Assessing the contribution of donor agencies to aid effectiveness: The impact of World Bank preparation on project outcomes

Christopher Kilby
Department of Economics, Villanova University, USA
chkilby@yahoo.com
September 12, 2012

Abstract:

This paper assesses the impact of World Bank project preparation on project outcomes. Using a stochastic frontier model, I generate a measure of World Bank project preparation duration based on variation in political economy factors that are exogenous to latent project quality. Panel analysis of project data finds that projects with longer preparation periods are significantly more likely to have satisfactory outcome ratings. This result is robust across a range of specifications but the effects are conditional on the degree of economic vulnerability. The impact of World Bank preparation is greater in countries experiencing debt problems. The impact of donor preparation identifies a positive contribution of development aid that is not subject to the usual fungibility critique.

Key words: Aid Effectiveness; Fungibility; Project Preparation; Stochastic Frontier Analysis; World Bank.

JEL codes: F35, F53, F55, O19

1. Introduction

Assessing the impact of development aid on economic development has proven difficult. The most direct measure of aid's impact, project-level assessment, is subject to a fungibility critique (Singer 1965). The project officially funded by an aid donor (Project A) might not be the additional project made possible via aid. If the recipient government would have undertaken Project A without donor funding, aid is fully fungible and actually finances some other activity (Project B). Thus the outcome of Project A may be irrelevant to assessing the impact of aid, giving only an upper bound that can be wildly optimistic. More generally, project-level assessments may tell very little about the overall impact of aid on economic development.

Despite mixed empirical evidence on the extent of aid fungibility (Pack and Pack 1990, 1993; Feyzioglu *et al.* 1998; van de Walle and Mu 2007), the fungibility critique has motivated research on the impact of aid flows at the aggregate level (economy-wide or across entire sectors within an economy). By abstracting from the individual projects nominally funded by donors, aggregate studies may capture the impact of aid, be it direct (from the projects officially funded) or indirect (from the additional activities made possible by easing the recipient government's budget constraint). The aggregate approach also allows for other spill-overs difficult to capture in disaggregate studies.

The results of aggregate studies have been disappointing, however. Questions about the utility of cross country regressions (particularly as the number of studies rivals the number of available data points) resurface periodically. Concerns about the endogeneity of aid in a growth regression dogged studies until Boone (1995) proposed geopolitical instruments as a solution. Yet this "solution" rests on a strong homogeneity assumption, i.e., the impact of aid is independent of donor motives (Deaton 2010; Kilby and Dreher 2010). For these and a host of

other reasons, studies report a wide range of results, leaving some scholars discouraged about the potential contribution of the aggregate approach (Doucouliagos and Paldam 2009).

Although project level analysis is unlikely to account for a project's full impact on the recipient economy, there remain plausible ways around the fungibility problem at the project level. Donor inputs into the preparation and supervision of projects may not be fungible in general if no comparable inputs from other sources (i.e., the recipient government) exist. Indeed, if recipient government preparation and supervision were substitutes for donor inputs, the measured impact of donor inputs should be biased toward zero when recipient government inputs are unobserved. This effectively reverses the direction of bias induced by the fungibility problem.

In this paper, I focus on measuring the impact of World Bank preparation on the outcome of World Bank-funded projects. This poses two problems. First, the amount of preparation is likely to be endogenous with extra preparation effort exerted when problems appear. Second, the World Bank does not make preparation data available to the public. The rest of this paper proceeds as follows to tackle these problems. Section 2 reviews the previous literature on World Bank project preparation as well as relevant work on the determinants of World Bank project performance. Key among these papers is Dreher *et al*. (forthcoming; henceforth DKVW) which also examines the impact on project outcomes of factors linked to project approval but without the explicit connection to preparation explored in this paper. Section 3 describes the stochastic frontier analysis (SFA) approach used in Kilby (2011b) and its application here to generate an exogenous measure of preparation. Section 4 presents estimation results from a project performance equation that includes preparation as an explanatory variable. Section 5 concludes.

2. Literature Review

Impact of Preparation

There have been a handful of studies that attempt to estimate the impact of project preparation on outcomes, all using World Bank data. The key challenge for these studies is the likely endogeneity of preparation. Donors have inside knowledge of project prospects (i.e., latent project quality) and so provide more inputs when project performance is in doubt. For example, when staff prepare a project that is high risk because it is a novel approach, is complex, takes place in a difficult environment, or previously has been poorly managed, they are likely to spend additional time to improve project design. As Denizer *et al*. (2011, 15; henceforth DKK) point out, "high risk" projects are more likely to receive intensive preparation but also are more likely not to be satisfactory on completion. To the extent that the researcher does not observe the underlying characteristics that signal risk, estimates of the impact of preparation on performance will suffer from omitted variable bias. This bias is likely to reduce the apparent impact of preparation and, if extra preparation is only partly successful in rectifying underlying shortcomings, the measured correlation or partial correlation may be negative.

Previous studies examining the impact of World Bank preparation have attempted to solve this endogeneity problem via an instrumental variables approach. Deininger *et al*. (1998) include the number of staff weeks of preparation in their analysis of the performance of World Bank-funded projects. A simple bivariate analysis finds higher average staff weeks of preparation in projects subsequently rated "unsatisfactory." In an instrumental variables analysis, World Bank project-specific inputs (preparation plus supervision) do not have a significant impact on a country's average performance although Deininger *et al*. note evidence that their instruments have not fully dealt with the endogeneity problem (footnote 3).

Looking just at World Bank-funded structural adjustment programs, Dollar and Svensson (2000) find that (instrumented) staff weeks of preparation do not influence program success rates. However, Dollar and Svensson demonstrate the exogeneity of their instruments (regional dummies, per capita income, and population) by showing that these variables are not significant in a performance equation that excludes preparation. That their instruments are uncorrelated with performance guarantees the later finding that instrumented preparation is insignificant and underscores the importance of theory-based exclusion restrictions.

DKK find a negative relationship between staff weeks of preparation and eventual project outcomes. The focus of their work, however, is to describe the data (to identify early warning signs of problem projects so that World Bank management can react in a timely fashion) rather than to uncover causal relations so the endogeneity of preparation is not a concern.

In sum, the small literature investigating the impact of project preparation on project performance is inconclusive. While it is intuitively appealing that poor or rushed preparation may lead to poor project selection or subsequent implementation problems (and, conversely, that good preparation pays real dividends), attempts to measure the impact of preparation are not wholly satisfactory because of limitations in the instrumental variables employed.[1]

[1] Kilby (1994) and Chauvet *et al*. (2006) use World Bank evaluations of the quality of preparation ("Quality at Entry") to assess the impact of preparation on project outcomes. Likewise, Limodio (2011) uses measures of World Bank "performance." However, as Kilby (1994) notes, these results are hard to interpret because of a halo effect, i.e., assessment of the project outcome may inform the evaluation of preparation (or other aspects of World Bank performance), inducing endogeneity. Focusing on project supervision, Kilby (2000) circumvents the feedback between performance and supervision by examining the link between supervision over a given year and the subsequent annual change in an intermediate measure of project performance. Because project performance is not assessed on an annual basis prior to implementation, this approach cannot be applied to preparation.

Determinants of Project Performance

Several previous papers examine the determinants of project performance as measured by World Bank project outcome ratings. DKVW is closest to the approach in this paper. The authors explore the impact of political factors reflecting the importance of the borrowing country (and hence privileged access to World Bank resources) on project outcomes. The basic question is whether favoritism shown to politically important countries in aid allocation has unintended negative consequences for the subsequent impact of that aid. This paper builds on DKVW by exploring shortened preparation time as the route by which political importance translates into lower performance.

The dependent variable in DKVW is a binary outcome rating. Key explanatory variables include temporary membership on the United Nations Security Council (UNSC), membership on the World Bank Executive Board, and measures of financial vulnerability (short term to total debt ratio and debt service to GDP ratio). In an analysis that includes country fixed effects, the authors find a robust link between temporary UNSC membership at the time of board approval and project outcomes, but only when the borrowing country was financially vulnerable (and hence in most need of immediate access to funds). This link persists even if the specification also includes similar political variables from the time of project evaluation, demonstrating that findings do not simply reflect rating bias.

DKK also use World Bank outcome ratings, either as a binary variable or a 1 to 6 scale. Explanatory variables include rating process variables (such as the lag between the end of implementation and evaluation and a dummy for ratings based on audits), macroeconomic/policy variables (including the World Bank's Country Policy and Institutional Assessment [CPIA] rating), basic project characteristics (project size, duration, preparation costs, and supervision costs), and early warning indicators. DKK find that 20 percent of the overall variation in project

performance is cross-country variation while a full 80 percent is within country variation, i.e., driven by project differences rather than country differences.

3. Stochastic Frontier Model

The above introduction identifies two challenges regarding World Bank preparation data. First, latent project quality may influence preparation, resulting in reverse causality (endogeneity). Second, the World Bank does not publish preparation data. This paper draws on Kilby (2011b) to circumvent both problems by constructing a predicted duration of project preparation that does not depend on project quality. Preparation duration is the length of time between project identification (unobserved) and project approval (observed). Because the identification date is not observed, I use sequentially generated Project Identification Numbers (Project IDs) as a noisy measure of the identification date in a stochastic frontier model (SFM) with the project approval date as the dependent variable. Independent variables include country and project characteristics that directly impact latent project quality but also geopolitical variables which do not. I then use this model to generate the predicted duration of preparation based on the geopolitical variables while holding country and project characteristics at their sample mean. The rest of this section motivates and summarizes this methodology.

