Preventing the Negative Externalities of Development: Aid Compliance, Incomplete Contracts, and State Capacity*

Michael Denly†

March 8, 2023

For the most recent version of this paper, please visit:
https://mikedenly.com/research/aid-compliance-externalities

Abstract

This paper examines the potential negative externalities of foreign aid projects: that is, costs that accrue to people outside the aid transaction between the recipient state (the implementer) and the aid financier (the agent). Related literature mostly focuses on governance relationships between agents and powerful donor countries (principals), but they do not explain compliance with policies to prevent negative aid externalities. Instead, due to genuine capacity challenges, incomplete contracts between the agent and implementer, and the latter's political incentives, compliance with agents' social and environmental risk management policies mostly relates to recipient state capacity. To test the hypothesis, I compile a new dataset on states' project-level compliance with World Bank safeguard policies on involuntary resettlement, indigenous peoples, and environmental protection. Statistical support for the hypothesis suggests availability and representativeness biases in how the media and previous literature approach safeguard compliance; a need for the literature to focus more on incomplete agent-implementer contracts; and a more limited role for principal-agent relations in foreign aid.

* I thank Ivan Besserer, Olgahan Cat, Danny Cowser, Mike Findley, Xiaobo Lü, Kyle McNabb, Dan Nielson, Xin Nong, Tyler Pratt, Alex Wais, Rachel Wellhausen, and participants APSA's 2021 Annual Meeting for valuable feedback or comments. I also thank the following Research Affiliates at the University of Texas at Austin's Innovations for Peace and Development lab for assistance with the data collection: Carter Adams, Destiny Alvarez, Alyssa Ashcraft, Roxanne Bankhead, Anita Basavaraju, Sam Bennett, Mary Margaret Burniston, Mariana Caldas, Claudia Canfield, Lauren Childress, Monique Cooper, Ricardo Corona, Ji Na Gil, Hannah Greer, Maria Fernanda Guerrero, Joelean Hall, Tara Jackson, Karan Kanatala, Samiya Javed, Judy Lane, Beomhak Lee, Julianna Leung, Vanessa Lizcano, Amanda Long, Aoife McDonnell, Ruth Mewhinney, Nicole Pownall, Briana Rocha, Mackenzie Sanderson, Kimberly Schuster, Taryn Shanes, Samantha Shoff, Mackenna Shull, Ayu Sofyan, Veronica Lulu Stracher, Adityamohan Tantravahi, Ethan Tenison, Skyler Thomas, Jade Tucker, and Jonathan Velasquez. All errors are uniquely my own.

†Postdoctoral Research Fellow, Institute for Advanced Study in Toulouse: ✉ michael.denly@iast.fr

In rural Paraguay, an aid project finances the construction of a road. Typically, such roads can help project beneficiaries access markets, schooling, and medical care. However, this road cuts through the land of an indigenous community, dividing its village in half. For its part, the staff from the multilateral development bank (MDB) supervising the project learn that the community does not hold a property title for the land. The three competing claims for the land, allegedly obtained via questionable means, thus endanger the community's rightful claim for compensation from involuntary resettlement. Nearby, another indigenous community with a property title cannot protect themselves from new agricultural settlers, who are a direct result of the road. Reportedly, the settlers threaten the community with violence, burn down its school, and convert precious forests into plots, destroying ecosystem services and people's livelihoods in the process (Tello, 2015, 37-40).

As the above example underscores, foreign aid projects can have severe negative externalities: that is, social and environmental costs that accrue to people outside the aid transaction between the aid-receiving state and the aid financier. These negative externalities are common, too. For example, each year development projects cause the forced resettlement of about 15 million people (Cernea, 2008; Negi and Ganguly, 2011).[1] That is why, nowadays, all MDBs and OECD countries giving bilateral aid have safeguard policies to prevent negative externalities for resettlement, indigenous peoples, and the environment (Greenstein, 2022, 173). Nevertheless, aid-receiving states often do not implement these policies, resulting in major scandals that can even involve the loss of life. It is thus crucial to know: Under what conditions do aid-receiving states comply with aid financiers' social and environmental risk management policies and thus prevent or mitigate the potential negative externalities of development?

In terms of which actor is most to blame for the negative externalities of aid projects, the media tends to put the onus on agents such as the World Bank (e.g., Huffington Post, 2015; International Consortium of Investigative Journalists, 2015). For its part, the academic

[1]For a study on the economic impacts of population resettlement, refer to Bazzi et al. (2016).

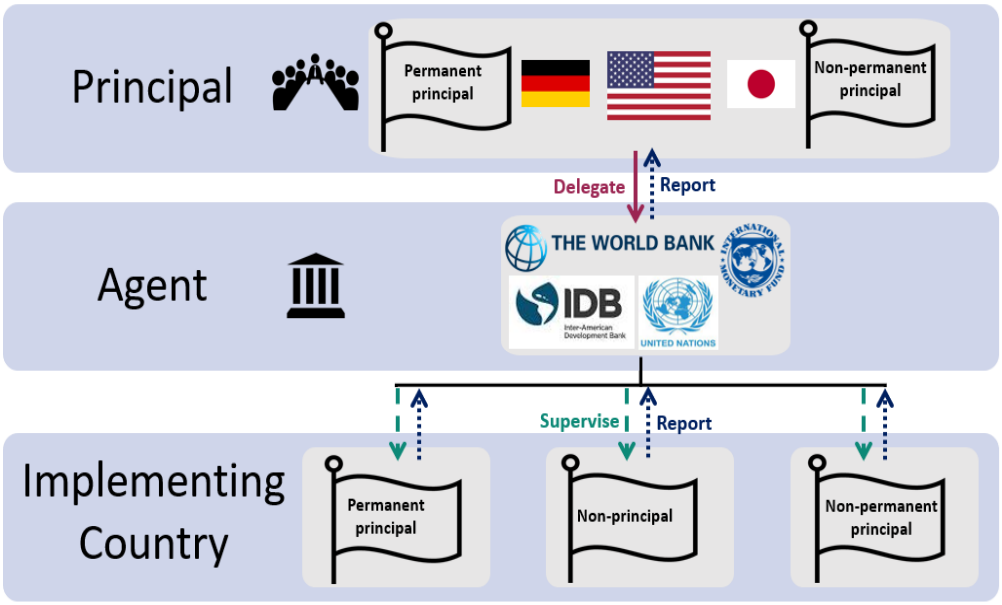
and policy literature has yet to quantitatively examine the determinants of environmental and social risk policy compliance in aid,[2] though it has examined overall levels of compliance (Girod and Tobin, 2016). The latter conclude that overall compliance is mainly a function of recipient political economy factors. By the same token, the overall academic consensus is that international organizations (agents) and their powerful donors' (principals') propensity to use aid for strategic purposes are mostly responsible for the adverse effects of aid (e.g., Easterly, 2006, 2015; Moyo, 2009; Dreher and Vreeland, 2014; Kersting and Kilby, 2016).

The received wisdom stressing principal-agent issues needs revision for two main reasons. First, consistent with the 2005 Paris Declaration on Aid Effectiveness and recent emphasis on country ownership in development, MDBs and bilateral aid providers generally do not implement their own development projects. Instead, they only design and supervise them (e.g., OECD, 2008; World Bank, 2012; Inter-American Development Bank, 2011; Asian Development Bank, 2012; African Development Bank, 2012).

The second reason why the received wisdom needs revision is that agent-implementer relationships are underpin by incomplete contracts.[3] As such, agents' abilities to engineer better project-level development outcomes are necessarily limited, because institutional constraints structure those outcomes. On that score, it is helpful to keep in mind Nielson and Tierney's (2005) clarification that agent-implementer relationships are not principal-agent relationships. Agents do not grant implementers a (conditional) delegation of authority. Almost the opposite is true: the implementer grants the agent authority to operate on its sovereign territory, so agents always supervise the implementing recipient country in a weakened position. In incomplete contracting terms, the recipient country has residual control by virtue of its ownership status/sovereignty and can "hold-up" the aid agency (e.g., Hart, 2017, 1732-1733). That becomes even more clear when considering agents' disbursement

[2] Buntaine (2016) codes his own quantitative score of environmental risk policy compliance but uses that as as independent variable, not a dependent variable, to examine aid selectivity.
[3] For an overview of incomplete contracts, see Hart (2017).

Figure 1: Principal-Agent and Agent-Implementer Interactions in Multilerateral Aid



imperative,[4] as well as the facts that agents often supervise powerful members of the UN Security Council or even their own principals (see Figure 1). In the case of the World Bank, aid recipients such as China and Brazil mostly enjoy permanent Board member positions, and countries such as Argentina generally rotate their Board positions within different country groupings (see Vreeland, 2011). Similar dynamics play out at other MDBs, too.

While some literature focuses on how recipient factors how affect aid compliance, that literature mostly focuses on political economy factors (e.g., Girod and Tobin, 2016), which is less relevant for environmental and social risk management aid compliance. It is different from overall aid compliance, because the negative consequences of the environmental and social variant have immediate political consequences. To be more precise, politicians all over the world routinely decry the World Bank and International Monetary Fund (IMF) for the conditionality on their loans, and that is often politically popular. By contrast, having an aid project destroy a rainforest or a sacred cultural site of indigenous peoples is rarely a strong political strategy. Accordingly, neither political economy factors nor principal-agent

[4]MDBs need to disburse to survive and, as a consequence, often do so when corruption or other issues should prevent them from doing so. See, for example, Booth (2011), Buntaine (2016), and Weaver (2008).

dynamics explain aid externality prevention. Instead, it is state capacity, because recipient implementers have a political incentive to address such issues.

