Growth	gdp_growth	Recipient country's GDP growth rate	Percentage
Recipient Log GDP per Capita	gdp_percapita	Log of recipient country's GDP per capita (in millions of US dollars)	Logarithmic
Recipient Aid/GNI	oda_to_gni	Recipient country's Aid to Gross National Income ratio	Ratio

Results of Conditional Models (Equation 2)

Table A3. Relationship between ATI Policy Appeals Mechanism and Project Success as Moderated by Police-Patrol Oversight

Outcome: Project Success$_t$, 1960-2016	(1)	(2)	(3)	(4)
Appeals Mechanism$_{t-3}$	0.379***	0.364***	0.378***	0.368***
	(0.099)	(0.096)	(0.095)	(0.102)
Executive Oversight$_{t-3}$	-0.049**			
	(0.018)			
Executive Oversight × Appeals Mechanism$_{t-3}$	-0.065*			
	(0.033)			
Legislature Investigates$_{t-3}$		-0.038		
		(0.022)		
Legislature Investigates$_{t-3}$ × Appeals Mechanism$_{t-3}$		-0.017		
		(0.028)		
Evaluation Unit$_{t-3}$			0.228**	
			(0.095)	
Evaluation Unit$_{t-3}$ × Appeals Mechanism$_{t-3}$			-0.012	
			(0.056)	
Legislature Questions$_{t-3}$				-0.002
				(0.027)
Legislature Questions$_{t-3}$ × Appeals Mechanism$_{t-3}$				-0.013
				(0.025)
Observations	17,213	17,251	19,554	17,257
R-squared	0.127	0.127	0.123	0.127
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. The control variables are Recipient GDP per Capita Growth$_{t-3}$, Recipient Log GDP per Capita$_{t-3}$ and Recipient Aid/GNI$_{t-3}$.

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

Table A4. Relationship between ATI Policy Appeals Mechanism and Project Success as Moderated by Fire-Alarm Oversight

Outcome: Project Success$_t$, 1960-2016	(1)	(2)	(3)	(4)
Appeals Mechanism$_{t-3}$	0.306**	0.220	0.179**	0.278***
	(0.111)	(0.128)	(0.068)	(0.086)
CSO Participatory Environment$_{t-3}$	0.014*			
	(0.008)			
CSO Participatory Environment × Appeals Mechanism$_{t-3}$	0.064			
	(0.049)			
Alternative Sources of Information$_{t-3}$		-0.066		
		(0.047)		
Alternative Sources of Information$_{t-3}$ × Appeals Mechanism$_{t-3}$		0.215		
		(0.150)		
Misconduct Mechanism$_{t-3}$			-0.251***	
			(0.067)	
Misconduct Mechanism$_{t-3}$ × Appeals Mechanism$_{t-3}$			0.367***	
			(0.095)	
Complaint Mechanism$_{t-3}$				-0.201*
				(0.097)
Complaint Mechanism$_{t-3}$ × Appeals Mechanism$_{t-3}$				0.026
				(0.074)
Observations	19,244	19,244	19,554	19,554
R-squared	0.122	0.122	0.124	0.123
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. The control variables are Recipient GDP per Capita Growth$_{t-3}$, Recipient Log GDP per Capita$_{t-3}$ and Recipient Aid/GNI$_{t-3}$.

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

Table A5. Relationship between ATI Policy and Project Success as Moderated by Police-Patrol Oversight

Outcome: Project Success$_t$, 1960-2016	(1)	(2)	(3)	(4)
ATI Policy$_{t-3}$	-0.002	-0.001	0.171	0.008
	(0.101)	(0.103)	(0.119)	(0.104)
Executive Oversight$_{t-3}$	-0.052**			
	(0.018)			
Executive Oversight × ATI Policy$_{t-3}$	-0.007			
	(0.019)			
Legislature Investigates$_{t-3}$		-0.039*		
		(0.020)		
Legislature Investigates$_{t-3}$ × ATI Policy$_{t-3}$		-0.001		
		(0.027)		
Evaluation Unit$_{t-3}$			0.247**	
			(0.109)	
Evaluation Unit$_{t-3}$ × ATI Policy$_{t-3}$			-0.272**	
			(0.120)	
Legislature Questions$_{t-3}$				0.004
				(0.024)
Legislature Questions$_{t-3}$ × ATI Policy$_{t-3}$				-0.018
				(0.023)
Observations	17,213	17,251	19,554	17,257
R-squared	0.125	0.125	0.122	0.125
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. The control variables are Recipient GDP per Capita Growth$_{t-3}$, Recipient Log GDP per Capita$_{t-3}$ and Recipient Aid/GNI$_{t-3}$.

