The Quality of Foreign Aid:

Country Selectivity or Donors Incentives?

Waly Wane

Development Research Group, World Bank, Room MC 3 – 537, 1818 H Street, NW, Washington D.C. 20433 USA. Tel: +1 202 458 0155; Fax: +1 202 522 1154; e-mail: wwane@worldbank.org. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author. They do not necessarily represent the views of the World Bank, its Executive Directors, or the countries they represent.

Abstract

This paper investigates the determinants of foreign aid quality. It shows that design effects are a crucial component of quality. It thus establishes that donors have an impact on the quality of the foreign assistance they provide. The paper also shows both theoretically and empirically that the quality of aid is endogenous to the relationship between the donor agency and the recipient government. Highly capable and accountable governments accept only well-designed projects, whereas governments with low accountability may accept poor quality projects either because they unable to assess the worth of the projects or they will benefit personally.

JEL Classification: F35; O19; O22; D21;
Keywords: Aid; Quality; Performance; Selectivity; Accountability

The objective of this paper is to show that the design of foreign aid is a key determinant of its quality and hence effectiveness. Indeed, aid design is shown to be as important as governance or economic management in determining its effectiveness. The paper also shows that the quality of aid is endogenous to recipient country's characteristics and the incentive system that prevails within the aid agency. In its theoretical part, the paper addresses the impact of the agency's internal incentive system on the effort its staff members devote to the design of aid projects depending on the recipient country's type. More able recipient countries receive better aid packages, because they have the capacity to screen and select projects conducive to development. Capacity is, however, not sufficient; governments also have to be accountable "enough" to their citizenry to be deterred from accepting bad aid packages.

The quality of foreign aid has, until very recently, been largely overlooked. The literature has relied almost exclusively on the volume of aid to assess its impact. Early attempts to measure the quality of aid include Mosley (1985), followed by White and Woestman (1994), who consider the quality of aid along four dimensions (1) its volume; (2) its terms and conditions; (3) the extent to which it is tied; and (4) its geographical allocation. More recent quality measures, while following the same spirit, have become more sophisticated with respect to the fourth dimension. The relevant criterion for allocating aid has evolved from equity to selectivity. Indeed, the early literature stressed that aid should be given in priority to countries that are most in need, i.e. the poorest, where aid supposedly could have the highest impact, see for instance Rao (1994). Trying to assess the effectiveness of aid, through its impact on GDP growth or poverty, the literature finds that aid have had no impact on average. Aid has not been found to spur growth or increase the quality of life.[1] This has been at the origin of donor fatigue and falling aid commitments from the donor's community who expressed concern about the quality of the assistance it provides to developing countries.[2]

An influential paper by Burnside and Dollar (2000) shows that aid works if it is targeted to appropriate countries, and a country is an appropriate candidate for the receipt of aid if it conducts good economic policies.[3] This important result has promoted the rise of selectivity. For aid to be most effective, it should target countries best able to translate a dollar of aid into growth or poverty reduction. As shown by Collier and Dollar (2002), being selective can produce large gains.[4] Indeed, according to these authors the impact of aid on reducing poverty could have been doubled had the available aid been used in a poverty-efficient manner.

As noted by Roodman (2003), although there is a general consensus that donors should be selective, what they should select for is still open to debate. Burnside and Dollar (2000) have established an index of "economic" selectivity. Boone (1996) shows that the impact of aid does depend on the political regime of the recipient country. Svensson (1999) and Kosack (2003) also show that politics is an important ingredient for how efficiently a country uses aid to generate improved development indicators. A country is therefore an appropriate candidate for foreign

---

[1] Hansen and Tarp (2000, 2001) are the rare exceptions to that rule. They show that aid *unconditionally* has a positive impact on growth rates.

[2] See Statement of the Development Committee of the World Bank and the International Monetary Fund during their sixty-sixth meeting in Washington D.C. on September 28, 2002.