To test the above hypothesis, I individually coded all World Bank investment lending projects approved between 2007 and 2015 with completed evaluations for social and environmental risk management (safeguard) policy compliance. These safeguard policies relate to potential environmental and social externalities, including resettlement, indigenous peoples, and the destruction of natural habitats. Although some scholars posit that aid fosters other externalities or unintended consequences, such as conflict, the prolongation of authoritarian rule, or currency fluctuations, all of these factors are either debated or are not immediate negative aid externalities. By contrast, non-compliance with safeguard policies meets the latter criterion.

I operationalize state capacity using Hanson and Sigman's (2021) new measure, which is based on a state-of-the-art Bayesian measurement model and improves upon previous measures in many ways. Then, I estimate relevant multilevel models that are able to appropriately capture that project compliance outcomes involve nested data—i.e., projects are nested within countries. Each of these models tests for multiple confounding variables at the project-level (e.g., preparation costs, commitments, and team leader skills); states' strategic relevance to powerful donors; and a battery of other controls (e.g., democracy, foreign direct investment, natural resource rents).

Consistent with the theory, I find that state capacity is what most explains compliance with safeguard policies and, by extension, the prevention of social and environmental aid externalities. Theoretically, the results suggest a need to for the academic literature to focus less on principal-agent issues and more on recipient capacity and how agents can make their incomplete contracts with implementers more complete (see Chayes and Chayes, 1993; Martens et al., 2002). That is particularly the case given that the new data suggest 22% of projects are subject to at least some safeguard policy non-compliance, and bureaucrats have

an incentive to avoid such outcomes given their career costs (see Buntaine, 2016).

From a policy perspective, the results suggests that the widespread outrage at the World Bank and other MDBs on the failure of safeguard policies to protect vulnerable people and the environment from the adverse impacts of aid may be misplaced (e.g., International Consortium of Investigative Journalists, 2015). In particular, the present paper underscores what happens in terms of broader patterns, not just the *selected sample* of projects that reach the Inspection Panel and receive more academic and media attention.[5] Essentially, the new data and empirical regularities that I show help overcome what Tversky and Kahneman (1974) famously described as availability and representativeness biases: that is, the ability of salient information and unrepresentative samples to shape human thinking in biased ways.

The article proceeds as follows. The next section conceptualizes negative externalities in foreign aid. Thereafter, I provide the main argument on state capacity and incomplete contracts. Next, I present the research design and results of the empirical analysis. The final two sections discuss the implications of the results for theory, policy, and external validity, as well as briefly conclude.

1. Negative Externalities and Foreign Aid

In the context of economics, "negative externalities are costs that accrue to parties other than [those] that produce them" (Krugman and Obstfeld, 2003, 277). Typically, negative externalities are social in nature, incur transaction costs for monitoring and governance, and outweigh the private benefits or rents that amass to the initiators of the transaction (Coase, 1960; Williamson, 1985). Accordingly, negative externalities are socially inefficient, arise from the lack of institutions such as property rights to correct for market failures, and thus harmfully affect societal provision of public goods (Arrow, 1970; North, 1981).

[5]For more on quasi-judicial bodies like the World Bank Inspection Panel, see Fox (2002) and (see Zvobgo and Graham, 2020).

Overcoming negative externalities in economics requires government regulation, a solution to the collective action problem, or an innovative arrangement between the affected parties (Olson, 1965; Ostrom, 1990; Ostrom, Walker and Gardner, 1992).

It is both possible and fruitful to apply what scholars know from negative externalities in economics toward foreign aid. In the context of foreign aid, negative externalities are the costs that accrue to the people outside the aid transaction between the aid financier (the agent) and aid-receiving state (the implementer). Generally, states have different capacities to prevent or mitigate negative externalities to people outside the aid transaction—i.e., project "beneficiaries". They are the very people that aid projects aim to support with, for example, public goods in health (e.g., vaccines), education (e.g., schooling), and infrastructure (e.g., roads, sanitation, flood protection).[6]

What, then, are the potential negative externalities of aid? Rajan and Subramanian (2011) argue that aid causes exchange rate woes and lowers economic competitiveness, but this is clearly only possible under certain circumstances: the aid flows have to be very large as compared a country's national GDP, which is usually not the case (Qian, 2015). Another set of scholars suggest that aid is a fungible non-tax revenue that prolongs the rule of authoritarian leaders, forestalls democratization, and fuels authoritarian reversals in a manner that is worse than oil (e.g., Djankov, Montalvo and Reynal-Querol, 2008; Bueno de Mesquita and Smith, 2009; Morrison, 2012). However, others contest these claims (Kono and Montinola, 2009; Altincekic and Bearce, 2014; Arndt, Jones and Tarp, 2015; Bermeo, 2016; Findley et al., 2017), and aid agencies nowadays take great measures to control corruption in their projects (Rose-Ackerman and Carrington, 2013). Yet another suite of studies argue that aid or aid shocks fuel civil conflict (Nielsen et al., 2011; Crost, Felter and Johnston, 2014; Nunn and Qian, 2014; Dube and Naidu, 2015; Wood and Sullivan, 2015). By the same token, aid did not cause conflict on its own in any of the countries under study (e.g., Colombia,

[6] OECD (2017) provides a classification of the economic sectors and activities that aid projects support. By definition, public goods are both non-rival and non-excludable, and a large percentage of aid activities meet these criteria.

Mali, the Philippines); all of these countries had pre-existing civil conflict or tensions, and aid only added fuel to the fire. Finally, Lee and Platas (2015) examine the U.S. President's Emergency Plan for AIDS Relief (PEPFAR) and find that the targeted program had adverse downstream consequences on neonatal health outcomes.

What, then, are the potential negative externalities that can apply to every aid project and do not require pre-existing country-level development challenges or the study of long and complex causal chains?[7] The negative externalities that accrue to project beneficiaries when the state does not implement the social and environmental risk management policies of the aid financier provide a more complete, direct answer. When states do not adequately implement these policies, the relevant negative externalities that arise might outweigh the benefits of providing aid in the first place. That is why all bilateral donors from OECD countries and all major MDBs have relevant safeguard policies to prevent social and environmental aid externalities (Greenstein, 2022, 173). Failure to adequately implement safeguard protections has resulted in some of the most egregious and embarrassing humiliations in the history of foreign aid and development more broadly (Nielson and Tierney, 2003; Weaver, 2008; Buntaine, 2016). Beatings, forced migration, and large-scale deforestation are just a few examples. What prevents or mitigates such externalities?

2. State Capacity, Incomplete Contracts, and Preventing the Negative Externalities of Foreign Aid

Aid-receiving states, not agents, implement projects, and the most used definitions of state capacity stress implementation (e.g., Soifer and vom Hau, 2008; Besley and Persson, 2010; Centeno, Kohli and Yashar, 2017).[8] To properly and equitably implement aid, which often comes in the form of a public good, requires not just a monopoly of violence but also

[7]For more on aid's long and complex causal chains, see Bourguignon and Sundberg (2007) and Denly (2021).
[8]All of these scholars draw from the Mann's (1984) definition of infrastructural power.

a capable, independent bureaucracy (see Weber, 1978; Evans and Rauch, 1999). States with these characteristics are generally able to enforce property rights, collect taxes in exchange for accountability, and overcome elite interests, internal conflict, and ethnic divisions that make public goods provision more difficult (Easterly and Levine, 1997; Fearon and Laitin, 2003; Miguel, 2004; Acemoglu, Johnson and Robinson, 2005; Besley and Persson, 2009, 2010). It thus follows that high-capacity states are more likely to better implement environmental and social risk management measures in foreign aid projects. Indeed, a working paper from Hanson and Sigman (2016) finds that state capacity is highly predictive of overall World Bank project outcomes as well.

Another reason why state capacity matters so much in social and environmental compliance is that aid projects are subject to incomplete contracts and recipients having a significant power advantage. Because agents need to disburse money to survive and closing poorly-performing projects is not in their economic interests, agents spend less effort on compliance matters that are not easily monitorable (Martens et al., 2002). Essentially, then, in the language of the incomplete contracts on firms, aid recipients can "hold up" agents (e.g., Hart, 2017). That becomes clearer when considering that agent-implementer interactions are not principal-agent relationships, because agents do not have sovereignty to operate on implementers' territory without their approval (Nielson and Tierney, 2005). When coupled with the fact that agents often have to supervise their principals (see Figure 1), it becomes even clearer that agents only have limited capacity to ensure compliance with its social and environmental protection policies.

The above suggests that there is scope for more attention to recipient factors. Along those lines, Girod and Tobin (2016) show that overall aid compliance is a function of recipient political economy features, such as ability to attract foreign direct investment (FDI) and natural resource wealth. However, environmental and social risk management compliance is different from overall aid compliance, as the negative consequences of the environmental

and social variant have immediate political consequences. To be more precise, politicians all over the world routinely decry the World Bank and International Monetary Fund (IMF) for the conditionality on their loans, and that is often politically popular—especially for left-wing governments (Smets, Knack and Molenaers, 2013). By contrast, having an aid project destroy a rainforest or a sacred cultural site of indigenous peoples is rarely a strong political strategy. Accordingly, neither political economy factors nor principal-agent issues are the primary determinants of negative aid externality prevention.

