

# State Capacity and World Bank Project Success

Jonathan K. Hanson  
Gerald R. Ford School of Public Policy  
University of Michigan

Rachel Sigman  
Naval Postgraduate School

Version: August 2019

## **Abstract**

In what types of institutional environments are international development projects most likely to succeed? Previous research suggests alternatively that liberal economic policies, democratic political institutions, and institutions characterized by “good governance” are all of primary importance. A serious drawback of this work is that measures of institutional quality tend to have poor coverage and conceptual fuzziness. This paper uses a new measure of state capacity with continuous coverage from 1960-2015 to evaluate the importance of different types of institutions on the success of over 10,000 World Bank projects. The results suggest that development projects are most likely to succeed where state capacity is relatively high, regardless of regime type and across different types of loans. Successful World Bank projects can, in turn, have a positive impact on State Capacity suggesting the possibility of a virtuous circle of development projects and state capacity building.

For some time now, scholars and practitioners of development have operated under the assumption that donor-funded development interventions are most likely to succeed in institutional contexts characterized by sound macroeconomic policies, democracy and “good governance.” It is in these environments where leaders have the political incentives to promote growth and provide goods and services to their citizens, and where citizens or groups can effectively demand and monitor the governments’ use of donor funds. Such findings have shaped the deployment of international development funds in at least two important ways. First, development projects are coupled with efforts to reform policy, promote democracy and improve governance. Second, as in the U.S.’s Millennium Challenge Account, donors are increasingly making development funding conditional on a country’s demonstrated commitment to democracy and good governance.<sup>1</sup>

While past research has provided a general sense of the institutional environments in which development projects may succeed or fail, competing accounts have produced some confusion about how, when and why these various types of institutions matter. In particular, it has been difficult for researchers to isolate the effects of these different types of institutions, all of which are likely to share some underlying latent characteristics. The use of aggregate indices that measure an array of facets related to institutional quality, often with limited temporal or geographic coverage, further complicates efforts to produce finer-grained knowledge about the ways that institutions affect donor investment in development projects. Moreover, while host-country capacity is often cited as an important issue for project success, the ability to measure its relationship to project outcomes has been limited by a dearth of reliable and conceptually valid measures.

In light of these issues, this paper re-examines the relationship between institutions and development project success. Using a new measure of state capacity with continuous coverage from 1960-2015, we are able to assess the relationship between state capacity and evaluations

---

<sup>1</sup>The World Bank’s IDA program uses the Country Policy Institutional Assessment (CPIA) for similar purposes and conditions tied to IMF loans have been in place for some time.

of more than 10,000 World Bank projects. We find that this new and improved measure of state capacity is a fairly consistent predictor of project success, regardless of other factors such as regime type, human capital of the borrowing country, and economic conditions.

Additionally, we examine variation in the impacts of these different institutions across policy sectors and loan instrument types. Given that program success is likely to vary depending on the nature of the intervention (Hirschman, 2011), and that host-country capacities inevitably vary across policy sectors (Whitfield et al., 2015), we seek to add further nuance to existing knowledge by analyzing the types of projects for which state capacity is most critical. In this analysis we find that state capacity is particularly important for the success of projects in the areas of human capital, the public sector, and infrastructure. State capacity does not have any clear relationship with the success of private sector projects. These results may have implications for donor decisions about what types of investments to make in high- and low-capacity states.

Finally, we examine the extent to which World Bank projects, many of which contain capacity-building components, have an effect on state capacity levels. The preliminary results suggest that World Bank projects can indeed have a positive impact on state capacity, particularly when projects have achieved higher outcome ratings. The presence of World Bank projects alone, however, is not sufficient to build capacity.

## **1 Institutional Determinants of Project Success**

Although there is widespread agreement that institutional environments are likely to shape the outcomes of international development interventions, debate persists about the types of institutions that matter most and why. At various times and in various contexts, studies have emphasized the importance of the macroeconomic policy environment, democracy and institutions promoting “good governance” as integral to development project success.

Understanding the institutional environments in which development projects are likely to succeed is critical for donors seeking higher returns on their investments and for minimizing risk of lending to countries that are unlikely to support donor-led projects. Indeed, as Burnside and Dollar (2004) find, the distribution of aid to countries conditional upon the quality of institutions and policies is a sensible response to the empirical reality that aid has a more positive impact on growth in these contexts.

One strand of research has emphasized the importance of macroeconomic policy environments. In particular, Isham and Kaufmann (1999) found that market-regarding policies such as undistorted prices, unmanipulated foreign exchange rates, free trade, and controlled fiscal deficits would significantly improve the payoffs from investing in new projects. Investments in policy environments that did not possess these attributes, on the other hand, would see considerably lower economic rates of return (1999: 177). In their widely cited study of the effects of foreign aid on economic growth Burnside and Dollar (1997, 2004) similarly find that in the presence of sound fiscal, monetary and trade policies, development aid tends to have a positive effect on economic growth, lending further support to the the notion that the macroeconomic policy environment is key to development project success. Subsequent challenges to this finding, advanced most notably by Boone (1996) and Easterly et al. (2003), have not only raised questions about the importance of the economic policy environment but have also questioned the potential effects of aid on economic growth more generally.

Other research focuses on the quality of institutions that underlie a strong macroeconomic policy environment such as the rule of law, property rights and contract enforcement. While Burnside and Dollar (1998) maintain the importance of macroeconomic policies, they expand their focus to economic management more broadly, combining insights about macroeconomic policy incentives with measures of institutional quality such as property rights, the absence of corruption and the quality of the bureaucracy. They find that the interaction of economic management and foreign aid has a meaningful effect on poverty reduction measures, namely

infant mortality. Concluding that countries that effectively put development policies into place are more likely to make productive use of aid monies (Burnside and Dollar, 1998: 14), the study implies the broader importance of the institutional environment, though without any attention to the specific roles or functions of these institutions.

The concept of economic management advanced by Burnside and Dollar (1998) has thus evolved to a broader focus on “governance” and project success. Factors such as corruption, rule of law, transparency, accountability and participatory decision-making are used to explain why projects tend to succeed in some countries but encounter many difficulties in others. The “governance” perspective has gained considerable popularity amongst the development community in shaping both how programs are designed and decisions about where and how to invest. Research has emphasized the importance of “good governance” for the success of World Bank development projects in particular (Ika et al., 2012; Burnside and Dollar, 2004; Dollar and Levin, 2005), as well as in the effectiveness of foreign aid more generally (Brautigam and Knack, 2004). Focusing more specifically on project outcomes, Dollar and Levin (2006) also finds an important role for these types of institutions in predicting project success.

A third strand of research focuses on the importance of democracy. In general, this research has come to the conclusion that interventions may be more successful in environments characterized by democratic political regimes, such as where citizens enjoy basic civil and political liberties including freedom of speech, assembly, and press freedom. In such environments, leaders have the incentives to pursue development projects that are desired by their populations and citizens have the ability to pressure and monitor their governments to implement projects effectively. For example, research by Isham et al. (1997) and Dollar and Levin (2006) have suggested a connection between higher rates of economic rates of return on development projects where citizen voice produces more effective government action. Although there is also evidence that no significant difference exists in aid effectiveness between

democratic and autocratic regimes Boone (1996), democracy has become an important precondition for some lending programs, such as the Millennium Challenge Account (MCA). The MCA's cancellation of their aid program in Tanzania in 2016 due to alleged electoral malfeasance, for instance, is testament to the strength of this commitment.