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

Table A6. Relationship between ATI Policy and Project Success as Moderated by Fire-Alarm Oversight

Outcome: Project Success$_t$, 1960-2016	(1)	(2)	(3)	(4)
Appeals Mechanism$_{t-3}$	-0.008	-0.002	0.032	0.015
	(0.109)	(0.110)	(0.084)	(0.078)
CSO Participatory Environment$_{t-3}$	0.014**			
	(0.006)			
CSO Participatory Environment × Appeals Mechanism$_{t-3}$	0.003			
	(0.029)			
Alternative Sources of Information$_{t-3}$		-0.048		
		(0.034)		
Alternative Sources of Information$_{t-3}$ × Appeals Mechanism$_{t-3}$		-0.008		
		(0.113)		
Misconduct Mechanism$_{t-3}$			-0.121	
			(0.152)	
Misconduct Mechanism$_{t-3}$ × Appeals Mechanism$_{t-3}$			-0.118	
			(0.175)	
Complaint Mechanism$_{t-3}$				-0.426***
				(0.123)
Complaint Mechanism$_{t-3}$ × Appeals Mechanism$_{t-3}$				0.207*
				(0.103)
Observations	19,244	19,244	19,554	19,554
R-squared	0.119	0.119	0.122	0.123
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. The control variables are Recipient GDP per Capita Growth$_{t-3}$, Recipient Log GDP per Capita$_{t-3}$ and Recipient Aid/GNI$_{t-3}$.

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

Results of Additional Robustness Checks

In this section, we conduct a series of additional analyses to probe the robustness of the results in Table 3.

Counterfactual Trends Assumption

To assess the plausibility of the assumption that outcome trends in treated and control units would have been the same in the absence of the treatment—the key identifying assumption of the difference-in-differences estimator—we employ two standard strategies. First, we add to Equation 2 1-10 year leads and lags as well as a contemporaneous version of $Appeals\ Mechanism_{d,t\text{-}3}$. As reported in Table A7, almost every coefficient on the leads is zero, implying that treated and controls units do not have varying pre-treatment outcome trends in the pretreatment period as well as providing further evidence that the treatment effect is causal.[25]

Table A7. Equation 2 with leads, lags, and contemporaneous measure of treatment

Outcome: Project Success$_t$, 1960-2016	(1)
Appeals Mechanism $_{t+10}$	0.008
	(0.095)
Appeals Mechanism $_{t+9}$	0.091
	(0.089)
Appeals Mechanism $_{t+8}$	0.209**
	(0.082)
Appeals Mechanism $_{t+7}$	0.120
	(0.084)
Appeals Mechanism $_{t+6}$	-0.052
	(0.085)
Appeals Mechanism $_{t+5}$	0.048
	(0.101)
Appeals Mechanism $_{t+4}$	-0.098
	(0.075)
Appeals Mechanism $_{t+3}$	0.045
	(0.099)
Appeals Mechanism $_{t+2}$	0.174
	(0.163)
Appeals Mechanism $_{t+1}$	0.014
	(0.125)
Appeals Mechanism $_{t}$	0.037
	(0.071)
Appeals Mechanism $_{t-1}$	0.146
	(0.145)
Appeals Mechanism $_{t-2}$	0.180***
	(0.038)
Appeals Mechanism $_{t-3}$	-0.101

[25] The positive and significant coefficient on the eight-year lead is most likely spurious, given that all of the more proximate leads are non-significant.

	(0.091)
Appeals Mechanism $_{t-4}$	0.093
	(0.113)
Appeals Mechanism $_{t-5}$	0.129*
	(0.068)
Appeals Mechanism $_{t-6}$	-0.163
	(0.093)
Appeals Mechanism $_{t-7}$	0.188*
	(0.100)
Appeals Mechanism $_{t-8}$	1.377***
	(0.141)
Appeals Mechanism $_{t-9}$	
Appeals Mechanism $_{t-10}$	
Observations	15,653
R-squared	0.123
Controls	Yes
Fixed Effects	Yes

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses. The control variables are Recipient GDP per Capita Growth$_{t-3}$, Recipient Log GDP per Capita$_{t-3}$ and Recipient Aid/GNI$_{t-3}$.
*** p<0.01, ** p<0.05, * p<0.1

Second, we control for donor- and recipient country-specific linear time trends, which allows treated and control units to follow different outcome trajectories. Table A8 shows that results are unchanged.