Aigner *et al*. (1977) introduced the SFM to estimate production functions and cost functions. The estimation procedure needs to account for two issues. First, some firms are inefficient and fall short of the efficient frontier. Second, real world data include measurement error so that measured values for efficient firms may fall short of or even exceed the true efficient frontier. To allow for this, the stochastic frontier model includes two stochastic terms, a one-sided error term that reflects firm-level inefficiency and a symmetric error term that allows for measurement error.

One can recast the SFM as a duration model with normally distributed measurement error. This proves particularly useful for the current application since Project IDs provide a noisy measure of the start of project preparation. Duration in this context is akin to cost where the most "efficient" projects—the ones with the shortest duration—define the frontier. Thus, the methodology is analogous to duration analysis that simultaneously estimates the starting date based on a noisy measure of that date.

To derive the SFM formally, define the start of preparation (identification) as *ID Date* and the end of preparation as *Approval Date*. Let u be the duration of preparation. Then the approval date for project j in recipient country i is given by

$$Approval\ Date_{ij} = ID\ Date_{ij} + u_{ij} \tag{1}$$

I model the duration u_{ij} as an independent exponential process with variance

$$\sigma^2_{u_{ij}} = e^{\beta x_{ij}} \tag{2}$$

where $\boldsymbol{x}_{ij}$ are country and project/loan characteristics.[2] *ID Date* is not observed but a sequentially issued *Project ID* is. For ease of notation, consider a linear equation linking *ProjectID* to *ID Date*[3]:

$$ID\ Date_{ij} = \alpha + \gamma Project\ ID_{ij} + v_{ij} \tag{3}$$

In Equation (3) $1/\gamma$ is the average number of project identification numbers issued per day and v is assumed iid $N(0,\sigma^2_v)$. Combining (1) and (3) yields the model to be estimated:

$$Approval\ Date_{ij} = \alpha + \gamma Project\ ID_{ij} + v_{ij} + u_{ij} \tag{4}$$

[2] The mean of an exponential distribution equals the square root of the variance. This parameterization fits with the stochastic frontier literature and ensures that u is non-negative. One could also include a constant term in Equation (1), i.e., a minimum duration greater than zero; this has no practical effect given the constant introduced by Equation (3).

[3] I experimented with up to quartic terms for *Project ID*; the estimated relationship proves essentially linear.

With the distributional assumptions specified for v and u, this is the SFM for a cost function (Aigner *et al*. 1977); estimation is via maximum likelihood.

[Figure 1 about here]

Figure 1 presents the results of estimating this SFM. The line at the lower edge of the cloud of data points is the estimated frontier, i.e., the estimated identification date. The vertical distance between any data point and that line is the estimated duration of preparation for each project (net of measurement error). The results presented below focus on project and country characteristics which influence this duration. Note that both the duration and the impact of the explanatory variables on that duration are estimated simultaneously so that standard errors are correct in the sense that they do not treat an estimated duration as the actual duration.

Table 1 describes the sample for the SFM estimation.[4] Several factors determine the estimation sample. *Project IDs* for projects approved before 1994 or with numbers below 20,000 follow an earlier (not fully sequential) numbering system and are excluded. I also drop supplemental loans that provide additional funding to existing projects because preparation for these loans is very different.[5] About 225 of the 1752 projects with id numbers above 75266 are identified but not yet approved (as of July 5, 2010); to avoid censoring issues, I exclude this entire region. This leaves 1607 project observations in 110 countries though results are similar without the last two restrictions (for a total of 3627 project observations in 119 countries).

[4] This repeats the specification in Table 4, Column 3 of Kilby (2011b) except that *Project Size* replaces *Loan Amount* in keeping with DKVW. *Project Size* is "Loan Project Cost" from the World Bank Independent Evaluation Group database which reflects the overall cost of the project including World Bank loan amount, co-financing from other external sources, and counterpart funds from the borrowing government. Results do not depend heavily on this particular specification and sample.

[5] Coefficient estimates for the preparation equation are not dramatically different if I include supplemental loans but the distribution of predicted durations are bimodal with supplemental loans averaging 112 days and non-supplemental loans averaging 669 days. In addition, the World Bank does not rate supplemental loans so it is appropriate to exclude them here.

Three broad categories of variables enter the analysis: project variables, country variables, and political economy variables. Project variables include *Approval Date* (the dependent variable), *Project ID*, *Project Size*, and various indicators of loan type and sector. Country variables are those likely to impact the speed of preparation, including macroeconomic and governance/institutional quality variables. I also consider a range of donor interest political economy variables: UN voting alignment, non-permanent UNSC membership, World Bank Executive Board membership, trade flows, military aid, and bilateral economic aid.

[Table 1 about here]

Approval Date ranges from March 10, 1994 to June 29, 2010 with a mean of May 27, 2000. I include total project cost as a measure of project size, importance, and complexity. *Project Size* is measured as the log of millions of constant 2005 dollars, averaging 4.16 ($64 million) and ranging from 0.49 ($1.6 million) to 8.85 ($7 billion). *IDA* equals one if the project includes any IDA funding, true for 56 percent of the sample. *SAL* equals one if the loan/credit is a development policy loan. Some 16 percent of the observations are development policy loans.

A number of country characteristics may be important determinants of preparation duration. *War* is a dummy variable indicating an on-going conflict that claims at least 1000 lives during the year. Country descriptors also include *Population* (log of population), *GDP per capita* (log of the purchasing power parity GDP per capita in 2000 dollars), the Cheibub *et al*. (2010) *Democracy* indicator, and *Freedom House* (an average of the political freedom and civil liberties measures).

The remaining variables in Table 1 are country-level political economy measures and associated control variables. *US important votes* measures alignment with the U.S. on United Nations General Assembly (UNGA) roll call votes identified as important by the U.S. State Department. *US other votes* covers all other UNGA regular session roll call votes on resolutions

that passed. Calculations follow Kilby (2011a) and yield a theoretical range of 0 to 1. Alignment is substantially higher on important votes (0.50 versus 0.37); U.S. alignment measures trend down over time as UN voting has become more polarized (Voeten 2004). I include corresponding variables for the other G7 countries as a group, *G7-1 important votes* and *G7-1 other votes*. These also use the U.S. designation of votes as important or other, the appropriate choice when they serve purely as control variables. I postpone until Section 4 discussion of why this is a reasonable way to include UN voting alignment.

US military aid is 1 if the country receives substantial U.S. military aid (more than $500,000 in 2005 dollars), 0 otherwise. *US economic aid* is the log of U.S. total official gross disbursements of economic aid in millions of 2005 dollars. *G7-1 economic aid* is the same but for the other G7 countries (averaged over these donors then logged). Fleck and Kilby (2006) note that economic aid may also proxy for recipient need in this setting and suggest including *Like-minded donor economic aid*, i.e., aid from Denmark, the Netherlands, Norway, and Sweden. These countries have relatively humanitarian aid policies and very limited power within the World Bank. *US trade* is the log of exports plus imports in constant 2005 dollars; *G7-1 trade* is the same variable for the other G7 countries. I also include *World trade* so *US trade* and *G7-1 trade* capture only the differential effect of trade with these countries.

The last two variables record international positions the country might hold that increase its importance or power. *UNSC non-permanent member* equals 1 for those years the country occupied one of the temporary UNSC seats. *World Bank Executive Director* equals 1 if the country held an Executive Director position in the current year or past three years.

Several of the country and geopolitical variables trend over time, raising the possibility of spurious correlation. To address this issue, I use detrended variables where appropriate. In

addition, all project performance equations estimated below include year dummies to avoid this spurious correlation issue in the final stage.[6]

One tricky issue in this estimation is timing. The relevant values of the explanatory variables are at the start of and during the preparation period but, of course, that period is uncertain. To address this, I include time-varying factors with a three year lag (unless otherwise noted in Table 1) to allow for the time elapsed during preparation. In most cases, the length of lag (up to 3 years) has little impact on the coefficient estimate (in part due to serial correlation) but in a few instances results are stronger with the three year lag. Averaging over the three year period approaching approval yields similar results.

SFM Estimation Results

Project ID enters Equation (4) directly; all other variables enter the conditional variance of the exponential term in Equation (2). Table 2 does not report the coefficient estimate for *Project ID* (0.0597 with a z-statistic of 67.58) as interpretation of this coefficient is not particularly enlightening.

[Table 2 about here]

Table 2 presents results in two columns. The left column reports coefficient estimates and z statistics for basic project and country characteristics; the right column reports coefficient estimates and z statistics for political variables. As expected, *Project Size* enters with a positive and significant coefficient estimate indicating longer preparation periods for larger, presumably

[6] Including an annual time trend in the conditional variance of the SFM (Equation (2)) yields similar results (though detrending variables individually deals with the issue more thoroughly). The model fails to converge with year dummies in the conditional variance, a typical problem with nonlinear models. Including an annual time trend or year dummies directly in Equation (4) makes little sense as the residual is simply within year variation, i.e., the number of days between the start of the approval year and board approval.

more complex projects.[7] Projects receiving IDA funds have shorter preparation periods than those that receive no IDA funding but the difference is insignificant. Structural Adjustment Loans (*SAL*) have substantially and significantly shorter preparation periods. *War* enters with a negative but insignificant coefficient. The preparation period is longer for larger countries. *GDP per capita* enters with a negative coefficient but is insignificant (with or without the IDA dummy in the specification). *Democracy* is insignificant while *Freedom House* enters with a significant negative estimated coefficient (with or without the *Democracy* dummy). These results are broadly consistent with a range of specifications and samples examined in Kilby (2011b).