[3] Easterly, et al. (2003) show that the Burnside and Dollar (2000) results are not robust to the addition of new data. Indeed, the significant positive statistical correlation between aid and growth in a sound policy environment falls apart when either more countries are added, or a longer time period is considered.

[4] Llavador and Roemer (2001) also show that aid should be allocated in a manner that rewards governments that make efforts to have a well managed economy, taking into account the country's 'circumstances' to equalise opportunities.

assistance if it implements sound economic policies and is democratic, or has good institutions. These results derived using a macro framework have been confirmed by micro studies. Indeed, Kaufmann and Wang (1995) and Isham and Kaufmann (1999) show, using World Bank-funded investment projects, that the probability of success and the economic rate of return (ERR) of projects are significantly lower in very distorted economies. The recently developed aid-quality indices, such as Easterly (2002) and Roodman (2003), mirror these results and have selectivity as their central element.

The quality measure used in this paper differs sharply from aid-quality indices developed in the literature. To understand the difference it is helpful to break up the aid delivery chain into the upstream end where donors operate and downstream end populated by recipient countries. The quality measures thus far devised, put all the emphasis on *who* gets aid, and are thus driven by the downstream actors characteristics. This paper investigates whether *how* aid is delivered matters and therefore considers aid design as a paramount element of quality. It shifts the focus to the donors in the process.

1. Measuring Donor's Impact

Among the shortcomings of the aid effectiveness literature is its treatment of aid as a pure income transfer, in which donors write a check to the recipient country to relax the country's budget constraint. Aid, however, is delivered as structural programs, sectoral projects, and budget support to name a few. The latter form of assistance is the closest to a pure income transfer. All other forms of assistance are mostly donor-packaged from conceptualisation to implementation. Foreign assistance is but the collection of projects and programs offered by different multilateral aid agencies or bilateral donors and is certainly not a pure income transfer. In effect, some analysts argue that a dollar of aid has a higher value than a domestic dollar because aid comes bundled with superior knowledge. The literature has thus been cognizant of a possible design effect, although, to the best of our knowledge, this paper offers the first attempt for its estimation. If there exists a design effect, the mere aggregation of the amounts disbursed under the different projects and programs leads to a biased estimate of the impact of aid on growth. In the same vein, donor performance ranking implicitly assumes that every dollar of aid has the same value. One can, however, legitimately assume that donors are not equally effective at designing of their programs. Even programs of the same donor are likely to display differences in quality.[5]

Several attempts have been made to measure whether donors have an impact on the outcome of their programs through variables under their control. Dollar and Svensson (2000) (hereafter DS) analyse World Bank–funded structural adjustment programs and show that once the variables under the control of the donor are properly instrumented for, they have no impact on the program's probability of success. Only the political variables of the recipient country matter. However, Deininger, et al. (1998) and Kilby (2000) also study World Bank–funded projects, and both convincingly demonstrate that donor efforts can improve the quality of aid. Deininger, et al. (1998) show that the stock of prior analytical work improves the quality of the World Bank lending portfolio, whereas Kilby (2000) suggests that the timing and intensity of supervision are important donor variables that may yield a positive effect on the project's probability of success. Thus, there are several strings the donor can pull to improve the quality of its aid. This paper also studies the effectiveness of World Bank–funded projects. The focus is on World Bank projects

[5] The quality indices implicitly consider aid to be donor-neutral. Indeed, two donors that give the same volume of aid to identical recipients would score the same in terms of quality irrespective of how well their projects fare.

because of the unavailability of the data on other donors', both bilateral and multilateral, projects and programs. The World Bank will thus be used as a proxy for the community of donors.