Whether the focus is on the macroeconomic policy environment, governance or democracy, the underlying logic is that these institutions constrain actors in recipient countries, namely government officials, in their ability to act in ways that undermine the goals of development projects. In this sense, these three types of institutions imply conditions that either incentivize leaders to pursue development, broadly defined, or constrain leaders' abilities to interfere with donor efforts to promote development. As one recent World Bank study states, "in many instances the underlying political drivers are too strong for technical constraints on politics to have their intended effects" (Fritz and Levy, 2014: 2). But to operate effectively in the development enterprise, these institutions must also possess *the capacity* to support and sustain development interventions by, for example, identifying areas of greatest need, developing the economic plans necessary to capitalize on the investments of foreign funds or support project goals over a sustained period of time. In particular, we would expect that countries possessing strong levels of control over their populations, effective planning processes and relatively capable bureaucracies to be able to participate in internationally-funded projects in a way that is likely to produce more successful project outcomes. These expectations arise from the widespread recognition that capable state institutions facilitate economic development (Johnson, 1982; Evans, 1995; Rauch and Evans, 2000; Kohli, 2004; World Bank Group, 1997). As such, the capacity of state institutions is likely to exert important influence over the extent to which governments become active and effective partners in international development interventions. Where state capacity is higher, we would expect to see better project outcomes, perhaps irrespective of other institutional variables.

Our contention is that the existing research has not adequately disaggregated the effects

of these different types of institutions on project success or failure. Though recent research has made great strides in producing knowledge about the ways that donor- or project-level variables affect outcomes, i.e. (Moll et al., 2015), attention to institutional environments has been somewhat less precise. In particular, studies of project success tend to use aggregate measures of the institutions such as the average of the World Bank’s Country Policy Institutional Assessment (CPIA) scores (Moll et al., 2015; Geli et al., 2014; Denizer et al., 2013; Bulman et al., 2015) or an index of the International Country Risk Guide’s institutional measures (Dollar and Levin, 2006), thereby clouding the ability to understand the specific types of institutions that matter most.<sup>2</sup> Additionally, these studies tend to be hampered by poor geographical or temporal coverage of available measures of institutional quality. The CPIA ratings become available only in 2005, and the International Country Risk Guide’s (ICRG) data, though available since 1982, cover only a limited set of countries.<sup>3</sup> Thus, our ability to gain analytic leverage to understand the relative importance of different types of institutions is severely hampered by limiting our inquiries to an abbreviated set of countries or years.

Additionally, both definitions and measures of governance or institutional quality often comprise a mix of concepts. It is not uncommon for the term “governance” to have multiple meanings such as corruption, rule of law, “good” policies, transparency, accountability, efficiency, and participation. As Grindle (2010) suggests, the “inflation” of the concept of good governance tends to cloud our understanding of the precise factors and processes that generate more productive development interventions. Most troubling, she notes, is that the concept of good governance tends to be conflated with the state’s *capacity* to grow the economy or the existence of democracy, resulting in the oversimplification of very complex relationships. One aim of our research is to effectively distinguish state capacity, democracy, and “good governance.” Thus, unlike past scholarship related to good governance, we avoid

---

<sup>2</sup>This problem is one that we do not address here.

<sup>3</sup>The ICRG have longer temporal coverage, beginning in 1982, but do not cover all countries.

definitions or measures of state capacity that relate to the political organization of decision making processes or the content of the decisions themselves.

Finally, efforts to understand the interaction of country-level and project-level variables have suggested that up to 80% of variation in project outcomes is a result of project-level factors such as the design, management or funding level of the project (Denizer et al., 2013; Bulman et al., 2015). But these studies have not fully taken into account the well-established idea that different projects have different levels of difficulty and will thus require different recipient-country *capacities*. Thus, one further element of our study is to examine how different types of institutions affect different types of projects. In particular, we draw on theories advanced by Hirschman (2011) about the importance of task specificity and the visibility of project results to achieve a finer-grained understanding of how the institutional environment, and specifically the capacity of state institutions, matters for predicting project success or failure.

In order to provide a stronger empirical basis for understanding the relationship between basic state capabilities and the success of development projects, we employ a new measure of state capacity with continuous coverage from 1960-2015 (Hanson and Sigman, March 21, 2013). To the extent possible, the indicators used to produce this Capacity measure are based on the definition of state capacity as the ability to achieve official goals. Though this conceptualization of state capacity may relate to governance concepts such as corruption, transparency and accountability, we see this notion of capacity as distinct from common conceptions of governance or institutional quality.

## **2 A New Measure of State Capacity**

In Hanson and Sigman (2013), we adopt a latent variable analysis of the kind that has been previously employed to assess measures of democracy (Treier and Jackman, 2008) and

governance (Arel-Bundock and Mebane, 2011; Bersch and Botero, 2011). The technique allows us to use multiple measurements of the same underlying concepts, even if noisy, to gain information about the distribution of the latent parameters that generate the observed indicators. With a recent update to the estimates, we now have a measure of state capacity with continuous coverage from 1960 through 2015, thereby enabling us to analyze the effects of state capacity on project outcomes over a much longer period than previously possible.

In selecting the indicators to use in the latent variable analysis, our goal was to capture the essence of state capacity while preserving analytical distinction from common variables of interest such as economic development, regime type, civil liberties, and good governance. We identified three basic capacity types that render the state's executive institutions more effective: coercive capacity, extractive capacity, and administrative capacity.<sup>4</sup> Additionally, we sought out indicators with broad coverage across countries and over time. Altogether, we employ 25 different indicators related to the three dimensions of state capacity. The data span 56 years (1960-2015) and up to 163 countries in a given year.<sup>5</sup> The indicators employed in this analysis are listed in Table 1.

The combination of these indicators allows us to make estimates over a broader span of countries and years than any one of these indicators individually. Wherever possible, we tried to avoid indicators prone to conflation with other variables of interest such as development, institutional quality, or regime type. Finally, we included a number of measures that we expect to relate principally to each of the three theorized dimensions of state capacity listed above to ensure that we capture multiple dimensions of the concept.

To produce the measure, we employ a latent variables estimation approach developed by Arel-Bundock and Mebane (2011) that uses Bayesian Markov-Chain Monte Carlo (MCMC)

---

<sup>4</sup>For a detailed explanation of why we chose these capacity types, see Hanson and Sigman (March 21, 2013).

<sup>5</sup>The number of countries is different in each year based on the sample definition using Polity IV data on the existence of countries. In some cases data for additional countries or years are available but not included because of the way we have defined the sample.

Table 1: Indicators of State Capacity

Variable	Countries	Years
Administrative Efficiency (Adelman and Morris, 1967)	69	1960-1962
Anocracy (calculated from Polity IV)	175	1960-2015
Bureaucratic Quality (Political Risk Services)	148	1982-2015
Census Frequency (calculated from UN 2011)	179	1960-2015
Efficiency of Revenue Mobilization (World Bank CPIA)	74	2005-2015
Information Capacity (Brambor et al., 2019)	166	1960-2015
Law and Order (Political Risk Services)	174	1984-2015
Military Personnel per 1,000 in population (COW)	171	1960-2015
(Log) Military Spending per 1,000 in population (COW)	168	1960-2015
Monopoly on Use of Force (Bertlesmann Transformation Index)	127	2003-2015
(Log) Paramilitary Personnel per 1,000 in population	164	1961-2015
Police Officers per 1000 in population (UN)	122	1973-2015
Quality of Budgetary & Financial Management (World Bank CPIA)	74	2005-2015
Quality of Public Administration (World Bank CPIA)	74	2005-2015
Rigorous and Impartial Public Administration (V-Dem v9)	177	1960-2015
State Antiquity Index, based on Bockstette et al. (2002)	162	1960-2015
State Authority over Territory (V-Dem v9)	174	1960-2015
Statistical Capacity (World Bank)	134	2004-2015
Taxes on Income as % of Revenue (IMF, WDI)	152	1970-2015
Taxes on International Trade as % Revenue (IMF, WDI)	155	1970-2015
Total Tax Revenue as % GDP (IMF, WDI, OECD)	152	1960-2015
Weberianess (Rauch and Evans, 2000)	34	1970-1990