> Hypothesis 1: State capacity is the most crucial factor to explain the prevention of negative social and environmental aid externalities, because recipient implementers have a political incentive to address such issues.

3. Research Design

3.1. Dataset and Dependent Variable

To test the above hypotheses regarding state capacity and negative externality prevention or mitigation, I coded a new dataset of states' compliance with social and environmental risk management—i.e., safeguard policies—in World Bank projects. The World Bank was the first MDB to adopt safeguard measures in the late 1980s following severe negative externalities for failed projects in Brazil and India. In Brazil, these externalities included large-scale deforestation and the spread of tuberculosis and malaria to local, indigenous populations. In India, forced displacement led to a "long-march" of protests, ultimately resulting in 140 arrests and beatings of affected populations by local authorities (Weaver, 2008, 22-23). Since these low moments in the history of the World Bank, the institution rebounded and has served as a leader in the development of safeguard policies, yielding significant policy emulation across the different MDBs (Buntaine, 2016; Greenstein, 2022).

The dataset in the present paper only covers World Bank investment projects, which

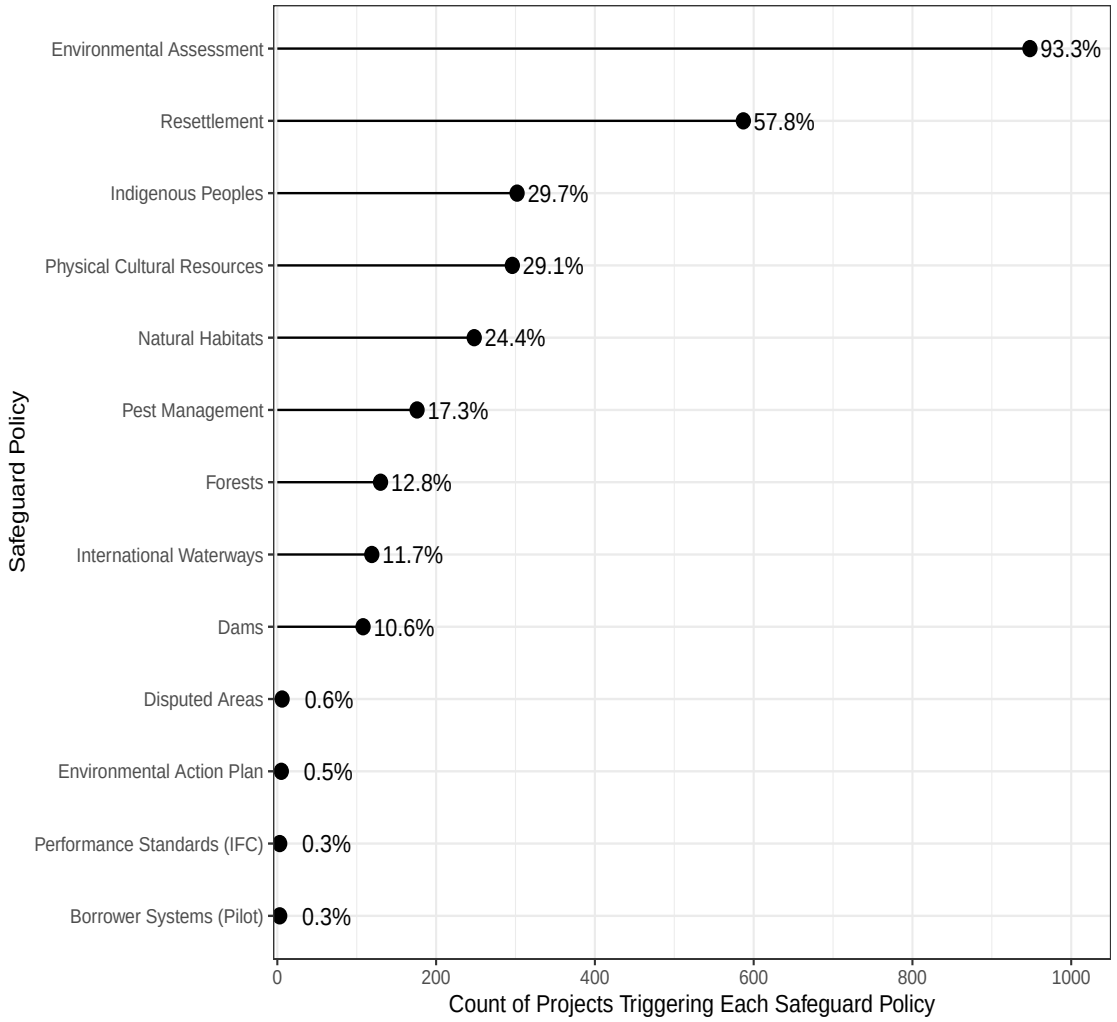
serve as the unit of analysis for this study. I exclude structural adjustment/policy projects because they do not have safeguards policies. For their part, Program for Results (PforR) projects do not have the same uniform usage of social and environmental risk management measures due to the flexible nature of the PforR lending instrument,[9] so I also exclude PforR loans.

With respect to the time period of study, the dataset covers all evaluated investment projects from 2007-2015. I chose 2007 as the starting year for two main reasons. First, although aid recipients implemented their own World Bank projects prior to the 2005 Paris Declaration on Aid Effectiveness, starting the analysis after 2005 made recipient ownership even more salient. Second, the World Bank finished converting the social and environmental risk management measures in Figure 2 from Operational Manual Statements to official Policies in 2006 (Independent Evaluation Group, 2010, 7). Starting with 2007 thereby prevents any potential problems with staff treating Operational Manual Statements different than Policies. Similarly, starting with 2007 reduces biases arising from implementers using the Operational Manual Statement and Policy distinction to reduce their safeguard policy compliance burden. Essentially, by 2007 safeguard policy inclusion happened a matter of course, as opposed to being subject to bargaining outcomes that may produce endogeneity. Given that it usually take 4-8 years to implement projects, another 6-12 months for relevant evaluation documents to be ready, and it is only possible to fully evaluate safeguard compliance on completed projects, the current ending year of the dataset is 2015. I also stopped in 2015, because the World Bank introduced a revised safeguards policy framework in 2016 (Greenstein, 2022).

Figure 2 provides a numerical breakdown of the safeguard policies triggered in the sample of 1,016 projects. The most frequently policies triggered include those regarding required environmental assessment (93%) and resettlement (58%). Projects trigger policies regard-

[9]See Winters (2010) and Winters and Kulkarni (2014) for more on the different types of World Bank lending instruments as well as when the institution decides to use one over the other.

Figure 2: Safeguard Policies Triggered in the Sample



Source: Own coding.

ing indigenous peoples (30%), physical cultural resources (30%), natural habitats (24%), and pest management (17%) with relatively high frequency as well. With relatively less frequency, projects sometimes trigger policies regarding forests (13%), international waterways (12%), and dams (11%). Projects in the sample almost never trigger policies on disputed areas, required environmental action plans, International Finance Corporation (IFC) performance standards, and pilot borrower compliance.

To measure the dependent variable, project-level compliance with World Bank safeguard policies, I coded the available evaluation documents for each project. Notably, I focused on

project Implementation Completion Reports (ICRs) and Project Performance Assessment Reports (PPARs) of the ICR and other project documents examined by the Independent Evaluation Group (IEG). Although IEG produces an overall borrower compliance score for each project that it evaluates, its scope is much broader than merely safeguards,[10] which is why the safeguard-specific coding was necessary. ICRs are generally written by consultants hired by each project's respective Task Team Leader (TTL), making them at least somewhat independent. However, the IEG PPARs provide another level of insulation against the potential downplaying safeguard issues in projects. Whenever both an ICR and an IEG PPAR is available, the coding team and I thus always base our final assessments on the IEG PPAR. By the same token, IEG does not evaluate all investment projects and tends to evaluate more positively performing projects (Kilby and Michaelowa, 2019), so the team and I always examine both the ICR and IEG PPAR. When other relevant documents are available, such as Project Papers detailing the safeguard performance of the first project in a supplemental financing project, we examine those documents as well.

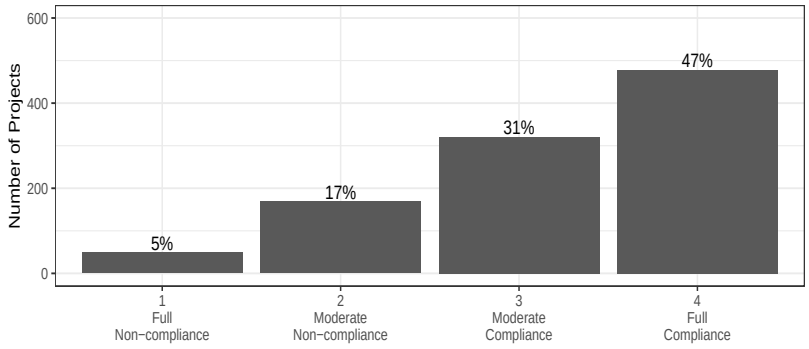
Following Buntaine (2016, 92-93), the safeguard compliance scores in the present study range from 1 (low compliance) to 4 (full compliance) corresponding with the policies in Figure 2. The difference between the present study's coding and that of Buntaine (2016) only pertains to the years under study. On that score, Buntaine (2016) examines 1990-2009, which is mostly prior to the World Bank's conversion of the safeguards Operational Directives to official Policies.