Turning to the political economy variables, the estimated coefficient for *US important votes* is negative and statistically significant while that for *US other votes* is substantially smaller, positive, and not statistically significant. For an otherwise typical project, an increase of one standard deviation in alignment with the U.S. on important UN votes corresponds to a 183 day (25%) reduction in the predicted duration of preparation. The picture is somewhat clouded by the positive and significant coefficient estimates for the other G7 countries. However, as Kilby (2011b) demonstrates, this latter result is not robust. If U.S. voting is omitted from the equation, the sample is modified, or the specification altered, other G7 votes cease to be significant.

The only other significant political economy variables are *UNSC non-permanent member* and *World Bank Executive Director*. Both enter with the expected negative coefficients. UNSC membership is associated with a 175 day (25%) reduction in preparation time while executive board membership is associated with a 157 day (22%) reduction in preparation time. Kilby

[7] *Project Size* has been decreasing over time (consistent with concerns about aid fragmentation). To account for this, the variable included in the equation is detrended.

(2011b) demonstrates that similar findings are robust to alternate approaches (without detrended data, with a wider sample of projects, and directly applying duration analysis to approximated duration data).

4. Project Performance

This section uses the SFM described above to construct an estimated preparation duration variable that is exogenous to latent project quality and then uses that measure of preparation as an explanatory variable in an analysis of World Bank project performance.[8] Below I first describe World Bank project performance ratings, then turn to the construction of a preparation variable based purely on variation in geopolitical factors, and finally use this variable to assess the impact of World Bank preparation on project outcomes.

Project Evaluation

The Independent Evaluation Group (IEG—formerly the Operations Evaluation Department or OED) is a semi-autonomous branch of the World Bank that reports directly to the Board of Executive Directors. The primary function of IEG is to conduct ex post evaluations of World Bank projects and policies.[9] In keeping with this mandate, IEG records performance ratings for virtually all completed World Bank-funded projects in a database (see IEG (2011c) and discussion below).

[8] This approach is similar to using a preparation prediction based on actual values of all variables in the SFM and then instrumenting this predicted value in the performance equation with the geopolitical characteristics from the SFM. However, the method used is more efficient than the IV approach. See Wooldridge (2002, 623-625) for a parallel approach with a first stage probit function. In this setting, bootstrapped standard errors may be appropriate (see below).

[9] OED was established as an independent department in 1973 (Grasso *et al*. 2003) and renamed IEG in November of 2005. IEG staff rules and procedures are designed to limit staff conflict of interest and promote objective evaluation (OED 2003).

A number of project ratings are available in the IEG database. At the end of the implementation phase (typically seven years after Board approval (Phillips 2009, 166)), the operational team leader in charge of supervising the project submits an Implementation Completion Report that includes categorical project ratings. Up through the end of 1996, these ratings appear in the IEG database as Project Completion Ratings (PCRs). Phased in starting in early 1995, IEG policy shifted to include an additional "validation" step before ratings—now termed Evaluation Summary or Evaluation Memorandum ratings—enter the database (DKK). IEG then audits some projects, generating a Project Performance Assessment Report (PPAR) and adds a new set of ratings (PARs) to the database.[10] Audit sample selection depends on a number of factors including particularly good or bad outcomes, projects in sectors subject to IEG review, and projects in audit clusters (to reduce audit expenses).[11] PPARs are typically completed 1 to 5 years after the project closes (i.e., the close of disbursement of the IBRD loan or IDA credit).[12]

The system used by IEG has evolved from a single dichotomous outcome rating into multiple, polychotomous ratings. However, most research and policy discussions focus on the

[10] Prior to FY 1983, IEG replaced old PCR ratings with new PAR ratings. Starting in FY 1983, IEG phased out this practice so that the replacement of initial ratings was virtually eliminated by 1995 with the introduction of the validation step. This is relevant for discussions below of changes from initial ratings to audit ratings, lags between initial ratings and audit ratings, etc. Calculating evaluation lags from closing date to evaluation date presents its own challenges since closing dates are not always accurate (missing, after evaluation dates, etc.).

[11] There is evidence that audits target projects with higher ratings. Initial outcome ratings average 72% satisfactory for projects that are not subsequently audited versus 80% satisfactory for projects that are later audited. When projects are audited, 10% are downgraded from satisfactory to not satisfactory while only 3% are upgraded from not satisfactory to satisfactory. As DKK note, the tendency of audits to downgrade may be due at least in part to the passage of time. Of course, it may be that critical evaluations simply take longer to prepare or that their release is delayed.

[12] The audit rate has declined over time and is currently at 25%. IEG devotes an average of six staff weeks to each PPAR, usually including a field mission to the borrowing country (IEG, 2011a). IEG budget for fiscal year 2011 was $34 million. Approximately $3 million was used for IBRD and IDA project evaluations; the remainder of the budget was spent on broader sector, thematic, or country reviews, evaluation of IFC and MIGA projects, and other initiatives (IEG, 2011b, 38).

original outcome rating reduced to a binary variable. Studies examining ratings in both raw and binary forms (e.g., DKK, DKVW) generally do not find compelling reasons to use the more fine-grained version. The analysis here follows the bulk of the literature, using the binary version of the most recent outcome rating (*Outcome*) for each project. This rating ostensibly measures project outcomes relative to objectives stated in the project appraisal and loan documents though there is evidence that an economic rate of return cut-off of 10% (i.e., an absolute standard) is used to distinguish between "Satisfactory" and "Not Satisfactory" where such figures are available (Kilby 2000).[13]

Over the period studied (approval dates between 1986 and 2008), not all entries in the World Bank Projects Database—the main source of other project data—have corresponding entries in IEG's ratings database. Over this period, there are 4691 unique entries in the World Bank Projects Database for country-specific IBRD/IDA projects with closing dates before 2011.[14] Of these, 417 have no matching entry in the IEG database. The vast majority of "missing" projects are not in the IEG database because they only recently closed. The share without IEG ratings is 85% for projects closing in 2010 but declines rapidly going back to earlier years so that, overall, only 4% of projects closed before 2010 still lack ratings. These few earlier projects may reflect cancellations. The World Bank Projects Database includes projects cancelled before significant implementation (e.g., the borrower never signed the project loan documents). For these cases, there would be no implementation to evaluate. Thus, the IEG

[13] This pattern is apparent in IEG (2010); Appendix B reports that only 12% of projects with an economic rate of return above 10% were rating "moderately unsatisfactory" or lower. DKK also argues that World Bank procedures promote applying relatively uniform standards to goal setting and evaluation.

[14] This count excludes supplemental loans, the preponderance of which do not report closing dates. IEG generally does not evaluate supplemental loans and I also rule these out because of their unusual preparation features.

sample covers virtually the entire relevant population so that sample selection does not appear to be an issue at this stage.

The estimation sample itself is largely determined by availability of the preparation variable constructed from the SFM. Preparation duration predictions use only the parameter estimates in conditional variance of the exponential term and so do not require project id numbers. Although the model in Section 3 could only be estimated using data after 1993 (when project ids became fully sequential), predictions—both in and out of sample—are possible as long as country and project data are available. The UN important vote alignment measure is the main limiting factor. The U.S. State Department began publishing its list of important votes in 1983; with the three year lag used, this means that projects must be approved after 1985 to be included in the sample. The latest approval date (2008) is driven by the availability of rating data discussed above. Measured by the year of IEG's evaluation, data run from 1989 to 2011. The estimation sample is reduced to 4147 due to missing data for country characteristics.

[Table 3 about here]

Table 3 reports descriptive statistics for this sample. The *Outcome* rating averages 72.5% satisfactory across the sample; performance varies considerably over time and across countries.[15] Preparation duration ranges from 0.9 years to 7.5 years with a mean duration just over two years. It is important to note that these predicted values of preparation duration are based on UN voting alignment with the U.S., UNSC membership, and World Bank Executive Board membership—which vary by country and year—but other variables that could reflect latent project quality (including project characteristics and macroeconomic factors) are held at the sample mean.

[15] Grouping projects by approval year, the satisfactory rate in the estimation sample ranges from 60% in 1986 to 81% in 2006; grouping instead by evaluation year, the range is from 59.0% in 1994 to 79.6% in 2009. For countries with at least ten projects, the success rate ranges from 10% to 100%.

I include only a very limited set of other variables in the core specification because later specifications include year dummies and fixed effects. These variables are *Project Size*, *Population*, *GDP per capita*, and *GDP growth*. *Project Size* is the log of total project cost and averages 4.326 ($75 million), ranging from -0.6488 ($0.5 million for an agribusiness project in Burundi in 1992) to 8.883 ($7.2 billion for a power project in Turkey in 1991). World Bank lending accounts for over 80 percent of the financing of these projects on average and results are similar if we use the World Bank loan amount instead. *Population*, again in logs, averages 16.95 (9.3 million) and ranges from 10.6 (40,000) to 20.98 (1.3 billion). *GDP per capita* (log) averages 7.849 ($2562) with a low of 5.968 ($390) and a high of 9.731 ($16,831). Finally, *GDP growth* has a sample mean of 2.6%, a low of -31% (Moldova 1994) and a high of 275% (Cambodia 1993). Estimation results are not sensitive to excluding the extreme values of these variables. The values of the country variables (population, GDP, and growth) are for the year of project approval.[16]

Performance Equation Estimation Results

Table 4 reports logit estimates for project performance. All specifications include year dummies. The first column presents a baseline specification that excludes preparation. *Population*, *GDP per capita*, and *GDP growth* all enter with positive and significant coefficient estimates. The coefficient for *Project Size* is positive but not significant once *GDP growth* is included.

[Table 4 about here]

[16] DKK find a strong partial correlation between World Bank CPIA ratings and IEG project ratings. I do not include CPIA ratings here as they are to some degree subjective and hence may reflect geopolitical factors themselves. Also, CPIA ratings are publicly available only for IDA countries and only since 2005.