Most models of donor behaviour consider the donor as a 'black box' unitary entity. All the staff in the donor's agency are identical and have preferences perfectly aligned with that of the donor's. The literature vastly considers donors to be altruistic, namely that they care only about poverty reduction or the development impact of the projects they implement. For instance Svensson (2000), and Azam and Laffont (2003)[6] consider a donor that cares about the consumption of the poor in the recipient country. The donor impacts the welfare of the poor by offering to the recipient government an aid contract contingent on raising the consumption level of the poor. For Adams and O'Connell (1999) also donors act in the public interest of the recipient country. In the few instances where the incentives of staff in aid agencies are addressed, the literature again views them as pure altruists. For instance, Deininger, et al. (1998) consider a country manager who wants to maximise the quality of the lending portfolio and its policy impact by choosing the optimal budget allocation between the volume of lending and analytical work. Their project manager also maximises the quality of her project by optimally allocating her time between preparation and supervision.[7] In Svensson (2003), although the project manager might have strong incentives to 'spend the budget' as in Mosley (1996), the objective of the manager is still to reduce poverty in the recipient country.

The assumption that donors are purely benevolent and pursue poverty reduction in their client countries is arguable, even more so is the assumption that all staff within the donor's agency share the same drive for poverty reduction. This paper differs from most of the aid literature by assuming that staffers of an aid agency maximise their own welfare, which is independent from poverty reduction or the developmental impact of the projects and programs they design. In a moral hazard framework, it will be shown that the quality of aid is endogenous to the relationship between the donor and the recipient country. Indeed, the quality of projects will depend not only on the amount of effort the staff in the donor's agency devotes to its design, but also on the "type" of government in the recipient country, in a sense that will be made clear later in the analysis. Promotion and staff career concerns will be at the core of the incentive system in the donor's agency. Two simple incentive systems and their impact on the quality of the projects will be studied in the theoretical part of the paper. The first one assumes that a staff member is promoted whenever she manages to get a project accepted, irrespective of the project's performance. In the second incentive system, the project's outcome matters and promotion happens only if the project is successful.

The remainder of the paper is as follows. Section 2 describes the data and provides summary statistics. Section 3 introduces the basic theoretical model. Section 4 offers a resolution of the model and presents the equilibrium that prevails depending on the incentive regime. Section 5 sets out the econometric model and the estimation strategy. Section 6 concludes.

[6] Azam and Laffont (2003) consider multilateral aid agencies (MAI) as non-benevolent intermediaries between the bilateral donor and its recipient country. The MAI has a superior information on the country's willingness to redistribute to poor people (its type) and "sells" this information to the bilateral donor.

[7] The country manager oversees the country's project portfolio, which is a collection of projects each managed by one staff under the authority of the country manager. The lending volume is simply the sum of the sizes (in dollars) of all the projects. Analytical work is a 'local' public good that benefits all project managers by providing them with valuable information on the country circumstances, which can improve the quality of all projects in the portfolio.

2. Data and Descriptive Statistics

As pointed out in Wapenhans (1992), project performance is a function of international, country, and project variables, which can be closely intertwined. This paper focuses on country and project factors. The data used for the analysis is drawn from two sources. For country-level variables, the paper uses indicators developed by Kaufmann, et al. (1999) and updated by Kaufmann, et al. (2003). These indicators, widely known as KKZ variables, are (a) voice and accountability, (b) political stability, (c) government effectiveness, (d) regulatory quality, (e) rule of law, and (f) control of corruption. For project-level variables ratings attributed by the independent Operations Evaluation Department (OED) of the World Bank are used. Each year OED assesses a number of projects completed two to three years earlier for their development impact, likelihood of sustainability, and so forth. OED's projects database also includes the size of the loan, and effort levels by donor and borrower at various stages of the project life cycle.