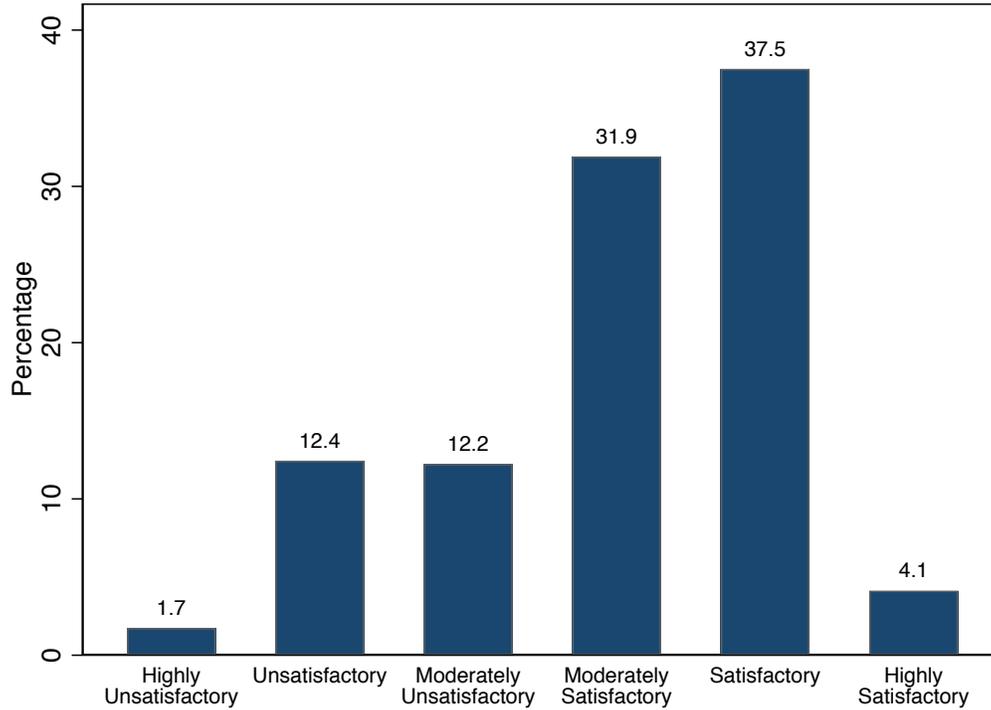
techniques to identify underlying factors. The various observed indicators are a linear function of the latent values of state capacity in each dimension measured with some error. Since there are many observed indicators measured in many countries over several years, we have multiple data points with which to obtain the posterior distributions of the latent parameters. For the purposes of the present analysis, we conduct the latent variable analysis to produce a 1-dimensional estimate of state capacity.

The resulting measure, labeled Capacity, appears to be a general-purpose measure of state capacity that draws from indicators representing all three theorized dimensions. Several of the indicators most strongly associated with Capacity are the World Bank’s measure of statistical capacity ( $r=.82$ ), Political Risk Services’ measures of bureaucratic quality ( $r=.80$ ) and law and order ( $r=.80$ ) and (log) military spending as a percentage of the size of the population ( $r=.78$ ). In a range of validity checks in Hanson and Sigman (March 21, 2013), we show that the Capacity measure performs as expected in predicting the state’s ability to perform core functions such as tracking its population and delivering a range goods and services.

### **3 Institutions and Project Success**

This section tests the influence of various institutions on the outcomes of World Bank development projects. We use the IEG World Bank Project Performance Ratings dataset (World Bank Group, 2019), which includes ratings for over 12,400 projects evaluated between 1973 and 2018. The IEG is an independent unit within the World Bank that assesses its activities. For projects rated prior to 1995, the dataset codes the outcome as either “satisfactory” or “unsatisfactory.” Starting in 1995, each lending project in a participating country is rated on a six-point scale that ranges from “highly unsatisfactory” to “highly satisfactory.” For some analyses, we use this six-point scale and thus limit the sample to projects rated 1995

Figure 1: Frequency Distribution of IEG Ratings post-1994



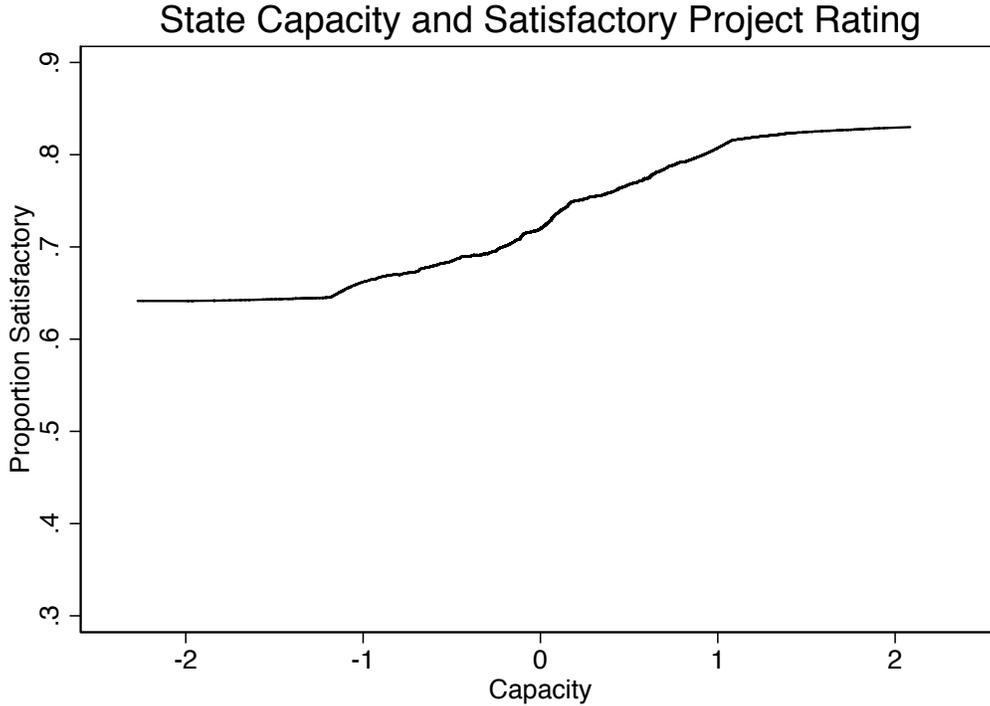
and later. The frequency distribution of these project ratings is provided in Figure 1. For analyses using the full dataset of project ratings, we convert the six-point scale to a dichotomous “satisfactory” vs. “unsatisfactory” format.<sup>6</sup> Overall, about 73% of projects received a rating of “satisfactory” upon recoding.

We employ the new Capacity measures as a predictor of project success. We find that higher levels of state capacity, as represented by Capacity, are associated with a significantly higher probability that a project will receive a rating of satisfactory. Figure 2 fits a Lowess curve plotting the proportion of projects that achieve a rating of “satisfactory” on the mean level of Capacity in a country during the period of project implementation. As is evident from the figure, the proportion of projects receiving satisfactory ratings rises steadily as the Capacity score moves from -1 to 1.

---

<sup>6</sup>The ratings of “moderately satisfactory” and “highly satisfactory” are combined with those coded as “satisfactory.” Ratings of “highly unsatisfactory” and “moderately unsatisfactory” are combined with those coded as “unsatisfactory.”