The second column introduces *Preparation*. In contrast to earlier attempts to measure the impact of preparation on performance (where endogeneity remained an issue, e.g., Deininger *et al*. 1998; DKK; Dollar and Svensson 2000), predicted preparation duration enters with a positive and significant coefficient estimate. This result holds across the increasingly demanding specifications of Table 4. Column (3) introduces region dummies (finding worse performance in Sub-Saharan Africa, South Asia, and Middle East-North Africa as compared to Europe and Central Asia). *Population* and *GDP per capita* cease to be significant factors once we account for regional differences (at least in part because much of the variation in these characteristics is by region) while *Project Size* becomes significant.[17] Column (4) adds a separate dummy for China (which has particularly strong performance). Column (5) replaces region dummies with country fixed effects in a conditional logit; the sample shrinks by 131 observations due to 19 countries with no variation in outcomes. *Project Size* and *GDP growth* cease to be significant.

The final specification takes this one step farther, introducing government fixed effects. I include a separate fixed effect for each government that differs substantially from its predecessor, i.e., when the government changes and the country's Polity score changes by more than 3 points. To understand the importance of government fixed effects in this context, a short digression is necessary.

The way in which I include UN alignment in the SFM used to generate the preparation variable is motivated by the vote buying model of Andersen, Harr, and Tarp (2006). Andersen, Harr, and Tarp differentiate between important votes—on which the donor lobbies other

[17] In DKK, the log of loan size (closely related to *Project Size*) enters with a negative and significant coefficient. The difference may be driven by different specifications (fixed effects or not), different covariates, and somewhat different samples. Some DKK variables are not publicly available; other covariates are relevant for World Bank decision making but not exogenous (e.g., project length, supervision costs, intermediate performance flags). In DKVW, *Project Size* is negative but not significant.

governments intensively—and other votes. Only in the second set of votes does the vote cast by the other government reflect that government's true preferences, free of donor influence. A government's alignment with the donor on these votes reflects the government's ideal location (relative to the donor) in the voting space. Conversely, votes on important resolutions (on which the donor lobbied intensively) reflect concessions made by the government. Andersen, Harr, and Tarp demonstrate that estimates of donor vote buying which do not control for the recipient government's ideal point will be biased. They argue that, to control for the ideal point, specifications should include either voting alignment on other votes or country fixed effects. The SFM in Section 3 takes the first approach.

Suppose, however, that vote buying does not take place. If a new government comes to power with a more internationalist, pro-western orientation, we would simultaneously see a shift in UN voting toward the U.S. position and a demand-driven acceleration in World Bank borrowing. If alignment on other votes is a good proxy for the borrowing government's ideal position on important issues, there is no problem: both alignment measures shift and measured concessions to the U.S. do not increase. However, if voting on non-important issues is not a good proxy, omitted variable bias becomes a real problem. Although the U.S. does not pressure the World Bank (in this scenario with no vote buying), a voting shift toward the U.S. goes hand in hand with accelerated preparation.

This suggests that country fixed effects also may not be sufficient because they do not capture within-country changes. Government fixed effects, however, should capture exactly the relevant within-country changes that predicted preparation might inadvertently include. Including government fixed effects again reduces the sample slightly but the fundamental result

remains.[18] Even in the government fixed effects specification, *Preparation* enters with a positive and significant coefficient estimate. The magnitude of the coefficient is relatively stable across all five specifications.

[Figure 2 about here]

Figure 2 gives a sense of the magnitude of the effect of preparation duration on project outcomes.[19] For a typical project (i.e., all values set at the sample mean), the probability derivative is 0.043. Put in more concrete terms, for an otherwise typical project with a preparation duration one standard deviation below the mean, the predicted probability of a satisfactory outcome is 70.1%. For the same project with preparation duration one standard deviation above the mean, the predicted probability of success rises to 77.8%. Looking instead at the extremes, for an otherwise typical project with the shortest preparation period (0.7 years), the predicted probability of success is 67.4%; this rises to 90.5% with the longest preparation period (7.5 years).

Finally, bootstrapped standard errors differ little from conventional estimates in this case. For example, in Column 6 of Table 4, the key variable of interest (*Preparation*) remains statistically significant at the 95% confidence level with the z statistic decreasing only slightly from 2.31 ($p = 0.021$) to 2.17 ($p = 0.030$).

[18] Switching from country to government fixed effects drops an additional 97 observations (29 governments in 24 countries), again due to lack of variation in project outcome. In 11 cases, this is because there was only one observation for the government. Despite these dropped observations, the number of countries only drops by one to 116. Following the argument above, the government in question is the one in power at the time of UN voting, i.e., three years before project approval (t-3). However, results are the same if I use the government in power at the time of project approval (t).

[19] This simulation is based on Table 4, Column 6; results are similar for other specifications.

Extensions

Tables 5 and 6 explore alternative specifications including those suggested by DKK and DKVW. All specifications build on the final column in Table 4, i.e., conditional logit with government fixed effects and year dummies. Table 5 presents simple extensions with additional covariates; the results are illuminating though the central preparation result is unchanged; *Preparation* enters with a positive and significant coefficient with roughly the same magnitude as in Table 4. Table 6 considers the "economic vulnerability" hypothesis raised by DKVW through interaction terms. These specifications yield results consistent with DKVW, suggesting a possible interpretation for my findings as well as for their findings.

[Table 5 about here]

The first three columns of Table 5 explore the role of the type of rating and the evaluation lag (the time between project closing and IEG's evaluation). *Audit Rating* equals 1 if the dependent variable is a PAR rating (true for 24.1% of the sample). *Evaluation Lag* averages 1.6 years (ranging from 6 days to 12.5 years); I treat it as missing if the project closing date is missing or reported as taking place after the evaluation date. Column (1) illustrates that audit ratings are not significantly different than ratings for projects that are not audited (though, as noted before, the prior rating average of the audit sample is higher). Column (2) demonstrates that longer evaluation lags are linked to worse outcomes. Column (3) allows for the effects of both (relevant because evaluation lags are typically longer for audits). In this specification, we see that, controlling for the difference in evaluation lags, audits are actually more likely to result in satisfactory ratings. For an otherwise typical project (including its evaluation lag), the probability of success is 73% if not audited, 82% if audited. Looking at evaluation lags instead, the predicted probability of success falls from 77% to 67% if the evaluation lag increases from 1 year (the average lag for an unaudited project rating) to 3.4 years (the average lag for an audit).

Of course, this addresses neither the direction of causation for lags (do projects deteriorate over time or does it take longer to evaluate—or report—a bad project) nor the selection issue for audits.[20]

Column (4) includes a dummy variable indicating IDA funding. For an otherwise typical project, satisfactory ratings are 6% more likely for IDA-funded projects (77% versus 71%) though the difference is only marginally significant. Structural Adjustment projects (Column (5)) are significantly more likely to receive a satisfactory rating (78% versus 73%). Finally, using a sectoral breakdown similar to DKVW and DKK, Column (6) finds that satisfactory ratings are more likely for transportation projects as compared to agricultural projects but less likely for energy and mining projects.

Table 6 explores the economic vulnerability hypothesis. Following Stone (2008), DKVW argue that governments will exercise their political power when their need is greatest, i.e., when they are economically vulnerable. The authors measure economic vulnerability with two debt variables, short term debt as a percent of total external debt (*% Short Term Debt*) and total debt service to GNI (*Debt Service/GNI*). The conditional effects of these variables can then be captured via interaction terms.

Looking at the duration of preparation, the story is as follows. Shorter preparation duration is correlated with, but does not always imply, the rushed delivery of an underprepared project. When a borrowing country is economically vulnerable, it is both more likely to make use of its political power and more willing to accept a weak project (due to rushed preparation) because the immediate cash flow benefits outweigh the long run costs. Following DKVW, I operationalize this by interacting economic vulnerability with *Preparation*.

[20] The interaction of *Audit Rating* and *Evaluation Lag* proves insignificant so the effects of time appear to impact audit and other ratings equally.

[Table 6 and Figures 3 and 4 about here]

Table 6, Column (1) introduces the vulnerability variable *% Short Term Debt*. The sample is reduced slightly due to limited availability of debt data. Short term debt enters with a very small positive but statistically insignificant coefficient. This indicates that, *ceteris paribus*, economic vulnerability at the time of project approval has little impact on the eventual outcome of the project. Column (2) introduces the interaction term though meaningful interpretation of the table itself is difficult at best (Ai and Norton 2003); see instead Figure 3 which depicts these results for the range of values of *% Short Term Debt*. The upward sloping solid line indicates the marginal effect of preparation conditional on the indicated level of economic vulnerability (here, short term debt). The dashed lines depict the 95% confidence interval for the marginal effect which does not include zero for values of *% Short Term Debt* at or above 7%. The histogram indicates the sample distribution of *% Short Term Debt*, showing that the marginal effect of preparation is significantly different from zero for about two thirds of the observations.[21] Looking at *% Short Term Debt* one standard deviation below its mean (at 1.6%), the change in the probability of a satisfactory rating when *Preparation* goes from one standard deviation below to one standard deviation above its mean is 2.9%. Repeating this calculation for *% Short Term Debt* one standard deviation above its mean (20.3%), the probability differential increases to15.2%, i.e., the impact of preparation on the probability of success is five times higher.

Table 6, Columns (3) and (4) and Figure 4 repeat the exercise for the other economic vulnerability measure with parallel results. The solid line indicates the marginal effect of preparation conditional on the indicated level of *Debt Service/GNI*. The histogram illustrates the sample distribution of *Debt Service/GNI* (ranging from 0 to 108%). For values of *Debt*

[21] It is also apparent from this histogram that there are a number of outliers in terms of *% Short Term Debt*. That said, the results are robust to excluding the 112 observations with very high short term debt (more than 30 percent of total debt).