To ensure quality in the coding of safeguard compliance, all projects underwent at least two rounds of coding. One team member performed an initial review, then a more experience team member or I performed a double-check of the first team member's work. Additionally, I performed random triple-checks of some projects. When I did so, I never changed any of

[10] As Girod and Tobin (2016, 220) explain, citing Smets, Knack and Molenaers (2013), overall borrower compliance captures "the extent to which the borrower complied with covenants and agreements. The following criteria are taken into account: government ownership and commitment to achieving objectives, adequacy of stakeholder involvement, timely resolution of implementation issues, adequacy of M&E [Monitoring and Evaluation] arrangements and relationship with donors/partners."

Figure 3: Summary Statistics of World Bank Safeguard Policy Compliance (2007-2015)

(a) Distribution of Compliance Levels



(b) Map of Average Compliance Levels by Country

Source: Own coding.

the final compliance scores.

Figure 3a provides summary statistics for projects that we have coded to date, and Figure 3b provides a map of average country-level compliance scores. As Figure 3a indicates, around 78% projects generally comply with the safeguard policies, as indicated by the share of projects with a compliance score of 3 (moderate compliance) or 4 (full compliance). Circa 22% of projects exhibit moderate-to-full non-compliance, as suggested by the shares of compliance scores with only 1 (full non-compliance) or 2 (moderate non-compliance). Table A1 provides the average results by country in numerical format.

3.2. Independent Variables

3.2.1. Primary Independent Variables

I employ Hanson and Sigman's (2021) new measure for the primary independent variable, state capacity. Hanson and Sigman's (2021) state capacity variable stands out relative to competing frequentist measures.[11] Notably, Hanson and Sigman's (2021) measure uses a Bayesian measurement model to combine multiple indicators and overcome missing data challenges in a way that competing frequentist measures cannot (see Fariss, Kenwick and Reuning, 2020). To capture the multiple, latent dimensions of state capacity, Hanson and Sigman (2021, 1502) include indicators on administrative efficiency, bureaucratic quality, budgetary quality, census frequency, fiscal capacity, information capacity, law and order, military expenditures, police expenditures, public administration, and taxation. In the present dataset, Hanson and Sigman's (2021) state capacity variable ranges from -1.667 to 1.55.

3.2.2. State-Level Control Variables

Perhaps the main impediment to the implementation of safeguard policies is corruption. To control for corruption, I include the World Bank's Worldwide Governance Indicators' Control of Corruption estimate.[12] Given that many resettlement and social safeguards issues often arise from lack of property rights (Tello, 2015), I also control for the Varieties of Democracy (V-Dem) project's measure property rights protection (see Lindberg et al., 2014). Because democracies are generally better at environmental protection than autocracies (e.g., Bernauer and Koubi, 2009), I further include V-Dem's overall democracy measure, which does not have the same measurement challenges as the commonly-used Polity measure (see Vreeland, 2008).

Another set of mandatory control variables include those related to the strategic interests of powerful donor countries on the World Bank's Board of Directors. To control for the

[11] See, for example, Hendrix (2010), Lee and Zhang (2017), and O'Reilly and Murphy (2022).
[12] See Kaufmann, Kraay and Mastruzzi (2011) for a fuller description of the WGI.

formal influence these donor countries, I follow Kaja and Werker (2010) and add an indicator of whether the aid-receiving state was a member of the World Board during the year of project approval. To capture the informal influence of these powerful donor countries, I follow Dreher, Sturm and Vreeland (2009) and include an indicator variable of whether the country was a temporary member of the UN Security Council at the time of project approval. Additionally, I include Bailey, Strezhnev and Voeten's (2017) measure of the aid-receiving state's ideal point distance from the United States, the World Bank's most powerful shareholder. As Bailey, Strezhnev and Voeten (2017) explain, their ideal point measure captures dynamic state preferences through UN General Assembly voting and correlates at 0.92 with votes deemed "important" votes by the US State Department.

The final set of state-level control variables come from the World Bank's (2017) World Development Indicators (WDI). Given that wealthier countries may have greater bureaucratic infrastructures to deal with potential externalities in aid projects, I include a measure of log GDP per capita, deflated in constant 2015 US dollars. In line Girod and Tobin (2016), I also control for natural rents as a share of GDP and foreign direct investment as a share of GDP. Although safeguards are only one very portion of Girod and Tobin's (2016) measure of borrower compliance, it is feasible that these variables are relevant for safeguard compliance as well. Generally, the UNU-WIDER (2021) Global Revenue Dataset generally has better coverage than the WDI on most revenue measures, but that is not the case for natural resource revenues. That is why I retain the WDI measure of resource rents for the present study.

3.2.3. Project-Level Control Variables

Following Denizer, Kaufmann and Kraay (2013) and Bulman, Kolkma and Kraay (2017), outcomes for safeguards, or any other project component, depend on project-specific features, especially the quality of the Task Team Leader (TTL). To capture the quality of the TTL, I made a transparency request to the World Bank to obtain full data on the name of

the TTL at each mandatory, 6-month Implementation Status Report (ISR) for each project. After three rounds of back-and-forth, the World Bank sent me the data. I then combined these data with IEG data on project outcomes to calculate a TTL quality score prior to the approval date of each respective project. Given that the TTL often switches during a project's implementation, by necessity the TTL quality score reflects a weighted average of the number of ISRs for each TTL is responsible for each project. First-time TTLs do not have a project outcome score history, so my weighted average of TTL quality prior to each project treats these instances as missing. When there are at least two TTLs on a project, and one of them has a project outcome history, I proportionally re-assign the relevant weighting to the non-first-time TTLs to overcome the missing data challenge.

Other potentially important project-level confounding variables include project preparation costs and commitment amounts. Higher amounts spent on preparation might help eliminate safeguard compliance issues, and projects with higher commitment amounts might garner more attention during preparation. I include and deflate both variables to constant 2015 US dollars and then take their natural logs to decrease the risk of overdispersion affecting the results. Finally, although supervision costs may proxy for supervision effort designed to reduce safeguard compliance issues, I exclude them because they are clearly post-treatment to any safeguard issues that arise—or what Angrist and Pischke (2008) call a "bad control". Additionally, the direction of post-treatment bias is hard to predict from a statistical perspective (Montgomery, Nyhan and Torres, 2018, 760).

3.3. Empirical Strategy

A primary empirical strategy of this paper involves a multilevel ordered logit model with robust standard errors:

$$Pr(y^*_{compliance(i,j)}) = \wedge\left(\alpha_{country(j[i])} + \beta_{\text{state capacity}_{(i,j)}} + \beta_{controls_{(i,j)}}\right) \tag{1}$$

where subscripts i refer to the project and j to the implementing country; $\beta_{\text{state capacity}_{i,j}}$ represents the primary independent variable of interest; $\beta_{controls_{i,j}}$ are the aforementioned control variables; and $\alpha_{country(j[i])}$ is a random intercept that captures the (mostly) time-invariant country-level factors j for project i. Because the dependent variable, $y^*_{compliance(i,j)}$, has four ordered categories, it is possible to classify $y^*_{compliance(i,j)}$ in the following way, where τ_i are the cutpoints for each imposed category:

$$y_{compliance(i,j)} = \begin{cases} 1 \text{ (full non-compliance)}, & if \quad y^*_{compliance} \leq \tau_2 \\ 2 \text{ (moderate non-compliance)}, & if \quad \tau_2 < y^*_{compliance} \leq \tau_3 \\ 3 \text{ (moderate compliance)}, & if \quad \tau_3 < y^*_{compliance} \leq \tau_4 \\ 4 \text{ (full compliance)}, & if \quad \tau_4 < y^*_{compliance} \end{cases} \tag{2}$$

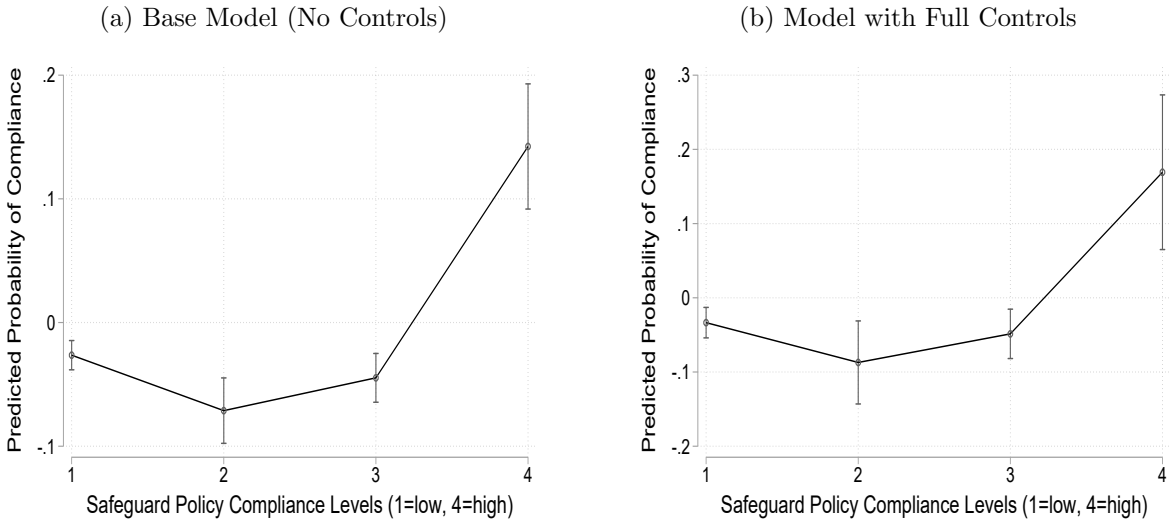
The multilevel specification is necessary because non-multilevel models assume independent error structures. Similar to Girod and Tobin's (2016) examination of overall, non-safeguard specific, aid compliance, such independent error structures are extremely unlikely to be relevant for the present estimations as well. Compliance rates are likely correlated across countries given different institutional factors, which makes the country-level nesting structure appropriate. As a robustness check, I also account for time variation by adding a random intercept $\nu_{year(t[i])}$, yielding the following specification:

$$Pr(y^*_{compliance(i,j,t)}) = \wedge \left(\beta_0 + \beta_{\text{state capacity}_{(i,j,t)}} + \beta_{controls_{(i,j,t)}} + \alpha_{country(j[i])} + \nu_{year(t[i])}\right) \tag{3}$$