Figure 2: IEG Project Success Rating Across Values of Capacity 1973-2018



We next turn to various multivariate tests with Capacity as a predictor of project success. Each model includes a set of control variables measured at their mean levels during the time between the project’s approval and completion dates: log GDP per capita (*lnGDPcap*) from Feenstra et al. (2015),<sup>7</sup> the GDP per capita growth rate, the average number of years of education in the population aged 15 and up (*YearsEduc*) from Barro and Lee (2010) and Bosworth and Collins (2003), and exchange rate shocks during the project period. Measures of political regime characteristics include the *Polity2* index (Marshall and Jaggers, 2009), the measures of political rights (*PolRights*) and civil liberties (*CivLib*) from Freedom House (Freedom House, 2009), and a dichotomous measure of *Democracy* from Boix et al. (2012).<sup>8</sup> Finally, we control for the length of the project in years, since some projects are much longer-term than others.

<sup>7</sup>The natural log of *rgpdna/pop*.

<sup>8</sup>We have rescaled the Freedom House variables such that they run from 0 to 6 with higher values meaning greater political rights or civil liberties. In all other variables higher values represent higher quality institutions.

In Table 2, we show the results of models that predict the probability that a project receives a rating of satisfactory, using the full IEG dataset with evaluations measured dichotomously. The four models each contain a different measure of political regime characteristics. In all four models, Capacity is associated with a higher probability of a satisfactory rating, all with high levels of statistical confidence ( $p < .01$ ). Substantively, in all four models, each one unit increase in Capacity, which is one standard deviation of the measure, is associated with an increase of approximately .08 to .1 in the probability that a project will get a satisfactory rating.<sup>9</sup>

None of the variables capturing characteristics of political regimes show a substantively large or statistically significant relationship to the probability of project success. There is solid evidence that growth-promoting economic policies support project success, as does avoidance of exchange rate shocks. Neither the overall level of wealth, as measured by the log value of GDP per capita, nor the level of education in the population are significant predictors of project success. Consistent with other research, we find that longer projects are a bit less likely to get satisfactory ratings.

Figure 3 depicts the predicted probability that a project is rated as “satisfactory” using the results from Model 1 in Table 3. All the other independent variables are held constant at their median values. As the figure illustrates, the predicted probability of a satisfactory rating rises dramatically when levels of Capacity are higher. The total effect of this variable, when Capacity goes from -2 to 2, is for the probability of a satisfactory rating to rise from .54 to .87. The bars around each point reflect the 95% confidence interval of the estimate.

To check the robustness of these findings, we divide the sample into different subsets based upon the type of loan agreement reached with borrowers (IBRD or IDA) and the lending instrument type (investment loans or development policy loans). International Bank for Reconstruction and Development (IBRD) loan agreements are made at interest rates

---

<sup>9</sup>This is the average marginal effect across all cases in the dataset, given their values of the independent variables.

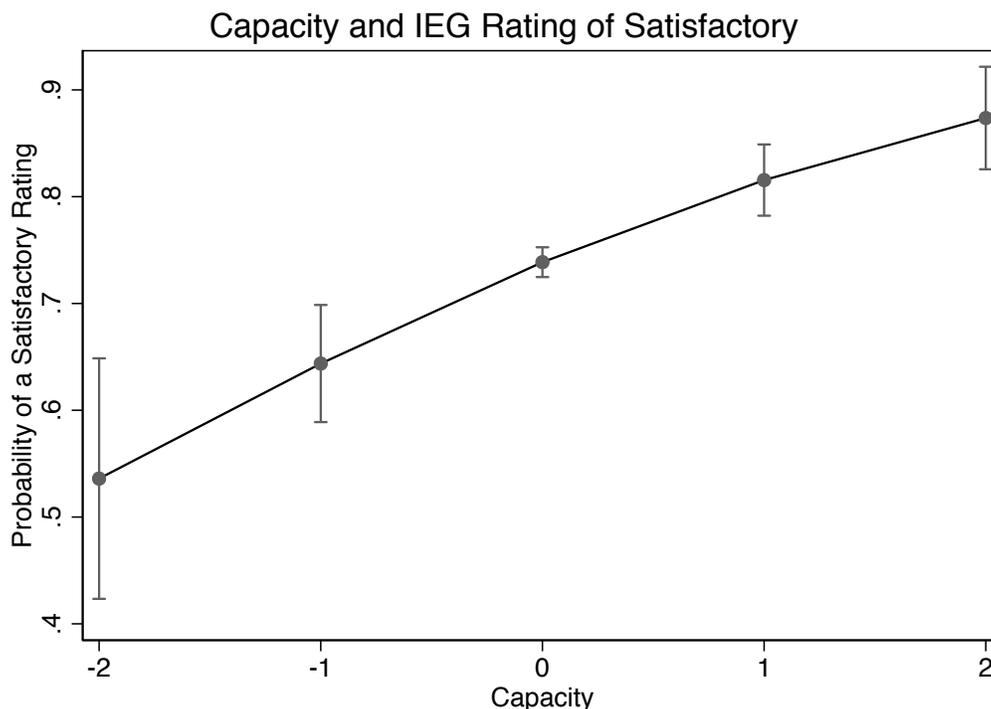
Table 2: Prediction of Satisfactory IEG Project Rating

	(1)	(2)	(3)	(4)
Capacity	0.45** (0.11)	0.46** (0.11)	0.46** (0.11)	0.45** (0.11)
ProjectLength	-0.05** (0.01)	-0.05** (0.01)	-0.05** (0.01)	-0.05** (0.01)
lnGDPcap	0.03 (0.08)	0.03 (0.08)	0.04 (0.08)	0.03 (0.08)
YearsEduc	-0.02 (0.02)	-0.03 (0.02)	-0.02 (0.02)	-0.03 (0.02)
Growth	0.10** (0.01)	0.10** (0.01)	0.10** (0.01)	0.10** (0.01)
ExRateShock	-0.30* (0.15)	-0.30^ (0.16)	-0.30^ (0.16)	-0.30* (0.15)
Polity2	-0.00 (0.01)			
PolRights		-0.03 (0.03)		
CivLibs			-0.05 (0.04)	
Democracy				-0.07 (0.09)
Constant	1.08^ (0.57)	1.13^ (0.58)	1.13* (0.57)	1.10^ (0.58)
N	10057	10045	10045	10058
Log-likelihood	-5576.59	-5570.79	-5570.23	-5576.70

Logistic Regression with standard errors clustered at the country level.

^  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$

Figure 3: Probability of Satisfactory Rating at Levels of Capacity



akin to commercial lending, with loans going predominantly to middle-income countries or low-income countries with good credit ratings. The International Development Association (IDA) provides credits and grants to the lowest-income countries as a form of international assistance. In general, we might expect that IDA projects are riskier propositions, and project ratings might be lower on average. Indeed, the success rate of IBRD agreements is about 76%, compared to 70% for IDA agreements. It is not clear, however, whether the effect of state capacity on these outcomes will matter more or less for one type of loan agreement or the other.

Moll et al. (2015) argue that the post-2014 shift from structural adjustment loans (investment loans) to development policy lending was a significant change in Bank operations. Loans were to be denied to countries that did not provide for basic macroeconomic management, and the purpose of the loans was to shift toward the achievement of development results. In theory, the requirement that countries should meet basic macroeconomic require-

ments could serve as a kind of selection effect that depresses the effect of state capacity on IEG success ratings. On average, loans might flow more readily to places with more effective institutions of economic management.

In the appendix (Table A1) we present the results using subsamples consisting of these four loan agreement or instrument types. The first and second models deal with IBRD and IDA agreements respectively. Capacity remains a positive and statistically significant predictor of a rating of “satisfactory” for both types, although the level of significance goes to .1 in the case of IDA agreements. The results suggest that Capacity has a stronger (i.e. more positive) effect for IBRD agreements than for IDA agreements (coefficients of .51 vs. .30), though we do not have sufficient statistical confidence to rule out the null hypothesis of no difference. Models 3 and 4 of Table A1 focus on the subsamples of investment loans (IPF) and development policy loans (DPF). In both cases, the effect of Capacity is positive and statistically significant. The coefficients are of similar magnitude. If anything, state capacity has a stronger positive effect for development policy loans than for investment loans.