Service/GNI of 5% or more (60% of the sample), the conditional marginal effect of preparation on performance is statistically significant.[22] Looking at *Debt Service/GNI* one standard deviation below its mean (at 0.44%), the change in the probability of a satisfactory rating when *Preparation* goes from one standard deviation below to one standard deviation above its mean is 3.8%. Repeating this calculation for *Debt Service/GNI* one standard deviation above its mean (11.1%), the probability differential is 11.8%. The impact of preparation on the probability of success is three times higher.

These findings are consistent with the DKVW economic vulnerability hypothesis. For borrowing countries in a strong economic position at the time the project is approved, the duration of preparation is not a significant factor in determining project outcomes. However, for the sizable share of World Bank borrowers that are economically vulnerable, longer World Bank project preparation is significantly and positively related to good project outcomes. This is controlling for government fixed effects and the level of economic vulnerability so it captures an independent effect of preparation rather than simply proxying for one of these covariates.

5. Conclusion

Accurately measuring the contribution of aid to economic development at the project level is notoriously difficult because money is fungible. The fungibility problem is one reason for the recent focus on macro level analysis and has been offered as an explanation for the micro/macro paradox (Hudson and Mosley 2001). This study takes an alternate approach, examining the impact of World Bank preparation (a non-monetary input) and effectively side-stepping the fungibility issue.

[22] Dropping outliers (*Debt Service/GNI* above 15 percent—99 observations) again does not change this result.

However, examining preparation poses new problems. First, preparation is likely to be endogenous. The World Bank devotes extra resources to problem projects in its pipeline in an attempt to bring these projects up to some minimum standard. Second, preparation data are not publicly available. In this paper, I tackle both problems by generating a measure of the duration of preparation using political economy variables—including UN voting alignment, UNSC non-permanent membership, and World Bank Executive Board membership—that influence preparation but are otherwise exogenous to latent project quality. To do this, I employ the stochastic frontier model from Kilby (2011b) that uses sequentially issued project id numbers and approval dates to back-out information about the unobserved length of preparation. I find that preparation has a sizeable positive, significant, and robust impact on project performance that increases with economic vulnerability (measured by the country's ratio of short to long term debt or the debt service ratio).

These findings fit well with past research on the political economy of World Bank lending which indicates that powerful donors (chiefly the U.S.) use access to World Bank resources to pursue their own interests. Borrowing countries that hold institutionally or geopolitically important positions (Executive Board membership or UNSC non-permanent membership) or make concessions to donors in UN voting receive more loans (Dreher et al. 2009), more funding (Andersen, Hansen, and Markussen 2006; Kaja and Werker 2010), and faster loan disbursement (Kilby 2009, 2013). As a bi-product, projects may be rushed through the preparation phase (Kilby 2011b). Dreher *et al*. (forthcoming) show that such politically motivated aid leads to significantly worse outcomes when countries are economically vulnerable.

Results for the duration of World Bank project preparation parallel this. When a government is in a strong position regarding its external debt, it avoids ill-prepared projects, either by being selective (picking only those projects that have been designed well despite the

compressed time frame) or by supplementing World Bank preparation with its own resources. With the luxury of a longer time horizon, these choices are sensible and feasible. However, when a country faces pressing debt problems, the government makes different choices, accepting new loans for their short term benefits regardless of the long term consequences. For the average developing country, this means World Bank preparation makes a real contribution to aid effectiveness. For particularly vulnerable countries, this contribution can be large.

References

Ai, C. Chunrong, and Edward C. Norton. 2003. "Interaction terms in Logit and Probit models." *Economics Letters* 80(1):123-129.

Aigner, Dennis, C.A. Knox Lovell, and Peter Schmidt. 1977. "Formulation and estimation of stochastic frontier production function models." *Journal of Econometrics* 6:21-37.

Andersen, Thomas Barnebeck, Henrik Hansen, and Thomas Markussen. 2006. "US politics and World Bank IDA-lending." *Journal of Development Studies* 42(5):772-794.

Andersen, Thomas Barnebeck, Thomas Harr and Finn Tarp. 2006. "On US politics and IMF lending." *European Economic Review* 50:1843-1862.

Boone, Peter. 1995. "Politics and the effectiveness of foreign aid." NBER Working Paper 5308.

Chauvet, Lisa, Paul Collier, and Andreas Fuster. 2006. "Supervision and project performance: A principal-agent approach." The Centre for the Study of African Economies Working Paper, Oxford University.

Cheibub, José Antonio, Jennifer Gandhi and James Raymond Vreeland. 2010. "Democracy and dictatorship revisited." *Public Choice* 143(1-2): 67-101.

Deaton, Angus. 2010. "Instruments, randomization, and learning about development." *Journal of Economic Literature* 48(2):424-455.

Deininger, Klaus, Lyn Squire, and Swati Basu. 1998. "Does economic analysis improve the quality of foreign assistance?" *World Bank Economic Review* 12(3):385-418.

Denizer, Cevdet, Daniel Kaufmann, and Aart Kraay. 2011. "Good countries or good projects? Macro and micro correlates of World Bank project performance." World Bank Policy Research Working Paper 5646.

Dollar, David and Jakob Svensson. 2000. "What explains the success or failure of structural adjustment programmes?" *The Economic Journal* 110(466):894-917.

Doucouliagos, C., and Martin Paldam. 2009. "The aid effectiveness literature: The sad results of 40 years of research." *Journal of Economic Surveys* 23(3):433-461.

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* 88:1-18.

Dreher, Axel, Stephan Klasen, James Raymond Vreeland, Eric Werker. Forthcoming. "The costs of favoritism: Is politically-driven aid less effective?" *Economic Development and Cultural Change*.

Feyzioglu, Tarhan, Vinaya Swaroop and Min Zhu. 1998. "A panel data analysis of the fungibility of foreign aid." *World Bank Economic Review* 12(1):29-58.

Fleck, Robert K. and Christopher Kilby. 2006. "World Bank independence: A model and statistical analysis of U.S. influence." *Review of Development Economics* 10(2):224-240.

Freedom House. 2009. *Freedom in the World*. http://www.freedomhouse.org/uploads/fiw09/CompHistData/FIW_AllScores_Countries2009.xls Accessed 5/17/2009.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand. 2002. "Armed conflict 1946-2001: A new dataset." *Journal of Peace Research* 39(5): 615-637. Updated data set: http://new.prio.no/CSCW-Datasets/Data-on-Armed-Conflict/UppsalaPRIO-Armed-Conflicts-Dataset/ file="Main Conflict Table v4_2009.xlsx" Accessed on 8/27/2009.

Grasso, Patrick G., Sulaiman S. Wasty, Rachel V. Weaving (Eds). 2003. *World Bank Operations Evaluation Department: The First 30 Years*. Washington, DC: The World Bank.

Heston, Alan, Robert Summers, and Bettina Aten. 2002. "Penn World Tables 6.1." Center for International Comparisons of Production, Income and Prices, University of Pennsylvania. Accessed 9/25/2006.

Heston, Alan, Robert Summers, and Bettina Aten. 2006. "Penn World Tables Version 6.2." Center for International Comparisons of Production, Income and Prices, University of Pennsylvania. Accessed 8/29/2009.

Hudson, John and Paul Mosley. 2001. "Aid policies and growth: In search of the Holy Grail." *Journal of International Development* 13(7):1023-1038.

IEG (World Bank Independent Evaluation Group). 2010. *Cost-Benefit Analysis in World Bank Projects*. Washington, DC: The World Bank.

IEG (World Bank Independent Evaluation Group). 2011a. "Evaluation methodology." IEG website. Accessed 10/13/2011. http://web.worldbank.org/external/default/main?theSitePK=1324361&piPK=64252979&pagePK=64253958&menuPK=5039271&contentMDK=20791122

IEG (World Bank Independent Evaluation Group). 2011b. *IEG Work Program and Budget (FY12) and Indicative Plan (FY13-14)*. Washington, DC: The World Bank.

IEG (World Bank Independent Evaluation Group). 2011c. IEG World Bank Project Performance Ratings. Washington, DC: The World Bank. http://databank.worldbank.org/databank/download/IEG-WB-Project-Ratings.xlsx

International Monetary Fund. 2009. *Direction of Trade CD-ROM*.

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.

Kilby, Christopher. 1994. "Risk management: An econometric investigation of project level factors." Background paper for the *World Bank's Operations Evaluation Department Annual Review of Evaluation Results 1994*.

Kilby, Christopher. 2000. "Supervision and performance: The case of World Bank projects." *Journal of Development Economics* 62(1):233-259.

Kilby, Christopher. 2009. "The political economy of conditionality: An empirical analysis of World Bank loan disbursements." *Journal of Development Economics* 89(1):51-61.

Kilby, Christopher. 2011a. "Informal influence in the Asian Development Bank." *Review of International Organizations* 6(3-4):223-257.

Kilby, Christopher. 2011b. "The political economy of project preparation: An empirical analysis of World Bank projects." Villanova School of Business Economics Working Paper #14.

Kilby, Christopher. 2013. "An empirical assessment of informal influence in the World Bank." *Economic Development and Cultural Change*, forthcoming.

Kilby, Christopher and Axel Dreher. 2010. "The impact of aid on growth revisited: Do donor motives matter?" *Economic Letters* 107(3):338-340.

Limodio, Nicola. 2011. "The success of infrastructure projects in low income countries and the role of selectivity." World Bank Policy Research Working Paper # 5694.

OECD Development Cooperation Directorate. 2006-2009. *International Development Statistics, CD-ROM*.