Projects closed since 1990 have been rated. Although OED ratings may be considered biased, several authors inside and outside the Bank, among others Deininger, et al. (1998), DS (2000), and Kilby (2000), deem them to be reliable measures of a project's success or failure. The performance of the World Bank during the supervision and implementation of the project is rated. One of the most important project ratings is quality at entry, which is also our measure of a project's quality. Quality at entry is intended to capture the suitability of the project's design to the country. It rates (1) the project concept, and (2) the realism of the objectives and approach used during identification and preparation. The client country's performance is rated according to the country's effort during the preparation and implementation stages of the project, and the country's compliance with its commitments and the project's covenants. All ratings run from Highly Satisfactory to Highly Unsatisfactory on a four-level ladder. The overall performance or outcome of the project after completion is also rated on the same scale. The KKZ variables have estimates for four periods 1996, 1998, 2000, and 2002. The score attributed to a country in this paper is a simple average of all four estimates. Other variables of interest collected on the project are its size in millions of current dollars, the amount of resources spent on preparation and supervision,[8] the sector in which the project has been undertaken, whether it is a concessional loan and so forth. From an original population of 3,179 projects, only 1,749 have complete ratings, of which 52 other projects display a size of zero and have been excluded. This leaves us with 1,697 projects for our sample.

Table 1 shows the KKZ governance indicators by region. KKZ indicators, by construction, range from –2.5 to 2.5, higher scores mean better governance. Clearly, Africa and South Asia have the worse environments with respect to governance. African countries display by far the worse environments, be it in terms of government's capacity, regulatory quality, or control of corruption. Africa falls far short of the average for all indicators. Compared to Africa, Latin American countries are on the opposite side of the governance spectrum. They have better regulatory environments and the most efficient and accountable governments.

Labelling satisfactory and highly satisfactory projects as successful, table 2 presents the environment in which projects succeed or fail along with few characteristics of these projects. Table 2 is reminiscent of DS (2000, table 2). Clearly, unfavourable political and economic

[8] Many papers consider the number of staff weeks as the relevant measure for preparation and supervision. However, the value of one staff week varies depending on the profile of the staff (experience, grade, and so forth). The time input of each staff is assigned a "price" depending on her profile. The resources on preparation and supervision reflect the value of the aggregate time devoted by the staff in the project, taking into account price differences. This measure is used in the paper instead of the number of weeks.

environments seem to jeopardize project's chances of success. Indeed, all six governance indicators are significantly lower for failed projects whereas all projects seem very similar otherwise. Most of the results in DS (2000) are confirmed here. For instance, with regard to the variables under the Bank's control, a project's size does not seem to be a good predictor of failure or success. The Bank also invests the same amount of resources for project preparation regardless of their subsequent result. The paper also finds that failed projects garner more supervision, this is quite certainly endogenous as observed by Deininger, et al. (1998) and DS (2000). Indeed, staff invests more resources to salvage bad projects, in vain.

Looking at tables 1 and 2 one would expect projects to fail more often in Africa, EAP, MENA, and SAR, and to succeed in LAC and ECA. Table 3, which also provides summary statistics on the projects by region, broadly confirms these expectations. Only EAP's projects fare better than expected. More than 38% of the projects fail in Africa, contrasting with a mere 19% in LAC. The odds for a project to succeed are 73.4%, but this probability falls to 61.5% for Africa and 68% for MENA, but reaches 80.9% for LAC projects. African projects are quite small, less than half the average project size. African countries are also usually small economies, so the project size may reflect the size of the economy. Moreover, size does not seem to matter for project's success as suggested by table 2. Gathering all this evidence, it seems that aid in fact cannot be effective unless the environment is economically and politically sound. These prima facie facts do indeed call for more selectivity in aid allocation.

How does project design fit in this picture? The lower panel of table 3 gives the distribution of quality at entry within regions, and table 4 presents the joint distribution of quality at entry and performance. Strikingly enough, projects' quality at entry and performance seem to have similar regional distributions. "Bad" environments have a higher rate of project failure relative to "good" environments. However, they also receive a higher percentage of poorly designed projects. Moreover, the correlation between quality at entry and performance is quite high, at around 0.6. Indeed, more than 72% (1,237) of the projects confirm, after implementation, the prediction OED gave on their rating at entry. Only about 18% of the projects rated in one category at time of entry (in grey) end up with a different rating after the project is completed (in white). Quality at entry therefore seems to be a powerful predictor of the project's success or failure. Projects are implemented in different environments and one might accept that they perform differently because of the differences in country environments. The explanation for the parallel regional variation in design quality is less obvious, however, because a single donor is responsible for the design across all regions. Because donors might have an impact on the quality of their aid, through the design effect, this calls for caution in the use of selectivity as exclusively based on recipient's characteristics.