Next, we turn to a second major set of tests involving World Bank projects that were rated on the six-point scale. The sample for these tests consists of project rated in the year 1995 or later, since that is the point at which the six-point rating scale was adopted. We thus estimate ordered logistic regressions. The results, as seen in Table 3, are generally consistent with those of the logit model in Table 2. The effect of Capacity is positive, meaning that it is associated with higher probability of receiving a rating at the upper end of the ordinal scale, and statistically significant at the .01 level. Each one-unit increase in Capacity is associated with an average increase in the probability of a “satisfactory” or “highly satisfactory” rating by about .11 to .12.

We again find that longer-term projects tend to have lower ratings, on average. Likewise, neither the level of GDP per capita nor average years of education in the population are significant predictors of project success. Faster economic growth is again statistically

Table 3: Prediction of Ordinal IEG Project Rating

	(1)	(2)	(3)	(4)
Capacity	0.50** (0.13)	0.51** (0.13)	0.56** (0.13)	0.49** (0.13)
ProjectLength	-0.05** (0.01)	-0.05** (0.01)	-0.05** (0.01)	-0.05** (0.01)
lnGDPcap	0.02 (0.08)	0.02 (0.09)	0.02 (0.08)	0.02 (0.09)
YearsEduc	-0.02 (0.02)	-0.02 (0.02)	-0.01 (0.02)	-0.02 (0.02)
Growth	0.08** (0.02)	0.08** (0.01)	0.07** (0.01)	0.08** (0.02)
ExRateShock	0.29 (0.18)	0.29 (0.18)	0.26 (0.18)	0.29 (0.18)
Polity2	-0.01 (0.01)			
PolRights		-0.03 (0.03)		
CivLibs			-0.10* (0.04)	
Democracy				-0.04 (0.10)
$\tau_1$	-4.15** (0.59)	-4.22** (0.62)	-4.42** (0.64)	-4.16** (0.60)
$\tau_2$	-1.75** (0.60)	-1.81** (0.63)	-2.02** (0.64)	-1.76** (0.60)
$\tau_3$	-0.98 (0.60)	-1.04 <sup>^</sup> (0.63)	-1.25 <sup>^</sup> (0.64)	-0.99 (0.61)
$\tau_4$	0.38 (0.60)	0.32 (0.63)	0.12 (0.64)	0.37 (0.60)
$\tau_5$	3.23** (0.61)	3.17** (0.64)	2.97** (0.64)	3.22** (0.61)
N	6395	6396	6396	6396
Log-likelihood	-9074.60	-9074.72	-9065.10	-9076.02

Ordered Logistic Regression with standard errors clustered at the country level.

<sup>^</sup>  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$

associated with a higher probability of project success, but exchange rate shocks are not found to be a significant factor in this set of estimates. In general, the variables measuring regime characteristics remain unimportant, but the coefficient on the Freedom House civil liberties measure, rescaled such that higher values mean greater civil liberties, is negative and statistically significant. Substantively, the finding is fairly small: a one-unit increase in the scale is associated with a reduction of .02 in the probability of a “satisfactory” or “highly satisfactory” project.

Thus far, most of the tests involve loans for projects in a wide variety of policy areas. Ignoring this heterogeneity is potentially problematic. As Denizer et al. (2011) find, project outcomes vary significantly within countries. One source of within-country variation may be systematic differences in project success across policy sectors. The IEG classifies projects into 19 different sectors. We combine some of these sectors to produce six broad categories of projects: agriculture, human capital (includes education, health, poverty reduction and social development projects), energy and mining, private sector development (includes financial sector), public sector development (includes public financial management), and infrastructure (includes IT/communications, transportation, and water). We exclude projects related to the environment, gender development, and social rights, because the number of projects conducted in these sectors were considerably smaller than other categories.

We continue with the six-point IEG outcome scale as the dependent variable, restricting the sample to each of the six categories of projects in sequence. The results are presented in Table 4. Divided by sector, the analysis shows that human capital projects, public sector project and infrastructure projects are the most responsive to (positively) to state capacity. This result is not surprising since large scale social interventions such as improvement in education and health systems often require complex coordination across territories and agencies, and significant involvement of actors from the apex of government all the way to street-level bureaucrats. In other sectors, such as agriculture and energy/mining, capacity

Table 4: Prediction of Satisfactory IEG Project Rating by Sector

	(1)	(2)	(3)	(4)	(5)	(6)
	Agriculture	Human Capital	Energy/Mining	Private Sector	Public Sector	Infrastructure
Capacity	0.35 <sup>^</sup> (0.21)	0.48* (0.19)	0.40 <sup>^</sup> (0.24)	-0.49 (0.70)	0.61** (0.24)	0.68* (0.28)
ProjectLength	-0.01 (0.03)	-0.12** (0.02)	-0.06 (0.04)	0.01 (0.18)	-0.06 <sup>^</sup> (0.03)	-0.13** (0.04)
lnGDPcap	0.07 (0.12)	0.16 (0.11)	-0.07 (0.18)	0.22 (0.82)	0.11 (0.15)	-0.25 (0.23)
YearsEduc	0.00 (0.04)	-0.08* (0.04)	0.05 (0.07)	0.24 (0.16)	-0.03 (0.06)	-0.02 (0.05)
Growth	0.13** (0.03)	0.05* (0.02)	0.10** (0.03)	0.14 (0.09)	0.08 (0.05)	0.07* (0.03)
ExRateShock	0.56** (0.14)	-0.00 (0.24)	-0.54* (0.25)	-0.14 (0.32)	0.15 (0.52)	0.21 (0.41)
Polity2	-0.00 (0.01)	-0.00 (0.01)	-0.02 (0.01)	0.04 (0.06)	0.02 (0.02)	-0.02 (0.02)
$\tau_1$	-3.08** (0.87)	-4.30** (0.82)	-4.43** (1.30)	-0.99 (5.74)	-3.34** (1.07)	-7.31** (1.53)
$\tau_2$	-0.71 (0.86)	-1.75* (0.82)	-1.80 (1.26)	2.28 (5.67)	-0.80 (1.06)	-4.98** (1.52)
$\tau_3$	-0.07 (0.86)	-0.82 (0.81)	-1.24 (1.27)	2.61 (5.66)	0.13 (1.13)	-4.15** (1.51)
$\tau_4$	1.27 (0.86)	0.56 (0.81)	-0.19 (1.25)	3.81 (5.72)	1.48 (1.13)	-2.68 <sup>^</sup> (1.49)
$\tau_5$	4.30** (0.94)	3.55** (0.83)	2.52* (1.22)	5.98 (5.66)	4.93** (1.19)	0.31 (1.36)
N	1117	1731	661	64	438	1011
Log-likelihood	-1571.39	-2406.14	-957.05	-91.47	-621.92	-1350.77

Logistic Regression with standard errors clustered at the country level.

<sup>^</sup>  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$

has a more modest positive effect, suggesting that state capacity is slightly less critical to project success in these areas. No significant effect of state capacity is found in private sector projects.

In the appendix, we present additional tests that show the importance of state capacity for project success. In Table A2, we use data from Honig (2018) containing project ratings from five other international development agencies: the Asian Development Bank; the UK’s Department for International Development (DFID); the Global Fund to Fight AIDS, Tuberculosis, and Malaria (GFATM); the Japan International Cooperation Agency (JICA), and the German development bank KfW. For projects involving three of the five agencies, Capacity is a positive and statistically significant predictor of project success. The two exceptions are DFID, for which Capacity has no predictive power, and JICA, for which Capacity is a positive but not significant predictor of project success.