4. Results

Figure 4 presents the results from the main test of Hypothesis 1 in terms of average marginal effects, using the main country random intercept specification from Equation (1). In the base model without covariates, a one-unit increase in state capacity increases the probability of that respective project has full safeguard compliance by 14%. That number is

Figure 4: Marginal Effects of State Capacity on Compliance (Country Intercept)



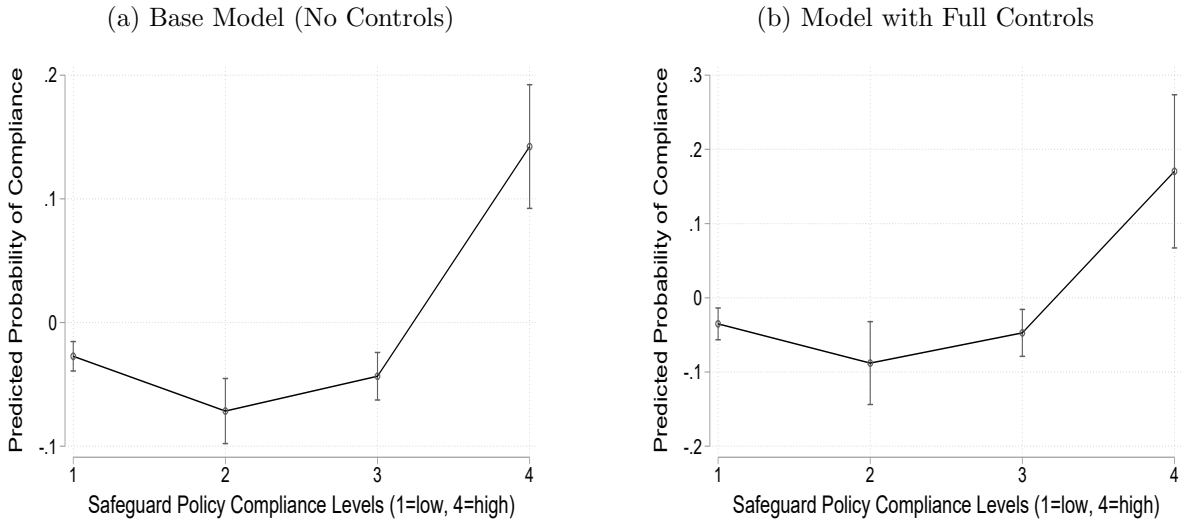
Note: Ordered logit model with a country random intercept as well as robust standard errors.

even higher for the full model with all controls, where it exceeds 17%. Given that Hanson and Sigman's (2021) state capacity variable ranges from circa -1.67 to 1.55 in this dataset, a one-unit increase in state capacity corresponds to a 32% increase in the variable.[13] Extrapolating further for the results from the full model, a 1% increase in state capacity leads to 0.53% increase in safeguard policy compliance.[14] These magnitudes are very high.

An indicator of the consistency in state capacity's ability to explain compliance is that it is statistically significant at the 1% level in both the base model and full models (see Table B1). Besides the logged and deflated commitment and GDP per capita variables, state capacity is the only variable that is statistically significant in both the base and the full models. Although it is difficult to fully interpret control variables (Cinelli and Hazlett, 2020), the statistical non-significance of these control variables may be informative (Abadie, 2020). Again, these control variables include ones that have formed the basis for the international relations literature examining foreign aid over the past 20 years.

[13] $1/(1.67 + 1.55) = 0.32$
[14] $0.17/0.32 = 0.53$

Figure 5: Marginal Effects of State Capacity on Compliance (Country and Year Intercepts)



Note: Ordered logit model with country and year random intercepts as well as robust standard errors.

Figure 5 presents results from the models with country and year random intercepts, consistent with Equation (3). Overall, the results are are almost indentical to those from Figure 4. As Figures 5a and 5b demonstrate, the marginal effects of a one-unit increase in state capacity are 14% in the base model and 17% in the full model. Both results are also statistically significant at the 1% level (see Table B2).

5. Implications for Theory, Policy, and External Validity

It is worth repeating that the above results correspond to the World Bank, which is the aid agency with the most developed environmental and social safeguard policies to prevent negative aid externalities. Over time, other MDBs and bilateral aid agencies have emulated the World Bank's policies (Greenstein, 2022), but that emulation and implementation after emulation takes time. It is thus clear that the above results correspond to what case study scholars call an "extreme case" (see Gerring, 2017). What that means here is that the World

Bank is more likely than other aid agencies to have the ability to design projects that prevent or mitigate aid externalities. Nevertheless, a factor like staff preparation costs is not what matters most from a statistical perspective; it is state capacity. By extension, state capacity is likely to be even more essential for determining safeguard outcomes at other MDBs and bilateral aid agencies with less safeguard policy experience.

More broadly, the importance of state capacity in driving results calls into question the focus of previous media exposés on safeguards as well as the previous academic literature. In line with the Paris Declaration on Aid Effectiveness (2005), the World Bank generally does not implement its own projects, so it also does not implement its own safeguard policies. Furthermore, the cases that reach quasi-judicial accountability bodies like the World Bank Inspection Panel are a *selected sample* extreme safeguard policy failures.[15] For this reason, the focus of media exposés—including that of the highly-respected International Consortium of Investigative Journalists (2015)—and the academic literature on inspection panels and their findings does not provide a full picture safeguard compliance. More specifically, previous media exposés and literature are subject to what Tversky and Kahneman (1974) famously called availability and representative biases. Only by examining the full range of safeguard policy outcomes can scholars and lay audiences understand that compliance, not significant policy failure, is the norm.

Another implication of the findings is that the dominant framework for theorizing about foreign aid, the principal-agent model, likely needs some revision to better capture dynamics with recipient countries. To be clear, the seminal work from Pollack (1997), Nielson and Tierney (2003), and Hawkins et al. (2006), among others, remains highly relevant and useful for capturing dynamics between principals and agents. However, the literature's focus on principals and agents has come at the neglect of recipient state capacity issues, and the statistical results in the present paper suggest that neglect is costly. To remedy this issue, I propose supplementing the principal-agent model with a second level focusing on the in-

[15]For more on inspection panels, see Fox (2002).

complete contract between the agent and aid recipient country implementer (see Figure 1). Doing so will enable the academic literature to make better sense of the true nature of foreign aid, which is subject to hold-up problems and other agent-implementer power dynamics that prevent agents from *engineering* outcomes in the way that the principal-agent model suggests. After all, all foreign aid—regardless of whether it is bilateral or multilateral in nature—has not two but three main actors: principals, agents, and implementers.

Finally, it is necessary to acknowledge two statistical limitations. First, the allocation of both of project and state capacity is far from randomly assigned, so it is unfortunately impossible to draw causal conclusions with the above statistical analyses. Second, as I specified above, I did not include any variable to directly capture the influence of World Bank supervision efforts due to post-treatment bias. Clearly, supervision efforts matter for aid externality prevention and mitigation, but it is unclear how to adequately show that statistically.

6. Conclusion

Even if an aid project accomplishes all of its objectives, its potential negative externalities—such as destruction of habitats, involuntary resettlement, and the loss of indigenous cultural property—can outweigh the benefits of undertaking an aid project in the first place. The present paper is thus not about reaching aid effectiveness targets but making sure that those targets are not achieved at all costs. In contrast to the previous literature on aid externalities, which focuses on externalities deriving from long causal chains, the present paper examines the direct social and environmental negative externalities of aid. It does so by examining aid recipients' project-level compliance with World Bank social and environmental safeguard policies (see Figure 2).

I find that state capacity is the strongest predictor of safeguard policy compliance and, by extension, the prevention or mitigation of negative aid externalities. These results are con-

sistent with my broader theory. It stresses that agent-implementer interactions are subject to incomplete contracts, hold-up problems, and political incentives that differ from typical aid compliance scenarios (see Girod and Tobin, 2016). In the case of safeguard policy compliance, noncompliance in the form of, say, destruction of a rainforest or indigenous people's land rarely constitutes strong political strategy. That is why aid recipients, regardless of their political regime, generally attempt to avoid negative aid externality outcomes if their state has the capacity to do so. Clearly, the large literature on compliance comes to a different conclusion, stressing that compliance is rarely so straightforward or self-enforcing (e.g., Simmons, 2010). By the same token, that is why environmental and social aid compliance is distinct from broader aid compliance and deserves further study.