Table A8. Table 3, Controlling for Unit-Specific Trends

Outcome: Project Success$_i$, 1960-2016	(1)
Appeals Mechanism$_{t-3}$	0.310***
	(0.066)
Observations	18,644
R-squared	0.306
Recipient Country Specific Trends	Yes
Recipient Country Fixed Effects	Yes
Donor Fixed Effects	Yes
Year Fixed Effects	Yes

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Lag Structure

We also explore sensitivity to the lag structure. The results are not at all sensitive to our choice of a 3-year lag; indeed, results with a 1-, 2-, or 4-year lag are even stronger, and results with a 5-year lag of similar magnitude.

Table A9. Varying of Time Lag for Table 3 Results

Outcome: Project Success$_t$, 1960-2016	(1)	(2)	(3)	(4)	(5)
Appeals mechanism$_{t-1}$	0.440***				
	(0.104)				
Appeals mechanism$_{t-2}$		0.415***			
		(0.088)			
Appeals mechanism$_{t-3}$			0.358***		
			(0.097)		
Appeals mechanism$_{t-4}$				0.365***	
				(0.088)	
Appeals mechanism$_{t-5}$					0.331***
					(0.088)
Observations	20,095	19,829	19,554	19,233	18,885
R-squared	0.124	0.123	0.122	0.123	0.123
All Controls in Model 4, Table 3	Yes	Yes	Yes	Yes	Yes
All Fixed Effects in Model 4, Table 3	Yes	Yes	Yes	Yes	Yes

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Placebo Test

It is also possible that the results in Table 3 are not in fact about an appeals mechanism that allows public engagement and appeal of adverse disclosure decisions per se, but about having features of an access to information policy that shift towards disclosure. To explore this possibility in Table A10 we re-run the analysis in Table 3, but examine instead of appeals mechanisms whether the disclosure policy contains a presumption of disclosure – that is, whether the policy states that the agency will disclose information absent a compelling reason not to do so. A presumption of disclosure is not associated with any positive changes in project performance, suggesting that there truly is something unique about the existence of an appeals mechanism in disclosure policies.

Table A10. Effect on Project Performance of Disclosure Policies with a Presumption of Disclosure

Outcome: Project Success$_t$, 1960-2016	(1)	(2)	(3)	(4)
Presumption of disclosure$_{t-3}$	-0.039	-0.320**	-0.046	-0.075
	(0.098)	(0.138)	(0.102)	(0.110)
Recipient GDP per Capita Growth$_{t-3}$			0.017***	0.002
			(0.004)	(0.002)
Recipient Log GDP per Capita$_{t-3}$			0.021	-0.252***
			(0.020)	(0.051)
Recipient Aid/GNI $_{t-3}$			-0.004	-0.003
			(0.002)	(0.002)
Constant	4.214***		4.052***	
	(0.081)		(0.185)	
Observations	19,554	19,554	19,554	19,554
R-squared	0.000	0.028	0.009	0.120
Recipient Country Fixed Effects	No	Yes	No	Yes
Donor Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes

OLS estimates with robust standard errors, clustered by donor and recipient country, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

In the main text we use the full richness of donors' project success ratings, following Honig's (2018, 2019) approach. However, in the literature on aid project success there is another stream of work that examines aid project success as a binary outcome. (e.g., Denizer, Kaufman, and Kraay 2013; Dollar and Svensson 2000; Dreher et al. 2013; Kilby 2009). While these scholars compress ratings to binary "success" and "failure" in a

variety of ways, the most common method is to examine whether projects are above or below the median on the scale (thus on a 6 point scale like that used by the World Bank, projects are considered successful if given a rating of 4 or above, unsuccessful if given a rating of 3 or below). In Tables A11 and A12 we employ a derivative this median approach with OLS and Logit models respectively, but in a way that takes account of differential donor interpretations of their scales, considering a project successful. That is, we variously model a project as successful if it is above the donor's mean project score (Model 1), median score (Model 2), or if and only if the project receives the highest project rating present in the data for that donor (Model 3). The results hold in all three specifications.