OED (World Bank Operations Evaluation Department). 2003. "Independence of OED." *OED Reach* February 24. Accessed 10/13/2011. http://lnweb90.worldbank.org/oed/oeddoclib.nsf/b57456d58aba40e585256ad400736404/50b9b24456b788be85256ce0006ee9f8/$FILE/Independence_Reach.pdf

Pack, Howard and Janet Rothenberg Pack. 1990. "Is foreign aid fungible? The case of Indonesia." *The Economic Journal* 100(March):188-194.

Pack, Howard and Janet Rothenberg Pack. 1993. "Foreign aid and the question of fungibility." *Review of Economics and Statistics* 75(2): 258-265.

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

Singer, Hans. W. 1965. "External aid: For plans or projects?" *The Economic Journal* 75:539-545.

Stone, Randall. 2008. "The scope of IMF conditionality." *International Organization* 62: 589-620.

United Nations. 2010. "UN Security Council members." http://www.un.org/sc/list_eng5.asp Accessed 9/1/2010.

USAID. 2009. U.S. overseas loans and grants: obligations and loan authorizations, July 1, 1945-September 30, 2007 (Greenbook). http://qesdb.usaid.gov/gbk/ Accessed 5/15/2009.

U.S. State Department. 1984-2010. *Voting Practices in the United Nations.* GPO, Washington.

van de Walle, Dominique and Ren Mu. 2007. "Fungibility and the flypaper effect of project aid: Micro-evidence for Vietnam." *Journal of Development Economics* 84(2):667-685.

Voeten, Erik. 2004. "Resisting the lonely superpower: Responses of states in the United Nations to U.S. dominance." *Journal of Politics* 66(3):729-754.

Voeten, Erik and Adis Merdzanovic. 2009. "United Nations General Assembly voting data." http://hdl.handle.net/1902.1/12379 UNF:3:Hpf6qOkDdzzvXF9m66yLTg== Access 7/23/2009.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

World Bank. 2009. *World development indicators.* http://www.worldbank.org/data/onlinedatabases/onlinedatabases.html Accessed 5/15/2009.

World Bank. 2010. *Projects database*. http://go.worldbank.org/IAHNQIVK30 Accessed 7/5/2010.

Table 1: Descriptive Statistics for SFM

Variable	Mean	StDev	Min	Max	Description
Approval Date	04/27/2000		3/10/1994	6/29/2010	Approval date (month/day/year)
Project ID	55297		31828	75256	World Bank Project ID
Project Size	4.163	1.384	0.493	8.85	Log of Total Project Cost in constant 2005 $ millions
IDA	0.5625		0	1	IDA funds dummy
SAL	0.1568		0	1	Structural Adjustment Loan dummy
War	0.08339		0	1	Dummy indicating major conflict (>1000 dead) t-3
Population	17.02	1.82	11.93	20.97	Log of population t-3
GDP per capita	7.868	0.8146	5.801	9.609	PPP GDP per capita in chained 2000 $ t-3
Democracy	0.5177		0	1	Democracy dummy t-3
Freedom House Index	4.143	1.491	1	7	Averaged Freedom House Rating t-3
US important votes	0.4993	0.1622	0.08333	0.85	Alignment with US on UN votes important to US t-3
US other votes	0.3682	0.1145	0.1349	0.6667	Alignment with US on other UN votes t-3
G7-1 important votes	0.7201	0.1388	0.4038	0.9848	Alignment with other G7 on UN votes important to US t-3
G7-1 other votes	0.7263	0.08259	0.5608	0.951	Alignment with other G7 on other UN votes t-3
US military aid	0.4624		0	1	Dummy for US military aid>0.5 (2005 $ millions) t-3
US economic aid	3.763	1.705	-2.204	7.902	Log of disbursements of US economic aid (2005 $ millions) t-3
G7-1 economic aid	3.915	1.644	-1.364	7.348	Log average disbursements of G7-1 economic aid (2005 $ millions) t-3
Like-minded donor economic aid	1.801	1.553	-3.532	4.652	Log average disbursements like-minded donor aid (2005 $ m.) t-3
US trade	6.856	2.56	-0.125	12.52	Log of US trade (imports+exports) with country (2005 $ millions) t-3
G7-1 trade	7.857	2.058	2.619	12.21	Log average of G7-1 (IM+EX) with country (2005 $ millions) t-3
World trade	9.246	1.956	5.036	13.61	Log World trade (imports+exports) with country (2005 $ millions) t-3
UNSC non-permanent member	0.07156		0	1	Indicates country holds non-permanent UNSC seat t-2
World Bank Executive Director	0.3167		0	1	Country held World Bank ED seat in current year or past 3 years

1607 observations
Time dependent variables measured relative to project approval year (t), e.g., t-3 is three years prior to approval.

Table 2: Stochastic Frontier Model of Preparation Duration
Dependent Variable: Approval Date

Project/Country Variables	
Project Size	0.142**
	(2.77)
IDA	-0.301
	(-1.52)
SAL	-0.750**
	(-4.96)
War	-0.0275
	(-0.13)
Population	0.275**
	(2.35)
GDP per capita	-0.178
	(-0.89)
Democracy	0.249
	(1.33)
Freedom House Index	0.0477
	(0.68)

Geopolitical Variables	
US important votes	-3.882**
	(-4.04)
US other votes	0.915
	(1.05)
G7-1 important votes	2.568**
	(2.44)
G7-1 other votes	2.369**
	(2.09)
US military aid	0.161
	(1.29)
US economic aid	0.0443
	(1.05)
G7-1 economic aid	-0.0447
	(-0.58)
Like-minded donor economic aid	0.00877
	(0.19)
US trade	0.0651
	(0.79)
G7-1 trade	0.0493
	(0.33)
World trade	-0.345
	(-1.63)
UNSC non-permanent member	-0.538**
	(-2.50)
World Bank Executive Director	-0.447**
	(-2.52)

Observations 1607
z statistics in parentheses; * p<.1, ** p<.05

Table 3: Descriptive Statistics for Performance Equation

Variable	Mean	StDev	Min	Max	Description
Outcome	0.7253		0	1	IEG overall project performance rating
Preparation	2.171	0.8996	0.7309	7.464	Estimated duration of World Bank project preparation in years
Project Size	4.326	1.362	-0.6488	8.883	Log of Total Project Cost in constant 2005 $ millions
Population	16.95	1.924	10.6	20.98	Log of population*
GDP per capita	7.849	0.8214	5.968	9.731	PPP GDP per capita in chained 2000 $*
GDP growth	0.02655	0.0781	-0.3074	2.747	Real growth rate of GDP per capita*

4147 observations

*Time dependent variables measured at project approval.

Table 4: Impact of Preparation Duration on Performance

	(1)	(2)	(3)	(4)	(5)	(6)
			Dependent Variable: Outcome			
Preparation		0.213**	0.281**	0.295**	0.226**	0.217**
		(3.62)	(4.43)	(4.66)	(2.80)	(2.31)
Project Size	0.0514	0.0489	0.0673*	0.0727**	0.0394	0.0415
	(1.50)	(1.43)	(1.95)	(2.11)	(1.09)	(1.13)
Population	0.0566**	0.0534**	0.0148	-0.0185	-1.265	-1.880*
	(2.39)	(2.25)	(0.54)	(-0.64)	(-1.47)	(-1.78)
GDP per capita	0.320**	0.381**	0.0733	0.0950	-0.698**	-0.801**
	(6.98)	(7.76)	(1.02)	(1.32)	(-2.09)	(-2.06)
GDP growth	3.513**	3.454**	2.705**	2.288**	0.776	0.649
	(5.17)	(5.07)	(3.89)	(3.31)	(1.11)	(0.86)
East Asia Pacific			0.169	0.00216		
			(1.05)	(0.01)		
Sub-Saharan Arica			-0.745**	-0.725**		
			(-4.89)	(-4.75)		
South Asia			-0.452**	-0.344**		
			(-2.63)	(-1.97)		
Latin America and Caribbean			0.153	0.166		
			(1.16)	(1.26)		
Middle East and North Africa			-0.538**	-0.534**		
			(-3.30)	(-3.27)		
China				1.028**		
				(3.50)		
Observations	4147	4147	4147	4147	4016	3919
z statistics in parentheses; * p<.1, ** p<.05						
Conditional logit estimation with:						
Year dummies	✔	✔	✔	✔	✔	✔
Country FE					✔	
Government FE						✔

Time dependent variables measured at project approval.

Table 5: Alternate Specifications

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Outcome					
Preparation	0.217**	0.211**	0.212**	0.219**	0.220**	0.210**
	(2.31)	(2.20)	(2.20)	(2.32)	(2.34)	(2.22)
Project Size	0.0399	0.0371	0.0265	0.0502	0.0170	0.0279
	(1.08)	(0.97)	(0.69)	(1.35)	(0.44)	(0.74)
Population	-1.882*	-2.148**	-2.179**	-1.904*	-1.867*	-1.801*
	(-1.78)	(-1.98)	(-2.01)	(-1.80)	(-1.77)	(-1.70)
GDP per capita	-0.793**	-0.819**	-0.792**	-0.737*	-0.770**	-0.856**
	(-2.04)	(-2.08)	(-2.01)	(-1.89)	(-1.98)	(-2.18)
GDP growth	0.654	0.731	0.782	0.622	0.686	0.620
	(0.87)	(0.96)	(1.01)	(0.83)	(0.90)	(0.83)
Audit Rating	0.0421		0.517**			
	(0.45)		(3.82)			
Evaluation Lag		-0.115**	-0.215**			
		(-4.28)	(-5.72)			
IDA				0.327*		
				(1.70)		
SAL					0.229**	
					(1.98)	
Energy/Mining Sector						-0.348**
						(-2.11)
Transportation Sector						0.554**
						(3.16)
Other Sectors						-0.117
						(-0.99)
Observations	3919	3788	3788	3919	3919	3919

z statistics in parentheses; * p<.1, ** p<.05

Conditional logit with year dummies & government fixed effects

Time dependent variables measured at project approval. Omitted sector is Agriculture.