More broadly, the present study responds to what Falleti (2021) called the "invisibility" of indigenous issues to political science and provides some cross-country, quantitative data to better understand these topics and related environmental ones.[16] I say "better understand", because the literature and media has to date focused on anthropological case studies and a very *selected sample* projects that reach accountability bodies like the Inspection Panels (e.g., Fox, 2002; International Consortium of Investigative Journalists, 2015; Tello, 2015; Zvobgo and Graham, 2020). To further overcome availability and representativeness biases, future research can use the new data advanced in this article to investigate when state capacity breaks down. By doing so, scholars and policymakers will be able to better protect indigenous communities and the environment from the negative externalities of foreign aid.

[16] Notably, Falleti (2021) finds that that the top-3 journals—*American Journal of Political Science*, *American Political Science Review*, and *Journal of Politics*—only published 11 articles with the word "indigenous" in the title or abstract from 1990-2020.

Appendix A Safeguard Compliance Scores by Country

Table A1: Compliance Scores by Country (Coded Projects)

Country	N	Mean	Std. Dev.	Minimum	Maximum
Afghanistan	16	2.75	.9309493	1	4
Albania	4	3.5	1	2	4
Angola	4	3	1.414214	1	4
Argentina	21	3.380952	.6690434	2	4
Armenia	9	3.444444	.7264832	2	4
Azerbaijan	10	3.5	.8498366	2	4
Bangladesh	19	3.526316	.6117753	2	4
Barbados	0	.	.	.	.
Belarus	4	3.75	.5	3	4
Benin	10	3.1	.7378648	2	4
Bhutan	5	3.6	.5477226	3	4
Bolivia	8	3.5	.7559289	2	4
Bosnia and Herzegovina	7	3.428571	.9759001	2	4
Botswana	1	2	.	2	2
Brazil	43	3.139535	.8885889	1	4
Bulgaria	4	3.75	.5	3	4
Burkina Faso	12	3.25	.7537784	2	4
Burundi	8	3.25	.8864053	2	4
Cape Verde	1	4	.	4	4
Cambodia	6	3.833333	.4082483	3	4
Cameroon	10	3	.942809	2	4
Central African	4	2.5	.5773503	2	3

Continued on next page

Table A1 : Compliance Scores by Country (Coded Projects) – *continued*

Country	**N**	**Mean**	**Std. Dev.**	**Minimum**	**Maximum**
Chad	4	3	1.414214	1	4
Chile	1	4	.	4	4
China	80	3.5625	.6721164	1	4
Colombia	10	3.6	.6992059	2	4
Comoros	1	4	.	4	4
Democratic Republic of Congo	8	2.875	1.125992	1	4
Republic of Congo	5	3.2	.83666	2	4
Costa Rica	2	3.5	.7071068	3	4
Cote d'Ivoire	9	3	1.118034	1	4
Croatia	8	3.625	.5175492	3	4
Djibouti	6	3.166667	.4082483	3	4
Dominican Republic	6	4	0	4	4
Ecuador	1	4	.	4	4
Egypt	13	2.769231	1.012739	1	4
El Salvador	4	3.5	.5773503	3	4
Eritrea	0	.	.	.	.
Ethiopia	21	2.380952	1.07127	1	4
Gabon	2	3.5	.7071068	3	4
Gambia, The	2	3.5	.7071068	3	4
Georgia	8	2.75	.8864053	1	4
Ghana	12	3	1.128152	1	4
Guatemala	3	4	0	4	4
Guinea	4	3.5	.5773503	3	4

Continued on next page

Table A1 : Compliance Scores by Country (Coded Projects) – *continued*

Country	**N**	**Mean**	**Std. Dev.**	**Minimum**	**Maximum**
Guinea-Bissau	4	3	1.414214	1	4
Guyana	1	4	.	4	4
Haiti	14	2.571429	.9376145	1	4
Honduras	10	3.5	.7071068	2	4
India	55	3.145455	.9313109	1	4
Indonesia	20	2.9	.8522416	1	4
Iraq	2	1.5	.7071068	1	2
Jamaica	4	3.75	.5	3	4
Jordan	3	4	0	4	4
Kazakhstan	5	2.8	1.30384	1	4
Kenya	15	2.733333	1.222799	1	4
Kosovo	3	4	0	4	4
Kyrgyz Republic	10	3.6	.6992059	2	4
Lao PDR	10	3.6	.6992059	2	4
Lebanon	1	4	.	4	4
Lesotho	5	3.4	.8944272	2	4
Liberia	8	2.75	1.38873	1	4
Macedonia	4	2.5	.5773503	2	3
Madagascar	5	2.8	.83666	2	4
Malawi	9	2.777778	.8333333	2	4
Maldives	2	2	1.414214	1	3
Mali	6	2.5	.83666	1	3
Mauritania	2	3	1.414214	2	4
Mauritius	1	4	.	4	4

Continued on next page

Table A1 : Compliance Scores by Country (Coded Projects) – *continued*

Country	**N**	**Mean**	**Std. Dev.**	**Minimum**	**Maximum**
Mexico	13	3.538462	.77625	2	4
Moldova	6	3.666667	.8164966	2	4
Mongolia	5	3.6	.8944272	2	4
Montenegro	7	3.571429	.5345225	3	4
Morocco	7	3	.5773503	2	4
Mozambique	14	3.214286	.8925824	2	4
Myanmar	1	4	.	4	4
Nepal	16	2.875	.7187953	2	4
Nicaragua	10	3.4	.6992059	2	4
Niger	8	3.125	.6408699	2	4
Nigeria	21	3.142857	.853564	2	4
Pakistan	15	3.066667	1.032796	1	4
Panama	5	2.4	1.140175	1	4
Papua New Guinea	7	3.285714	.7559289	2	4
Paraguay	3	3	1	2	4
Peru	13	3.230769	1.012739	1	4
Philippines	8	3.5	.9258201	2	4
Poland	0	.	.	.	.
Romania	1	4	.	4	4
Russia	2	3.5	.7071068	3	4
Rwanda	8	3.375	1.187735	1	4
Sao Tome and Principe	2	3.5	.7071068	3	4
Senegal	8	2.875	.6408699	2	4
Serbia	6	3.166667	1.169045	1	4

Continued on next page

Table A1 : Compliance Scores by Country (Coded Projects) – *continued*

Country	**N**	**Mean**	**Std. Dev.**	**Minimum**	**Maximum**
Sierra Leone	4	3	.8164966	2	4
Solomon Islands	2	4	0	4	4
South Africa	0	.	.	.	.
South Sudan	2	2.5	.7071068	2	3
Sri Lanka	12	3.25	.9653073	1	4
Swaziland	2	3	0	3	3
Tajikistan	6	3.333333	1.21106	1	4
Tanzania	15	3.066667	.7037316	2	4
Thailand	0	.	.	.	.
Timor-Leste	5	3.4	.8944272	2	4
Togo	4	3	.8164966	2	4
Tunisia	6	3.166667	.9831921	2	4
Turkey	8	3.375	.9161254	2	4
Uganda	13	2.461538	.9674179	1	4
Ukraine	4	4	0	4	4
Uruguay	5	3	0	3	3
Uzbekistan	7	3	1	1	4
Vietnam	37	3.594595	.5990483	2	4
Yemen, Republic	5	2.4	1.140175	1	4
Zambia	6	2.333333	1.21106	1	4
Total	944	3.202331	.8944004	1	4

Source: Own coding. Note: "0" values indicate country is included in dataset but not coded yet.

Appendix B Additional Regression Tables

Table B1: Compliance with World Bank Safeguard Policies 2007-2015
(Ordered Logit Model with Country Random Intercept)

	(1)	(2)	(3)	(4)	(5)	(6)
State capacity	1.8623***	1.5220**	1.6792**	1.6905**	1.6853**	2.1280***
	(0.2209)	(0.3052)	(0.3993)	(0.4003)	(0.3975)	(0.5038)
GDP per capita (log)		1.2622**	1.2599**	1.2674*	1.2955**	1.3090**
		(0.1462)	(0.1461)	(0.1581)	(0.1560)	(0.1611)
Democracy		0.4486*	0.5076	0.4804	0.4699	0.6175
		(0.2131)	(0.2600)	(0.2547)	(0.2477)	(0.3352)
Property rights		1.1283	1.1221	1.1006	0.9432	0.6010
		(0.5923)	(0.5787)	(0.5806)	(0.4792)	(0.3299)
Corruption control			0.8151	0.8132	0.6937	0.6078*
			(0.2232)	(0.2229)	(0.1948)	(0.1639)
US ideal point dist.				1.0343	0.9566	0.9234
				(0.1863)	(0.1741)	(0.1678)
Temp. UNSC				0.9525	0.9930	1.0733
				(0.1902)	(0.1873)	(0.2214)
Civil war					0.7109**	0.7819
					(0.1161)	(0.1375)
FDI					1.0299*	1.0240
					(0.0156)	(0.0154)
Natural resources					0.9875	0.9876
					(0.0092)	(0.0089)
Commitment (log)						0.7706***
						(0.0584)
Preparation cost (log)						0.9009
						(0.0735)
Cutpoint 1	0.0463***	0.1944*	0.2347	0.2149	0.3066	0.0317**
	(0.0081)	(0.1876)	(0.2315)	(0.3084)	(0.4284)	(0.0536)
Cutpoint 2	0.2837***	1.1846	1.4304	1.3031	1.8668	0.1981
	(0.0275)	(1.0845)	(1.3416)	(1.8187)	(2.5507)	(0.3278)
Cutpoint 3	1.3079***	5.3183*	6.4190*	5.8673	8.4290	0.8985
	(0.1187)	(4.9275)	(6.1128)	(8.2744)	(11.6462)	(1.4954)
Country Intercept	1.2897***	1.1914**	1.1833**	1.1789**	1.1187*	1.1411**
	(0.1067)	(0.0973)	(0.0920)	(0.0896)	(0.0733)	(0.0708)
Observations	990	983	983	978	978	957