3. Theoretical Model

This section proposes a simple theoretical framework that rationalises the fact that aid only works in good environments. It is claimed here that this outcome is endogenous to the type of the borrower and the incentive system the project's designer faces. The two main actors in the model are the staff in the international aid agency, which designs the project, and the borrowing country's government. It is a simple two-period game, at period $t = 0$, the staff designs the project and the client country accepts or rejects the proposal. If the project is refused the staff is not promoted and the country does not have any project. If the project is accepted, at $t = 1$ the performance of the project is observed and the staff's promotion decision is taken, depending on the incentive regime that prevails. The government and the population assess the state of the economy irrespective of whether a project is implemented and elections take place.

The Aid Agency proposes and finances development projects and programs in a number of developing countries. The aid agency's mission is to help countries escape from their poverty trap, or put them on a development growth path. To accomplish this mission the aid agency hires its staff[9] and sets up its internal incentive system within which the staff operates. An incentive system determines the promotion criteria for the staff based on their (relative) performance. Only two types of incentive systems will be considered here: (a) promotion occurs if the staff member manages to sell a project to a borrower country; and (b) promotion occurs only if the sale of a project to a borrower country also results in a successful project. Objective performance measures are readily available for both incentive systems. The mere observation of an implemented project is sufficient for promotion under the first incentive system. For the second system, international aid agencies have developed sophisticated measures of project success or failure to evaluate performance.

The staff of the aid agency is assumed to maximise their welfare and take the internal incentive system as given. This assumption is in contrast with most of the foreign aid literature. As pointed out by Deininger, et al. (1998) the staff in international aid agencies are heavily involved in the design and conceptualisation of projects and programs. Projects are designed by the staff that then negotiate with the borrower country's government that may take or leave the offer. The promotion of the staff depends on the internal incentive system, and means a higher wage in the following period. The welfare of the staff is a function of wage and effort exerted for project design. Each staff is also characterised by an ability parameter known only to them which influences the probability of a successful project. Let β the ability of the staff be distributed on the compact support $[\beta_L, \beta_H]$ according to the density *f(β)* and the c.d.f. *F(β)*. A staff maximisation program can be written as

$$\max_e u(s) + \delta \cdot \left[u(S) \cdot \Pr(promotion \mid \beta, e) + u(s) \cdot (1 - \Pr(promotion \mid \beta, e))\right] - \psi(e) \quad (1)$$

where *u( )* is a concave utility function, *s* is the entry salary, *S (>>s)* is the salary if and when promoted, δ is the discount rate, *e* is the effort level which also satisfies $0 \le e \le 1$, and $\psi(\cdot)$ is a convex function representing the disutility of exerting effort. Consistent with the career concern literature, the staff does not exert any effort in the second and last period of the game. The probability of being promoted depends, in part, on the incentive system in place in the international aid agency. The salary increase following a promotion is assumed to be high enough for the staff to be willing to exert maximum effort if it secures promotion. This is satisfied if the following condition holds:

$$u(s) + \delta \cdot u(S) - \psi(1) > 0. \quad (A1)$$

For the sake of simplicity, all projects are the same regarding their cost *C* and potential "full" benefit *B*. The costs C[10] are to be repaid by the borrowing member country as soon as the project is accepted. On the other hand the benefit that accrues to the country will depend on the

[9] It is assumed that the aid agency has the resources to (and will) hire highly qualified individuals who are competitive on the international labor market. In its recruitment process the agency aims at identifying both highly skilled and dedicated individuals. The hiring process is imperfect, however, so staff will vary in terms of ability and degree of dedication to development.