Table A11. Fitting OLS Model to Binary Measures of Project Success

Outcome: Binary measure of Project Success, 1960-2016	(1)	(2)	(3)
Appeals Mechanism$_{t-3}$	0.205***	0.201***	0.099***
	(0.063)	(0.064)	(0.018)
Observations	19,554	19,554	19,554
R-squared	0.114	0.144	0.217
All Controls in Model 4, Table 3	Yes	Yes	Yes
All Fixed Effects in Model 4, Table 3	Yes	Yes	Yes
DV=1 if project rating greater or equal to donor's	mean rating	median rating	highest rating

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

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

Table A12. Fitting Logit Model to Binary Measures of Project Success

Outcome: Binary measure of Project Success$_i$, 1960-2016	(1)	(2)	(3)
Appeals Mechanism$_{t-3}$	0.960***	0.957***	1.118***
	(0.127)	(0.134)	(0.155)
Constant	3.320***	2.891***	-2.791
	(0.882)	(0.830)	(1.951)
Observations	19,492	19,484	18,406
All Controls in Model 4, Table 3	Yes	Yes	Yes
All Fixed Effects in Model 4, Table 3	Yes	Yes	Yes
DV=1 if project rating greater or equal to donor's	mean rating	median rating	max rating

Estimates from logistic regression with robust standard errors, clustered by recipient country, in parentheses.

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

The Quality of Project Success Data

As mentioned in describing the project outcome ratings in the main text, project ratings have conventionally been seen as a noisy measure of project success. Part of this noise is the potential bias for project staff to rate their own projects as having performed more successfully than they actually are in practice. While in expectation there is no reason to imagine this dynamic would lead to spurious findings (indeed, inasmuch as noise inflates standard errors ceteris paribus, it would augur against spurious findings), in Table A13 we examine whether results are robust to only examining projects evaluated in a donor-year when a given donor had an independent evaluation unit.[26] Consistent with viewing this potential for more rosy self-evaluation as noise, restricting the sample leads to a lower standard error (and also a higher point estimate).

[26] While we would like to be able to examine whether a given donor project was in fact evaluated by an independent evaluation unit, we have the project-level data to determine the project's assessor for only a small subset of the donors we examine. As such we examine the presence of an evaluation unit, which we imagine greatly raises the likelihood of an independent evaluation and – given many independent evaluation units' ability to audit ratings generated by project staff – may reduce bias in project staff-generated ratings even where projects are not themselves in fact evaluated by the independent evaluation unit.

Table A13: Examining Only Projects Rated When Independent Evaluation Unit is Present

Outcome: Project Success$_i$, 1960-2016	(1)
Appeals Mechanism$_{t-3}$	0.458***
	(0.054)
Observations	13,720
R-squared	0.133
All Controls in Model 4, Table 3	Yes
All Fixed Effects in Model 4, Table 3	Yes

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

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

Another variant of the data quality concern, given the broad time period covered by these data, is that very old projects conducted when rating standards were potentially very different may be driving the results. To address this concern Table A14 restricts the sample to projects rated from 1990-2016 and 1995-2016. The substantive findings are unchanged.

Table A14. Restricting the Time Period to More Recent Decades

Outcome: Project Success$_i$, sample period	(1)	(2)
Appeals Mechanism$_{t-3}$	0.314**	0.314***
	(0.104)	(0.100)
Observations	13,802	11,553
R-squared	0.140	0.147
All Controls in Model 4, Table 3	Yes	Yes
All Fixed Effects in Model 4, Table 3	Yes	Yes
Sample Period	1990-2016	1995-2016

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

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

Sensitivity to Alternative Coding of Treatment

As mentioned in a footnote in the main text, when treatment variables are introduced, the World Bank is not considered by evaluators of disclosure policies to have an appeals mechanism, because the World Bank appeals mechanism limits the range of information regarding which appeals can be made. This runs slightly counter conventional wisdom amongst aid practitioners, who often think of the World Bank's appeals mechanism as a robust one. As such, we examine whether results are sensitive to including the World Bank's World Bank's appeals mechanism to be present (that is, equal to one) in every year where the restricted appeals mechanism existed. Table A15 indicates that the results are not sensitive to this coding decision.

Table A15. Results with Alternate Coding for World Bank Appeals Mechanism

Outcome: Project Success$_i$, 1960-2016	(1)
Appeals mechanism present $_{t-3}$	0.305***
	(0.080)
Observations	19,554
R-squared	0.122
All Controls in Model 4, Table 3	Yes
All Fixed Effects in Model 4, Table 3	Yes

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

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