Coefficients are odds ratios; robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B2: Compliance with World Bank Safeguard Policies 2007-2015
(Ordered Logit Model with Country and Year Random Intercepts)

	(1)	(2)	(3)	(4)	(5)	(6)
State capacity	1.9076***	1.5445**	1.7127**	1.7248**	1.7171**	2.2180***
	(0.2326)	(0.3215)	(0.4241)	(0.4246)	(0.4210)	(0.5571)
GDP per capita (log)		1.2733**	1.2703**	1.2821**	1.3128**	1.3285**
		(0.1486)	(0.1487)	(0.1616)	(0.1592)	(0.1653)
Democracy		0.4226*	0.4819	0.4581	0.4437	0.5878
		(0.2117)	(0.2607)	(0.2557)	(0.2445)	(0.3315)
Property rights		1.1955	1.1855	1.1753	1.0084	0.6225
		(0.6353)	(0.6186)	(0.6261)	(0.5236)	(0.3482)
Corruption control			0.8076	0.8051	0.6820	0.5918*
			(0.2271)	(0.2266)	(0.1960)	(0.1632)
US ideal point dist.				1.0133	0.9361	0.8946
				(0.1866)	(0.1733)	(0.1674)
Temp. UNSC				0.9441	1.0019	1.0903
				(0.1972)	(0.2012)	(0.2354)
Civil war					0.7025**	0.7830
					(0.1240)	(0.1506)
FDI					1.0317**	1.0254*
					(0.0158)	(0.0156)
Natural resources					0.9870	0.9871
					(0.0095)	(0.0092)
Commitment (log)						0.7564***
						(0.0616)
Preparation cost (log)						0.8961
						(0.0769)
Cutpoint 1	0.0430***	0.1934*	0.2346	0.2360	0.3435	0.0315**
	(0.0079)	(0.1905)	(0.2364)	(0.3443)	(0.4852)	(0.0544)
Cutpoint 2	0.2739***	1.2279	1.4896	1.4929	2.1832	0.2070
	(0.0277)	(1.1406)	(1.4195)	(2.1160)	(3.0134)	(0.3492)
Cutpoint 3	1.3352***	5.8379*	7.0798**	7.1340	10.4555*	1.0029
	(0.1239)	(5.4345)	(6.7877)	(10.1950)	(14.5591)	(1.6945)
Country Intercept	1.2665***	1.1583*	1.1516*	1.1440	1.0836	1.1061
	(0.1104)	(0.1029)	(0.0977)	(0.0950)	(0.0804)	(0.0776)
Year Intercept	1.2346*	1.2502*	1.2493*	1.2604*	1.2580*	1.2891**
	(0.1419)	(0.1495)	(0.1480)	(0.1537)	(0.1580)	(0.1648)
Observations	990	983	983	978	978	957

Coefficients are odds ratios; robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

References

Abadie, Alberto. 2020. "Statistical Nonsignificance in Empirical Economics." *American Economic Review: Insights* 2(2):193–208.

Acemoglu, Daron, Simon Johnson and James A. Robinson. 2005. Institutions as a Fundamental Cause of Long-Run Growth. In *Handbook of Economic Growth*, ed. Philippe Aghion and Steven N Durlauf. Vol. 1A Elsevier B.V. pp. 385–472.

African Development Bank. 2012. "Rules and Procedures for the Use of Consultants.".

Altincekic, Ceren and David H. Bearce. 2014. "Why There Should Be No Political Foreign Aid Curse." *World Development* 64(October 2012):18–32.

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

Arndt, Channing, Sam Jones and Finn Tarp. 2015. "Assessing Foreign Aid's Long-run Contribution to Growth and Development." *World Development* 69:6–18.

Arrow, Kenneth J. 1970. Political and Economic Evaluation of Social Effects and Externalities. In *The Analysis of Public Output*, ed. Julius Margiolis. NBER chapter 1, pp. 1–23.

Asian Development Bank. 2012. "Handbook on Project Implementation.".
URL: *http://www.adb.org/documents/handbook-project-implementation*

Bailey, Michael A., Anton Strezhnev and Erik Voeten. 2017. "Estimating Dynamic State Preferences from United Nations Voting Data." *Journal of Conflict Resolution* 61(2):430–456.

Bazzi, Samuel, Arya Gaduh, Alexander D. Rothenberg and Maisy Wong. 2016. "Skill Transferability, Migration, and Development: Evidence from Population Resettlement in Indonesia." *American Economic Review* 106(9):2658–2698.

Bermeo, Sarah Blodgett. 2016. "Aid Is Not Oil: Donor Preferences, Heterogeneous Aid, and the Aid-Democratization Relationship." *International Organization* 70(1):1–32.

Bernauer, Thomas and Vally Koubi. 2009. "Effects of Political Institutions on Air Quality." *Ecological Economics* 68(5):1355–1365.

Besley, Timothy and Torsten Persson. 2009. "Property Rights, Taxation, and Politics: The Origins of State Capacity." *American Economic Review* 99(4):1218–1244.

Besley, Timothy and Torsten Persson. 2010. "State Capacity, Conflict, and Development." *Econometrica* 78(1):1–34.

Booth, David. 2011. "Aid, Institutions and Governance: What Have We Learned?" *Development Policy Review* 29:s5–s26.

Bourguignon, François and Mark Sundberg. 2007. "Aid Effectiveness: Opening the Black Box." *American Economic Review: Papers and Proceedings* 97(2):316–321.

Bueno de Mesquita, Bruce and Alastair Smith. 2009. "A Political Economy of Aid." *International Organization* 63(2):309–340.

Bulman, David, Walter Kolkma and Aart Kraay. 2017. "Good Countries or Good Projects? Macro and Micro Correlates of World Bank and Asian Development Bank Project Performance." *Review of International Organizations* 12:335–363.

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

Centeno, Miguel Angel, Atul Kohli and Deborah J. Yashar. 2017. Unpacking States in the Developing World: Capacity, Performance, and Politics. In *States in the Developing World*, ed. Miguel Angel Centeno, Atul Kohli and Deborah Yashar. Cambridge: Cambridge University Press chapter 1, pp. 1–31.

Cernea, Michael M. 2008. Compensation and Investment in Resettlement: Theory, Practice, Pitfalls, and Needed Policy Reform. In *Can Compensation Prevent Impoverishment? Reforming Resettlement Through Investments and Benefit-Sharing*, ed. Michael M. Cernea and Hari Mohan Mathur. Oxford: Oxford University Press chapter 1, pp. 15–82.

Chayes, Abram and Antonia Handler Chayes. 1993. "On Compliance." *International Organization* 47(2):175–205.

Cinelli, Carlos and Chad Hazlett. 2020. "Making Sense of Sensitivity: Extending Omitted Variable Bias." *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 82(1):39–67.

Coase, Ronald H. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3(1):1–44.

Crost, Benjamin, Joseph H. Felter and Patrick B. Johnston. 2014. "Aid Under Fire: Development Projects and Civil Conflict." *American Economic Review* 104(6):1833–1856.

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(2013):288–302.

Denly, Michael. 2021. Institutional Autonomy and Donor Strategic Interest in Multilateral Foreign Aid: Rules vs. Influence. In *Political Economy of International Organizations*.

Djankov, Simeon, Jose G. Montalvo and Marta Reynal-Querol. 2008. "The Curse of Aid." *Journal of Economic Growth* 13(3):169–194.

Dreher, Axel and James Raymond Vreeland. 2014. *The Political Economy of the United Nations Security Council: Money and Influence*. New York: Cambridge University Press.

Dreher, Axel, Jan-Egbert Sturm and James Raymond Vreeland. 2009. "Development Aid

and International Politics: Does Membership on the UN Security Council Influence World Bank Decisions?" *Journal of Development Economics* 88(1):1–18.

Dube, Oeindrila and Suresh Naidu. 2015. "Bases, Bullets, and Ballots: The Effect of U. S. Military Aid on Political Conflict in Colombia." *Journal of Politics* 77(1):249–267.

Easterly, William. 2006. *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good*. New York: Penguin Books.

Easterly, William. 2015. *The Tyranny of Experts: Economists, Dictators, and the Forgotten Rights of the Poor*. New York: Basic Books.

Easterly, William and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112(4):1203–1250.

Evans, Peter and James E. Rauch. 1999. "Bureaucracy and Growth: A Cross-National Analysis of the Effects of "Weberian" State Structures on Economic Growth." *American Sociological Review* 64(5):748–765.

Falleti, Tulia G. 2021. "Invisible to Political Science: Indigenous Politics in a World in Flux." *Journal of Politics* 83(1):e5–e12.

Fariss, Christopher J., Michael R. Kenwick and Kevin Reuning. 2020. Measurement Models. In *SAGE Handbook of Research Methods in Political Science and International Relations*, ed. Luigi Curini and Robert J. Franzese. London: Sage Publications chapter 20, pp. 353–370.

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(01):75–90.

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

Fox, Jonathan A. 2002. The World Bank Inspection Panel and the Limits of Accountability. In *Reinventing the World Bank*, ed. Jeffrey A Winters. Ithaca and London: Cornell University Press pp. 131–163.

Gerring, John. 2017. *Case Study Research: Principles and Practices*. Second ed. Cambridge: Cambridge University Press.