Table 6: Economic Vulnerability

	(1)	(2)	(3)	(4)
	Dependent Variable: Outcome			
Preparation	0.240**	0.0578	0.232**	0.111
	(2.51)	(0.50)	(2.42)	(0.97)
Project Size	0.0333	0.0322	0.0361	0.0356
	(0.89)	(0.86)	(0.96)	(0.95)
Population	-1.896*	-2.373**	-2.231**	-2.584**
	(-1.72)	(-2.12)	(-2.01)	(-2.29)
GDP per capita	-0.863**	-0.854**	-0.919**	-0.985**
	(-2.18)	(-2.15)	(-2.28)	(-2.43)
GDP growth	1.250	1.279	1.362	1.330
	(1.34)	(1.37)	(1.42)	(1.39)
% Short Term Debt	0.00794	-0.0326*		
	(0.93)	(-1.94)		
× *Prep*		0.0199**		
		(2.71)		
Debt Service/GNI			-0.00757	-0.0605**
			(-0.77)	(-2.08)
× *Prep*				0.0200*
				(1.86)
Observations	3761	3761	3726	3726

z statistics in parentheses; * p<.1, ** p<.05
Conditional logit with year dummies & government fixed effects
Time dependent variables measured at project approval.

Figure 1: World Bank Preparation SFM

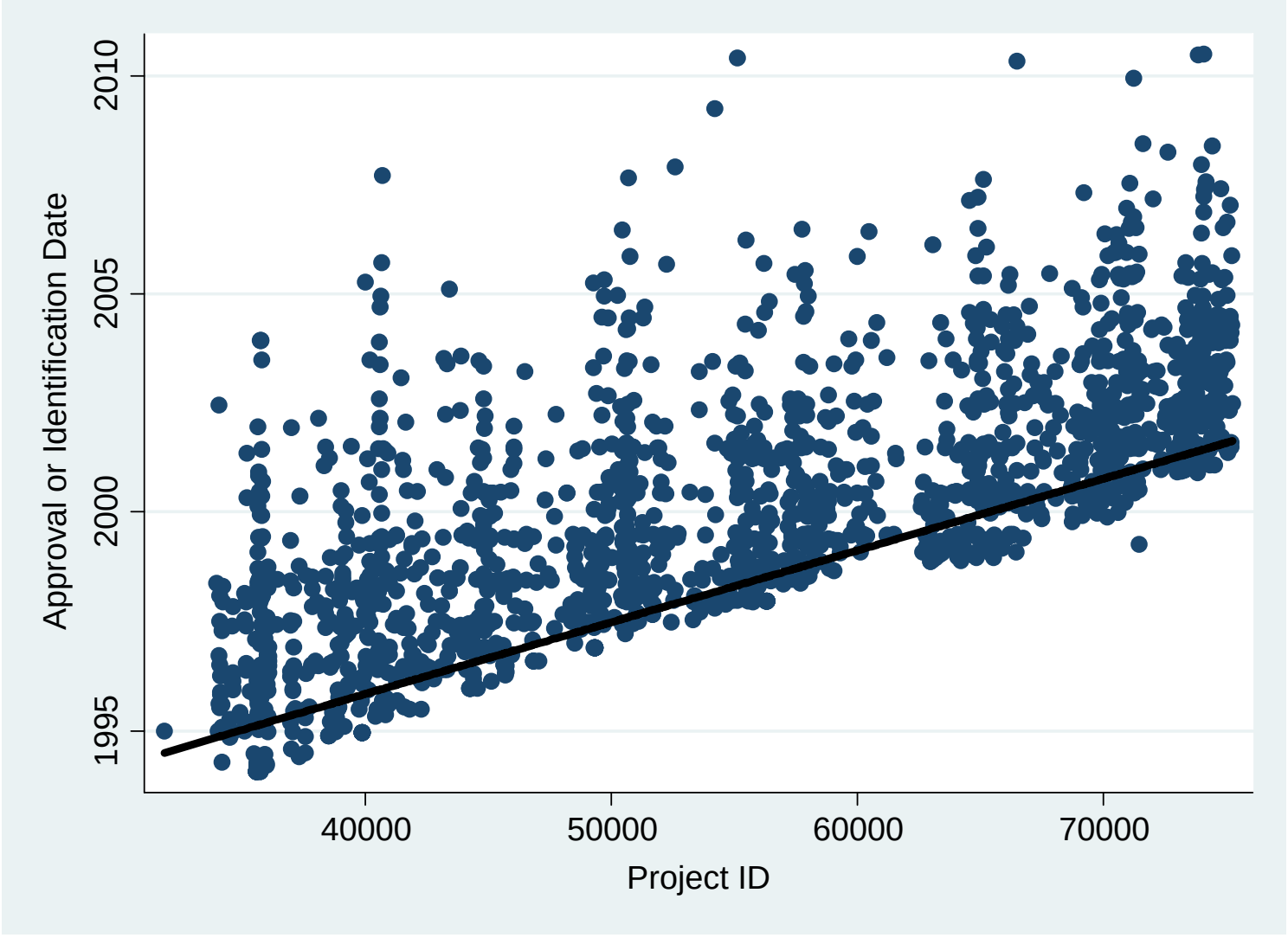
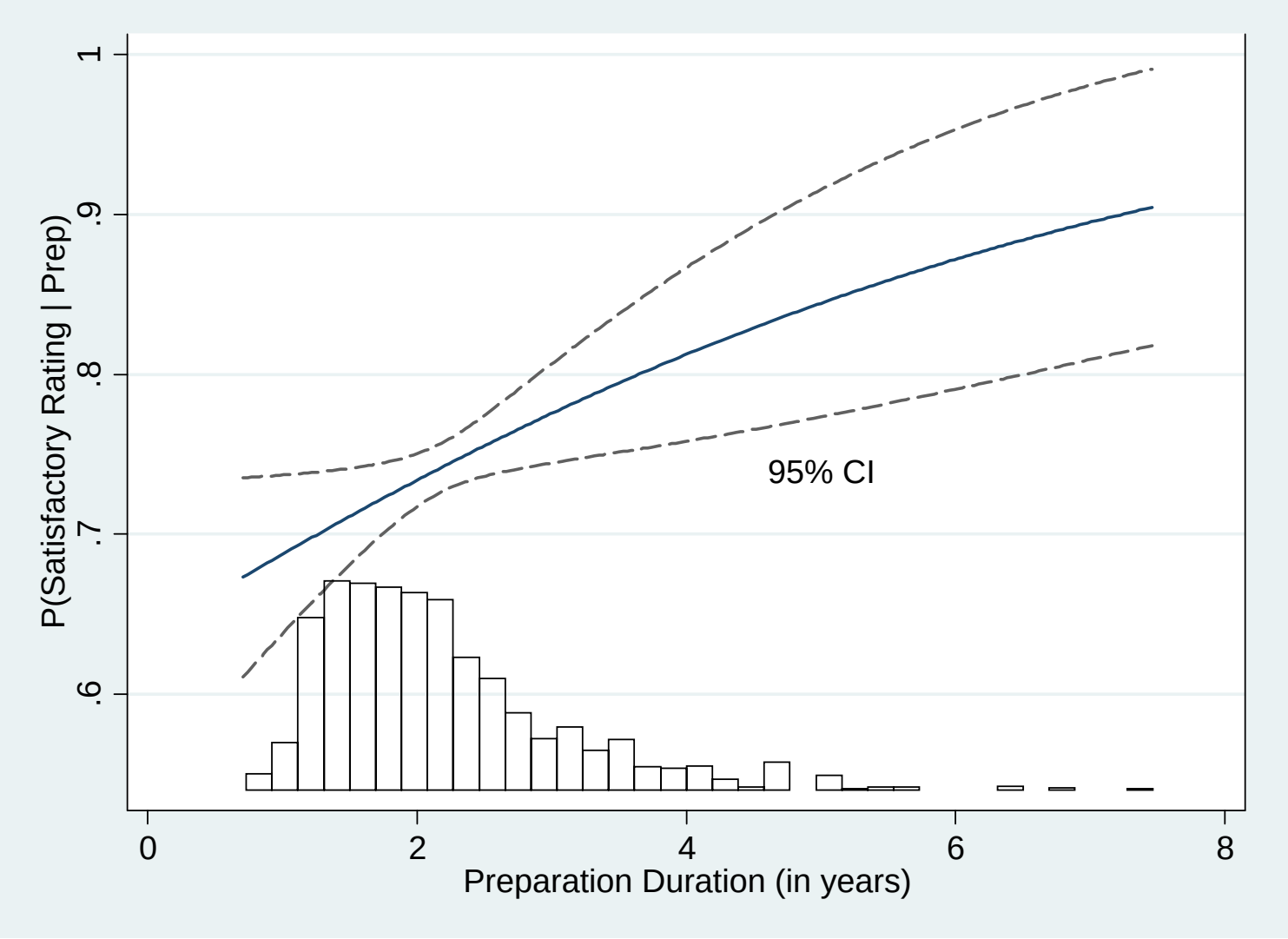
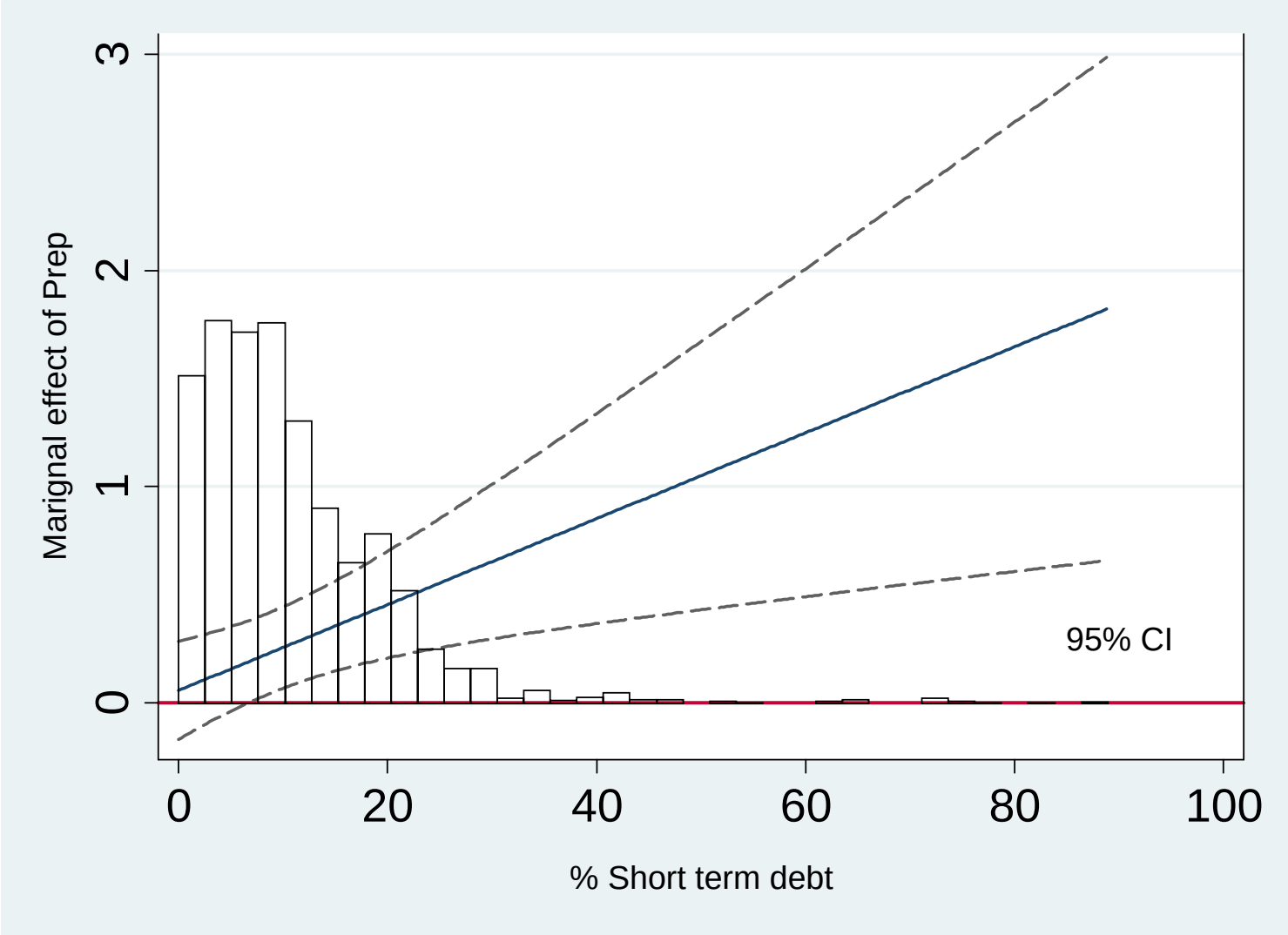