A final set of robustness checks switches the dependent variable from overall project success to the IEG’s rating of the performance of the borrower government. We expect that Capacity should be a strong predictor of borrower government performance, and the estimates confirm these expectations. See Table A3.

Overall, across a variety of different tests, we find that Capacity is a fairly solid and consistent predictor for the success of development projects. We do not find that measures of democracy have any consistent effects. Of course, capacity-building is an important goal in many World Bank projects, so we are also very interested in whether these projects, over the long run, can raise the level of state capacity. We turn to that next.

## **4 Effects of World Bank Projects on State Capacity**

Seeing that state capacity has consistently positive effects on the success of World Bank development projects, we test to see whether or not successful projects can, in turn, lead to

greater levels of state capacity, thereby creating a virtuous circle in which projects promote capacity and vice versa. The integration of capacity-building elements into World Bank project design has been ongoing for quite some time. For example, a 2005 OED report estimated that the World Bank had spent approximately \$10 billion on capacity building throughout the public sector in sub-Saharan Africa alone.

To conduct this analysis, we transformed the project-level data into time-series-cross-sectional data by country-years. For each country-year observation, we recorded the number of projects that were ongoing in that year, the number of projects happening in that year that were ultimately rated as satisfactory, as well as the mean of the IEG outcome ratings of those projects. The number of projects happening in a given country-year is represented by the variable *Project*, the number of satisfactory projects is represented by the variable *Satisfactory*, and the mean IEG outcome rating is represented by the variable *Outcome*. The data were then collapsed into five-year groups from 1960-2015 with mean values across the five year time period.

In the first set of tests, displayed in Table 5, the dependent variable is the Capacity score in time period  $t$  and independent variables are lagged by one five-year period ( $t-1$ ). We use a random effects generalized least squares regression with robust standard errors to test the effects of Capacity alongside other institutional variables. For each model, we check for robustness of the findings by testing not only the average of the Outcome rating over that period, but also the number of Satisfactory projects that were ongoing or completed in a given country in the five year period. The results indicate that the execution of projects with higher ratings in the previous five-year period have a consistent, positive and statistically-significant effect on the Capacity level of that country in the subsequent period. Likewise, democratic characteristics of political regimes seem to have a consistently positive and statistically significant effect.

Next, we calculate the change in Capacity over each five-year period and over each ten-

Table 5: Effect of World Bank Projects on Level of State Capacity

	(1)	(2)	(3)	(4)	(5)	(6)
Project <sub>t-1</sub>	0.00* (0.00)	-0.00* (0.00)	0.00* (0.00)	-0.00* (0.00)	0.00* (0.00)	-0.00* (0.00)
Outcome <sub>t-1</sub>	0.05** (0.01)		0.05** (0.01)		0.05** (0.01)	
Satisfactory <sub>t-1</sub>		0.01* (0.00)		0.01** (0.00)		0.01** (0.00)
Polity2 <sub>t-1</sub>	0.00* (0.00)	0.00* (0.00)				
CivLib <sub>t-1</sub>			0.02** (0.01)	0.03** (0.01)		
PolRights <sub>t-1</sub>					0.02** (0.01)	0.02** (0.01)
lnGDPcap <sub>t-1</sub>	0.02^ (0.01)	0.02 (0.01)	0.02 (0.01)	0.02 (0.01)	0.02 (0.01)	0.02 (0.01)
YearsEduc <sub>t-1</sub>	0.01** (0.00)	0.01** (0.01)	0.01** (0.00)	0.02** (0.00)	0.01** (0.00)	0.02** (0.00)
Corruption <sub>t-1</sub>	0.01 (0.01)	0.02 (0.01)	0.01 (0.01)	0.02 (0.01)	0.01 (0.01)	0.01 (0.01)
Capacity <sub>t-1</sub>	0.85** (0.02)	0.85** (0.02)	0.85** (0.02)	0.85** (0.02)	0.85** (0.02)	0.85** (0.02)
Constant	-0.42** (0.11)	-0.23* (0.11)	-0.44** (0.11)	-0.25* (0.10)	-0.42** (0.11)	-0.23* (0.10)
N	427	427	427	427	427	427
Countries	87	87	87	87	87	87
R <sup>2</sup>	0.95	0.95	0.95	0.95	0.95	0.95

Random-effects Generalized Least Squares regression with robust standard errors.

^  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$

year period. These change in Capacity variables serves as the dependent variables for the next set of tests. Again, we use a random effects GLS regression with robust standard errors. Because the dependent variable is a difference, we do not lag the independent variables. Instead we test the effects on change in Capacity in the current period as a function of projects ongoing during that period, as well as the level of Capacity at the beginning of the priod. In Table 6, Models 1 and 3 represent the 5-year change, and models 2 and 4 represent the 10-year change.

Table 6: Number of World Bank Projects and Change in State Capacity (Polity)

	(1)	(2)	(3)	(4)
Project	0.00 (0.00)	0.00** (0.00)	-0.00* (0.00)	-0.00 (0.00)
Outcome	0.02** (0.01)	0.07** (0.02)		
Satisfactory			0.00* (0.00)	0.01* (0.00)
Polity2	0.00** (0.00)	0.01** (0.00)	0.00** (0.00)	0.01** (0.00)
Corruption	0.01^ (0.01)	0.03 (0.02)	0.02* (0.01)	0.03^ (0.02)
GDPcap	-0.00 (0.00)	0.00* (0.00)	-0.00 (0.00)	0.00* (0.00)
YearsEduc	0.00 (0.00)	0.02** (0.01)	0.00 (0.00)	0.02** (0.01)
CapStart	-0.04* (0.02)	-0.27** (0.04)	-0.04* (0.02)	-0.27** (0.04)
Constant	-0.10* (0.04)	-0.45** (0.11)	-0.02 (0.03)	-0.18* (0.07)
N	518	427	519	427
Countries	88	87	88	87
$R^2$	0.07	0.15	0.07	0.12

Random-effects Generalized Least Squares regression with robust standard errors.

^  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$

The results are remarkably consistent. The raw number of World Bank projects does not have a consistent effect on the growth of Capacity over each time period. Higher average IEG ratings, however, have a positive and statistically significant relationship to the change in State Capacity for both 5-year and 10-year variables. The effects are particularly strong when the 6-point ordinal IEG Outcome scale is used in the analysis (Models 1 and 2).

These tests are repeated in the Appendix with different measures of democracy See Tables A4 and A5. In all of these tests, countries in which IEG project ratings were higher tended to experience faster growth in Capacity. Similarly, the measures of democracy are positively associated with the growth of state capacity in a way that is consistent across the measures.

## 5 Discussion and Conclusion

In examining the relationship between state capacity, political institutions and the success of World Bank development projects, this paper provides evidence that there is considerably more to learn about the institutional environments in which development projects are deployed. Democratic political institutions such as electoral competition, civil liberties, and political rights do not necessarily have the theorized effects on development project success, warranting the need for further inquiry into the hypothesized mechanisms and incentives at work in these relationships. The results also indicate that there may be some systematic variation across project sectors, whereby pre-existing levels of state capacity are particularly important to the success or failure of investments in human capital such as health and education, as well as infrastructure projects.

Understanding how varied institutional environments shape development interventions is particularly important in light of the rise of randomized control trials (RCTs) to assess the effectiveness of development interventions. Since experiments across multiple countries (or other jurisdictions) is not always possible, advancements in theory about the relationship

between institutions and development project outcomes can help to provide insight about the external validity of RCT findings.

Given the strength of the evidence suggesting the importance of state capacity in producing favorable project outcomes, we also inquired into the effects of World Bank projects on state capacity. With some surprise, the findings were very consistent. When executed successfully, World Bank projects appear to facilitate positive changes in state capacity, as well as overall higher levels of state capacity. These findings hold up with numerous controls, including pre-existing levels of state capacity, economic wealth, and a variety of political institutional arrangements. Taken together, the two sets of analyses suggest that virtuous circles of development interventions and capacity building exist, but that project selection and successful execution are necessary. Future work should further explore these dynamics.