Girod, Desha M. and Jennifer L. Tobin. 2016. "Take the Money and Run: The Determinants of Compliance with Aid Agreements." *International Organization* 70(1):209–239.

Greenstein, Gus. 2022. "The Influence of Alternative Development Finance on the World Bank's Safeguards Regime." *Global Environmental Politics* 22(3):171–193.

Hanson, Jonathan K. and Rachel Sigman. 2016. State Capacity and World Bank Project Success. In *Midwest Political Science Association Annual Meeting*. Chicago: .

Hanson, Jonathan K. and Rachel Sigman. 2021. "Leviathan's Latent Dimensions: Measuring State Capacity for Comparative Political Research." *Journal of Politics* 83(4):1495–1510.

Hart, Oliver. 2017. "Incomplete Contracts and Control." *American Economic Review* 107(7):1731–1752.

Hawkins, Darren G., David A. Lake, Daniel L. Nielson and Michael J. Tierney, eds. 2006. *Delegation and Agency in International Organizations*. Cambridge: Cambridge University Press.

Hendrix, Cullen S. 2010. "Measuring State Capacity: Theoretical and Empirical Implications for the Study of Civil Conflict." *Journal of Peace Research* 47(3):273–285.

Huffington Post. 2015. "World Bank Safeguards.".
URL: *http://www.huffingtonpost.com/news/world-bank-safeguards/*

Independent Evaluation Group. 2010. Safeguards and Sustainability Policies in a Changing World: An Independent Evaluation of the World Bank Group Experience. Technical report Washington, DC: .

Inter-American Development Bank. 2011. Policies for the Procurement of Goods and Works Financed by the Inter-American Development Bank. Technical report Inter-American Development Bank Washington: .

International Consortium of Investigative Journalists. 2015. "Evicted and Abandoned: The World Bank's Broken Promise to the Poor.".
URL: *http://www.icij.org/project/world-bank*

Kaja, Ashwin and Eric Werker. 2010. "Corporate Governance at the World Bank and the Dilemma of Global Governance." *World Bank Economic Review* 24(2):171–198.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi. 2011. "The Worldwide Governance Indicators: Methodology and Analytical Issues." *Hague Journal on the Rule of Law* 3(2):220–246.

Kersting, Erasmus and Christopher Kilby. 2016. "With a Little Help from My Friends: Global Electioneering and World Bank Lending." *Journal of Development Economics* 121:153–165.

Kilby, Christopher and Katharina Michaelowa. 2019. What Influences World Bank Project Evaluations? In *Lessons on Foreign Aid and Economic Development: Micro and Macro Perspectives*, ed. Nabamita Dutta and Claudia R. Williamson. Cham, Switzerland: Palgrave Macmillan chapter 6, pp. 109–150.

Kono, Daniel Y. and Gabriella R. Montinola. 2009. "Does Foreign Aid Support Autocrats, Democrats, or Both?" *Journal of Politics* 71(2):704–718.

Krugman, Paul and Maurice Obstfeld. 2003. *International Economics: Theory and Policy*. Seventh ed. Boston: Pearson Education.

Lee, Melissa M. and Melina R. Platas. 2015. "Aid Externalities: Evidence from PEPFAR in Africa." *World Development* 67:281–294.

Lee, Melissa M. and Nan Zhang. 2017. "Legibility and the Informational Foundations of State Capacity." *Journal of Politics* 79(1):118–132.

Lindberg, Staffan I., Michael Coppedge, John Gerring and Jan Teorell. 2014. "V-Dem: A New Way to Measure Democracy." *Journal of Democracy* 25(3):159–169.

Mann, Michael. 1984. "The Autonomous Power of the State: Its Origins, Mechanisms and Results." *Archives européennes de sociologie* 25(1984):185–213.

Martens, Bertin, Uwe Mummert, Peter Murrell and Paul Seabright. 2002. *The Institutional Economics of Foreign Aid*. Cambridge: Cambridge University Press.

Miguel, Edward. 2004. "Tribe or Nation? Nation Building and Public Goods in Kenya versus Tanzania." *World Politics* 56(3):327–362.

Montgomery, Jacob M., Brendan Nyhan and Michelle Torres. 2018. "How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do about It." *American Journal of Political Science* 62(3):760–775.

Morrison, Kevin M. 2012. "What Can We Learn about the Resource Curse from Foreign Aid?" *World Bank Research Observer* 27(1):52–73.

Moyo, Dambisa. 2009. *Dead Aid: Why Aid is Not Working and How There Is a Better Way*. New York: Farrar, Straus and Giroux.

Negi, Nalin Singh and Sujata Ganguly. 2011. "Development Projects vs. Internally Displaced Populations in India: A Literature Based Appraisal.".

Nielsen, Richard A., Michael G. Findley, Zachary S. Davis, Tara Candland and Daniel L. Nielson. 2011. "Foreign Aid Shocks as a Cause of Violent Armed Conflict." *American Journal of Political Science* 55(2):219–232.

Nielson, Daniel L. and Michael J. Tierney. 2003. "Delegation to International Organizations: Agency Theory and World Bank Environmental Reform." *International Organization* 57(02):241–276.

Nielson, Daniel L. and Michael J. Tierney. 2005. "Theory, Data, and Hypothesis Testing: World Bank Environmental Reform Redux." *International Organization* 59(3):785–800.

North, Douglass C. 1981. *Structure and Change in Economic History*. New York: W.W. Norton and Company.

Nunn, Nathan and Nancy Qian. 2014. "US Food Aid and Civil Conflict." *American Economic Review* 104(6):1630–1666.

OECD. 2008. "Paris Declaration of Aid Effectiveness and Accra Agenda for Action.".
URL: *http://www.oecd.org/dac/effectiveness/34428351.pdf*

OECD. 2017. "Purpose Codes: Sector Classification.".
URL: *http://www.oecd.org/dac/stats/purposecodessectorclassification.htm*

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, Massachusetts: Harvard University Press.

O'Reilly, Colin and Ryan M. Murphy. 2022. "An Index Measuring State Capacity: 1789-2018." *Economica* 89(355):713–745.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.

Ostrom, Elinor, James Walker and Roy Gardner. 1992. "Covenants With and Without a Sword: Self-Governance is Possible." *American Political Science Review* 86(2):404–417.

Pollack, Mark A. 1997. "Delegation, Agency, and Agenda Setting in the European Community." *International Organization* 51(1):99–134.

Qian, Nancy. 2015. "Making Progress on Foreign Aid." *Annual Review of Economics* 7(1):277–308.

Rajan, Raghuram G. and Arvind Subramanian. 2011. "Aid, Dutch Disease, and Manufacturing Growth." *Journal of Development Economics* 94(1):106–118.

Rose-Ackerman, Susan and Paul D. Carrington. 2013. *Anti-Corruption Policy: Can International Actors Play a Constructive Role?* Durham, N.C.: Carolina Academic Press.

Simmons, Beth A. 2010. "Treaty Compliance and Violation." *Annual Review of Political Science* 13(1):273–296.

Smets, Lodewijk, Stephen Knack and Nadia Molenaers. 2013. "Political Ideology, Quality at Entry and the Success of Economic Reform Programs." *Review of International Organizations* 8(4):447–476.

Soifer, Hillel and Matthias vom Hau. 2008. "Unpacking the Strength of the State: The Utility of State Infrastructural Power." *Studies in Comparative International Development* 43(3-4):219–230.

Tello, Rodolfo. 2015. *Social Safeguards: Avoiding the Unintended Impacts of Development.* Arlington, VA: Amakella Publishing.

Tversky, Amos and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185(4157):1124–1131.

UNU-WIDER. 2021. "Government Revenue Dataset.".

Vreeland, James Raymond. 2008. "The Effect of Political Regime on Civil War: Unpacking Anocracy." *Journal of Conflict Resolution* 52(3):401–425.

Vreeland, James Raymond. 2011. "Foreign Aid and Global Governance: Buying Bretton Woods - the Swiss-bloc Case." *Review of International Organizations* 6(3):369–391.

Weaver, Catherine E. 2008. *Hypocrisy Trap: The World Bank and the Poverty of Reform*. Princeton, New Jersey: Princeton University Press.

Weber, Max. 1978. *Economy and Society: An Outline of Interpretive Sociology*. Berkeley, California: University of California Press.

Williamson, Oliver. 1985. *The Economic Institutions of Capitalism*. New York: Simon and Schuster.

Winters, Matthew S. 2010. "Choosing to Target: What Types of Countries Get Different Types of World Bank Projects." *World Politics* 62(3):422–458.

Winters, Matthew S. and Shyam Kulkarni. 2014. The World Bank in the Post-Structural Adjustment Era. In *Handbook of Global Economic Governance: Players, Power and Paradigms*, ed. Manuela Moschella and Catherine Weaver. New York: Routledge chapter 17, pp. 251–266.

Wood, Reed M. and Christopher M. Sullivan. 2015. "Doing Harm by Doing Good? The Negative Externalities of Humanitarian Aid Provision during Civil Conflict." *Journal of Politics* 77(3):736–748.

World Bank. 2012. "International Bank for Reconstruction and Development Articles of Agreement (amended as effective June 27, 2012).".

World Bank. 2017. "World Development Indicators.".
URL: *http://data.worldbank.org/data-catalog/world-development-indicators*

Zvobgo, Kelebogile and Benjamin A.T. Graham. 2020. "The World Bank as an Enforcer of Human Rights." *Journal of Human Rights* 19(4):425–448.