Figure 2: Impact of Preparation on Outcome



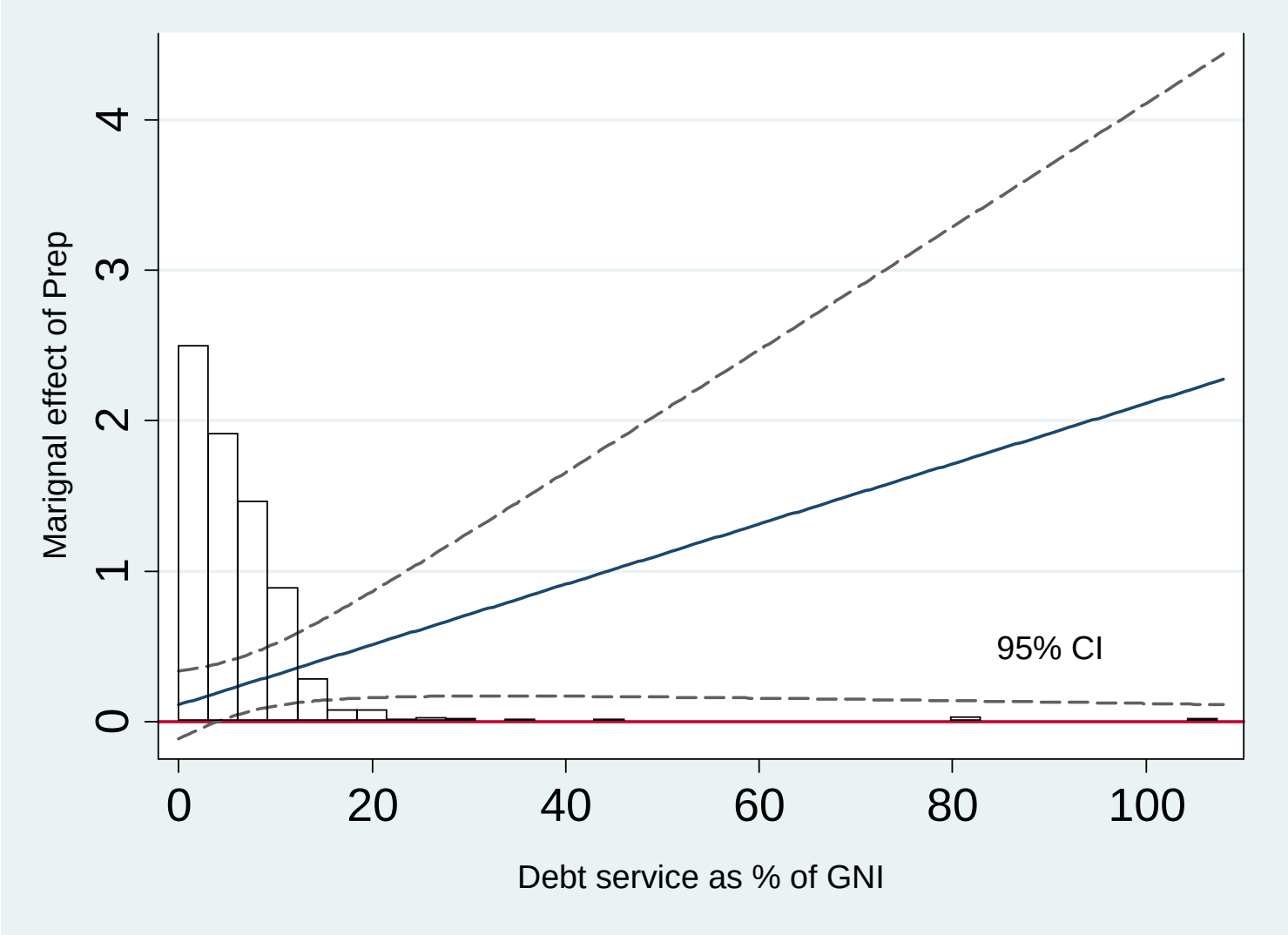
Histogram indicates distribution of variable

Figure 3: Marginal Effect of Preparation Conditional on Short Term Debt (with government fixed effects)



Histogram indicates distribution of conditioning variable

Figure 4: Marginal Effect of Preparation Conditional on Debt Service (with government fixed effects)



Histogram indicates distribution of conditioning variable

Data Appendix

Variable	Data sources
% Short Term Debt	World Bank (2009)
Approval Date	World Bank (2010)
Audit Rating	IEG (2011c)
Debt Service/GNI	World Bank (2009)
Democracy	Cheibub *et al*. (2010)
Energy/Mining Sector	IEG (2011c)
Evaluation Lag	IEG (2011c)
Freedom House Index	Freedom House (2009)
G7-1 economic aid	OECD Development Cooperation Directorate (2006-2009)
G7-1 important votes	U.S. State Department (1984-2010), Voeten and Merdzanovic (2009)
G7-1 other votes	U.S. State Department (1984-2010), Voeten and Merdzanovic (2009)
G7-1 trade	International Monetary Fund (2009)
GDP growth	Heston *et al*. (2002, 2006), World Bank (2009)
GDP per capita	Heston *et al*. (2002, 2006), World Bank (2009)
IDA	World Bank (2010)
Like-minded donor economic aid	OECD Development Cooperation Directorate (2006-2009)
Loan Amount	World Bank (2010)
Other Sectors	IEG (2011c)
Population	Heston *et al*. (2002, 2006), World Bank (2009)
Project ID	World Bank (2010)
Project Size	IEG (2011c)
SAL	World Bank (2010)
Transportation Sector	IEG (2011c)
UNSC non-permanent member	United Nations (2010)
US important votes	U.S. State Department (1984-2010)
US trade	International Monetary Fund (2009)
US military aid	USAID (2009)
US other votes	Voeten and Merdzanovic (2009)
US economic aid	OECD Development Cooperation Directorate (2006-2009)
War	Gleditsch *et al*. (2002)
World Bank Executive Director	Kaja and Werker (2010)
World trade	International Monetary Fund (2009)