Finally, the findings also suggest that additional work can be done to sort out the relationships between capacity and governance, and to develop better measures that can consistently and reliably measure these concepts. To this end, the Bayesian latent variable methods used to develop a new measure of state capacity in this paper could be applied to other institutional concepts to provide even more nuanced knowledge about the ways that institutional environments shape development project outcomes.

## References

- Adelman, Irma and Cynthia Taft Morris. 1967. *Society, Politics, & Economic Development: A Quantitative Approach*. Baltimore: The Johns Hopkins Press.
- Arel-Bundock, Vincent and Walter Mebane. 2011. "Measurement Error, Missing Values and Latent Structure in Governance Indicators." Presented at the 2011 Annual Meeting of the American Political Science Association.
- Barro, Robert and Jong-Wha Lee. 2010. "A New Data Set of Educational Attainment in the World, 1950-2010." NBER Working Paper No. 15902.
- Berkman, Heather, Carlos G. Scartascini, Ernesto Stein, and Mariano Tommasi. 2008. "Political Institutions, State Capabilities, and Public Policy: An International Dataset." Washington, D.C., United States: IDB Research Department.
- Bersch, Katherine and Sandra Botero. 2011. "Measuring Governance: A Defense of Disaggregation." Prepared for the American Political Science Association Conference, September 1-4, 2011.
- Bockstette, Valerie, Areendam Chanda, and Louis Putterman. 2002. "States and markets: The advantage of an early start." *Journal of Economic Growth* 7(4): 347-369.
- Boix, Carles, Michael Miller, and Sebastian Rosato. 2012. "A Complete Data Set of Political Regimes, 1800-2007." *Comparative Political Studies* 46(12): 1523-54.
- Boone, Paul. 1996. "Politics and the Effectiveness of Foreign Aid." *European Economic Review* 40(2): 289-329.
- Bosworth, Barry P. and Susan M. Collins. 2003. "The Empirics of Growth: An Update." *Brookings Papers on Economic Activity* 2003(2): 113-179.
- Brambor, Thomas, Agustín Goenaga, Johannes Lindvall, and Jan Teorell. 2019. "The Lay of the Land: Information Capacity and the Modern State." *Comparative Political Studies* pp. 1-19.

- Brautigam, Deborah and Stephen Knack. 2004. "Foreign Aid, Institutions, and Governance in Sub-Saharan Africa." *Economic Development and Cultural Change* 52(2): 255–285.
- Bulman, David, Woalter Kolkma, and Aart Kraay. 2015. "Good Countries or Good Projects? Comparing Macro and Micro Correlates of World Bank and Asian Development Bank Project Performance." *World Bank Policy Research Working Paper No. 7245* .
- Burnside, Craig and David Dollar. 1997. "Aid, Policies and Growth." *The World Bank Policy Research Working Papers* 1777.
- . 1998. *Aid, the Incentive Regime, and Poverty Reduction*. Washington: The World Bank.
- . 2004. "Aid, Policies and Growth: Revisiting the Evidence." *World Bank Policy Research Paper* O-2834.
- Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Steven Wilson, Agnes Cornell, Lisa Gastaldi, Haakon Gjerløw, Nina Ilchenko, Joshua Krusell, Laura Maxwell, Valeriya Mechkova, Juraj Medzihorsky, Josefine Pernes, Johannes von Römer, Natalia Stepanova, Aksel Sundström, Eitan Tzelgov, Yi ting Wang, Tore Wig, and Daniel Ziblatt. 2019. "V-Dem [Country-Year/Country-Date] Dataset v9." Varieties of Democracy (V-Dem) Project.
- 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.
- Denizer, Cevdet, Daniel Kaufmann, and Art Kraay. 2011. "Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance." *World Bank Development Research Group Macroeconomics and Growth Team, Policy Research Working Paper 5646* .
- Dollar, David and Victoria Levin. 2005. "Sowing and Reaping: Institutional Quality and Project Outcomes in Developing Countries." World Bank Policy Research Working Paper 3524.

- . 2006. “The Increasing Selectivity of Foreign Aid, 1984-2003.” *World Development* 34(12): 2034–2046.
- Easterly, William, Ross Levine, and David Roodman. 2003. “New Data, New Doubts: A Comment on Burnside and Dollar’s “Aid Policies and Growth” (2000).” *National Bureau of Economic Research* 9846.
- Evans, Peter. 1995. *Embedded Autonomy: States and industrial transformation*. Princeton: Princeton University Press.
- Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. “The Next Generation of the Penn World Table.” *American Economic Review* 105(10): 3150–3182.
- Freedom House. 2009. *Freedom in the World Country Ratings: 1972 to 2008*. Washington, D.C.: Freedom House, Inc.
- Fritz, Verena and Bryan Levy. 2014. *Problem-Driven Political Economy Analysis: The World Bank’s Experience*. Washington DC: The World Bank.
- Geli, Patricia, Art Kraay, and Hoveida Nobakht. 2014. “Predicting World Bank Project Outcome Ratings.” *World Bank Policy Research Working Paper No. 7001* .
- Grindle, Merilee S. 2010. “Good Governance: The Inflation of an Idea.” *HKS Faculty Research Working Paper Series, John F. Kennedy School of Government, Harvard University*. RWP10-023.
- Hanson, Jonathan K. and Rachel Sigman. March 21, 2013. “Leviathan’s Latent Dimensions: Measuring State Capacity for Comparative Political Research.” *World Bank Political Economy Brown Bag Lunch Series* .
- Hirschman, Albert O. 2011. *Development Projects Observed*. Washington DC: The Brookings Institution.
- Honig, Dan. 2018. *Navigation By Judgment: Why and When Top Down Management of Foreign Aid Won’t Work*. New York: Oxford University Press.

- Ika, Lavagnon A., Amadou Diallo, and Dennis Thuillier. 2012. "Critical Success Factors for World Bank Projects: An Empirical Investigation." *International Journal of Project Management* 30: 105–116.
- Isham, Jonathan and Daniel Kaufmann. 1999. "The Forgotten Rationale for Policy Reform: The Productivity of Investment Projects." *The Quarterly Journal of Economics* 114(1): 149–184.
- Isham, Jonathan, Daniel Kaufmann, and Lant H. Pritchett. 1997. "Civil Liberties, Democracy, and the Performance of Government Projects." *The World Bank Economic Review* 11(2): 219–240.
- Johnson, Chalmers. 1982. *MITI and the Japanese Miracle*. Stanford: Stanford University Press.
- Kohli, Atul. 2004. *State-Directed Development: Political power and industrialization in the global periphery*. Cambridge: Cambridge University Press.
- Marshall, Monty G. and Keith Jagers. 2009. *Polity IV Project: Dataset Users' Manual*. Center for Systemic Peace, George Mason University.
- Moll, Peter, Patricia Geli, and Pablo Saavedra. 2015. "Correlates of Success in World Bank Development Policy Lending." World Bank Policy Research Working Paper 7181.
- Rauch, James and Peter Evans. 2000. "Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries." *Journal of Public Economics* 75(1): 49–71.
- Transparency International. 2009. "Corruption Perceptions Index." URL <http://www.transparency.org/>.
- Treier, Shawn and Simon Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 52(1): 201–217.
- Whitfield, Lindsay, Ole Therkildsen, Lars Buur, and Anne Mette Kjær. 2015. *The politics of African industrial policy: A comparative perspective*. Cambridge University Press.
- World Bank Group. 1997. *World Development Report: The State in a Changing World*. Washington, DC: The World Bank.

———. 2019. “IEG World Bank Project Performance Ratings, 1964-2016.” URL <http://data.worldbank.org/data-catalog/IEG>. Last updated on 18-Feb-2019.

Table A1: Prediction of Satisfactory IEG Rating by Loan/Agreement Type

	(1)	(2)	(3)	(4)
Capacity	0.51** (0.16)	0.30 <sup>^</sup> (0.16)	0.41** (0.11)	0.49* (0.24)
ProjectLength	-0.05** (0.01)	-0.06** (0.02)	-0.04** (0.01)	-0.11* (0.06)
lnGDPcap	-0.12 (0.10)	0.15 (0.11)	0.05 (0.07)	-0.09 (0.19)
YearsEduc	-0.00 (0.03)	-0.04 (0.03)	-0.04* (0.02)	0.05 (0.05)
Growth	0.09** (0.02)	0.11** (0.02)	0.12** (0.01)	0.06* (0.02)
ExRateShock	-0.23 (0.15)	-0.59* (0.30)	-0.36 <sup>^</sup> (0.18)	-0.04 (0.26)
Polity2	-0.01 (0.01)	0.00 (0.01)	-0.00 (0.01)	0.00 (0.02)
Constant	2.17** (0.75)	0.29 (0.80)	0.85 <sup>^</sup> (0.50)	1.66 (1.36)
N	5462	4222	8736	1310
Log-likelihood	-2907.10	-2452.47	-4884.81	-673.62
Type	IBRD	IDA	IPF	DPF

Logistic Regression with standard errors clustered at the country level.

<sup>^</sup>  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$

Table A2: Prediction of Ordinal Project Rating by Donor

	(1)	(2)	(3)	(4)	(5)
Capacity	0.40* (0.18)	0.01 (0.24)	0.47^ (0.27)	0.43 (0.34)	0.53* (0.21)
ProjectLength	-0.00** (0.00)		0.00** (0.00)	-0.00** (0.00)	-0.00** (0.00)
lnGDPcap	0.04 (0.24)	-0.23 (0.14)	0.02 (0.21)	0.18 (0.25)	-0.04 (0.12)
YearsEduc	-0.04 (0.05)	0.16* (0.06)	0.03 (0.07)	-0.11^ (0.06)	0.00 (0.05)
Growth	0.19** (0.05)	0.04* (0.02)	0.09** (0.03)	0.18** (0.06)	0.11** (0.03)
ExRateShock	2.09^ (1.13)	1.03 (0.98)	-4.57 (3.64)	1.50 (1.52)	-0.25 (0.33)
Polity2	0.01 (0.02)	-0.03* (0.01)	0.01 (0.02)	0.01 (0.02)	0.03* (0.01)
$\tau_1$	-2.36 (1.56)	-4.78** (0.86)	-1.90 (1.41)	-2.14 (1.80)	-5.62** (0.79)
$\tau_2$	-0.42 (1.56)	-3.60** (0.83)	-0.21 (1.43)	-0.53 (1.82)	-3.57** (0.81)
$\tau_3$	3.53* (1.48)	-1.69* (0.84)	1.83 (1.41)	1.28 (1.88)	-2.06** (0.80)
$\tau_4$		0.85 (0.87)			-0.44 (0.78)
$\tau_5$					1.57* (0.76)
N	1366	1431	427	658	1506
Log-likelihood	-1292.29	-1644.13	-459.91	-683.89	-2137.85
	AsianDB	DFID	GFATM	JICA	KfW

Ordered Logistic Regression with standard errors clustered at the country level.

^  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$

Table A3: IEG Rating of Borrower Government Performance (Ordinal)

	(1)	(2)	(3)	(4)
Capacity	0.50** (0.14)	0.49** (0.14)	0.55** (0.15)	0.50** (0.14)
ProjectLength	-0.09** (0.01)	-0.09** (0.01)	-0.09** (0.01)	-0.09** (0.01)
lnGDPcap	0.11 (0.10)	0.11 (0.10)	0.11 (0.09)	0.12 (0.10)
YearsEduc	-0.07** (0.02)	-0.08** (0.02)	-0.07** (0.02)	-0.08** (0.02)
Growth	0.05** (0.02)	0.06** (0.02)	0.05** (0.02)	0.06** (0.02)
ExRateShock	0.29 (0.29)	0.29 (0.29)	0.26 (0.29)	0.30 (0.29)
Polity2	-0.01 (0.01)			
PolRights		-0.01 (0.04)		
CivLibs			-0.08 <sup>^</sup> (0.05)	
Democracy				-0.10 (0.12)
$\tau_1$	-3.98** (0.67)	-4.03** (0.70)	-4.22** (0.72)	-4.00** (0.67)
$\tau_2$	-1.41* (0.66)	-1.46* (0.68)	-1.66* (0.70)	-1.44* (0.66)
$\tau_3$	-0.93 (0.66)	-0.98 (0.68)	-1.18 <sup>^</sup> (0.70)	-0.96 (0.65)
$\tau_4$	-0.24 (0.66)	-0.29 (0.68)	-0.49 (0.70)	-0.26 (0.65)
$\tau_5$	3.08** (0.66)	3.02** (0.68)	2.84** (0.69)	3.05** (0.65)
N	5966	5967	5967	5967
Log-likelihood	-7832.48	-7835.36	-7828.60	-7834.25

Ordered Logistic Regression with standard errors clustered at the country level.

<sup>^</sup>  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$

Table A4: Number of World Bank Projects and Change in State Capacity (CivLib)

	(1)	(2)	(3)	(4)
Project	0.00 (0.00)	0.00** (0.00)	-0.00* (0.00)	-0.00 (0.00)
Outcome	0.02** (0.01)	0.08** (0.02)		
Satisfactory			0.00** (0.00)	0.01* (0.00)
CivLib	0.02** (0.00)	0.04** (0.01)	0.02** (0.01)	0.04** (0.01)
Corruption	0.01^ (0.01)	0.03 (0.02)	0.02* (0.01)	0.04^ (0.02)
GDPcap	-0.00 (0.00)	0.00^ (0.00)	-0.00 (0.00)	0.00 (0.00)
YearsEduc	0.00 (0.00)	0.03** (0.01)	0.00 (0.00)	0.03** (0.01)
CapStart	-0.05** (0.02)	-0.28** (0.04)	-0.05** (0.02)	-0.27** (0.04)
Constant	-0.16** (0.05)	-0.57** (0.12)	-0.07* (0.03)	-0.30** (0.08)
N	518	427	519	427
Countries	88	87	88	87
$R^2$	0.08	0.16	0.07	0.14

Random-effects Generalized Least Squares regression with robust standard errors.

^  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$

Table A5: Number of World Bank Projects and Change in State Capacity (PolRights)

	(1)	(2)	(3)	(4)
Project	0.00 (0.00)	0.00** (0.00)	-0.00* (0.00)	-0.00 (0.00)
Outcome	0.02** (0.01)	0.08** (0.02)		
Satisfactory			0.00** (0.00)	0.01* (0.00)
PolRights	0.01** (0.00)	0.03** (0.01)	0.01** (0.00)	0.03** (0.01)
Corruption	0.01 (0.01)	0.02 (0.02)	0.01^ (0.01)	0.03 (0.02)
GDPcap	-0.00 (0.00)	0.00^ (0.00)	-0.00 (0.00)	0.00 (0.00)
YearsEduc	0.00 (0.00)	0.03** (0.01)	0.00 (0.00)	0.03** (0.01)
CapStart	-0.05** (0.02)	-0.27** (0.04)	-0.05** (0.02)	-0.27** (0.04)
Constant	-0.14** (0.05)	-0.53** (0.11)	-0.05* (0.03)	-0.26** (0.07)
N	518	427	519	427
Countries	88	87	88	87
$R^2$	0.07	0.16	0.07	0.14

Random-effects Generalized Least Squares regression with robust standard errors.

^  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$