[10] The international financial institution actually lends an amount *C/(1+r)* to the country where *r* is the borrowing interest rate the countries face. We are concerned here with neither the optimal reward system of the agency nor the optimal size of its loans and interest rates. These questions are left for future research.

design of the project. The country reaps a portion $\pi + \beta \cdot e$ of the potential full benefit B, where π is sheer luck and commonly known, and *e* is the costly effort the staff puts into the design of the project. The expected return of a project designed by a staff of talent β and who devotes effort *e* to the project is simply:

$$\text{Expected Return} \mid \beta, e = (\pi + \beta \cdot e)B - C.$$ [11]

For a similar level of effort, staff with higher abilities bring a higher rate of return. The ability of the staff and the effort they provide multiplicatively determine the value added of the staff into the project. In this model, unlike the additive one, staff ability is valuable only insofar as they put effort into the project. Indeed, staff who make no effort in the design of their project have no value added irrespective of how talented they are. To quote Dewatripont, et al. (1999), "talent matters little if the [staff] shirks, but makes an important difference if the [staff] 'tries to make things happen.'"[12] To avoid a degenerate problem, the following assumptions are necessary:

$$\pi \cdot B < C \text{ and } \beta_L < \frac{C - \pi \cdot B}{B} < \beta_H . \qquad \text{(A2)}$$

The first part of assumption A2 simply states that for a project to be beneficial, the staff has to put a minimum effort into its design. In other words, projects with no staff input have a negative return, and only increase the country's level of indebtedness. The second part of A2 tells us that staff at the bottom of the talent distribution cannot design a successful project no matter how much effort they make. However, a portion of the staff have the ability to design successful projects. That marginal staff who by exerting maximum effort can generate a zero-return project is $\beta^* = C/B - \pi,$ all staff with higher ability can, if they put the required effort, design good projects.

The government of the client country borrows the funds and implements the project. The government is characterised both by its *capacity* (ability to screen proposed projects) and its level of *accountability* (the willingness of the government to accept projects likely to improve the welfare of the population).

The government can be of either high (c_H) or low (c_L) capacity.[13] A high capacity government is able to exactly pinpoint the expected economic return of a proposed project, whereas for a low capacity government projects are indistinguishable. In the jargon of Sah and Stiglitz (1986) the government's capacity is its screening function. A high capacity government is therefore a perfect screener, whereas a low capacity government has no discriminating capability. Although it can assess the worth of any project, the high capacity government knows neither the ability of the staff nor is it able to determine the amount of effort invested in the design of the project. Low capacity governments on the other hand rely on "trust" to judge the project. Trust can be parameterised by the (minimum) proportion of the project's full benefit, π^*, the

[11] An alternative and equivalent specification would be to consider the staff's effort to influence the probability of success of the project, which is $P(success \mid \beta, e) = \pi + \beta \cdot e$ for a project that yields B in case of success and zero when it fails. The country would then consider the project's expected net benefit.
[12] See Dewatripont, Jewitt, and Tirole (1999).
[13] Screening capacity could have been represented along a continuum without changing the basic result of the paper, but at the expense of analytical complications.

government believes it will get when dealing with the aid agency. Therefore the government expects the project to have a return $\pi^* B - C$. The trust parameter may depend on the history of the country and aid agency relationship, and/or the aid agency's reputation. It is supposed to be known to all parties for the sake of simplicity.

Following Ferejohn (1986), we assume that the government is judged by its current performance and is re-elected if the welfare of the population is greater than *W*, the minimum (increase in) level of welfare the population. This threshold defines the level of accountability of the government to its population, and is distributed on the support $W \in [\underline{W}, \overline{W}]$.[14] For instance a negative threshold would mean that even if the welfare of the population deteriorates, the government will safely be re-elected. For dictatorial regimes, *W* certainly tends to negative infinity. A high and positive threshold signals a very demanding population that wants good performers and is able to exercise its voice and vote non-performing administrations out of office. In case the government accepts the project, the (increase in) population's welfare is represented by the expected return of the projects. Should the government refuse the project, the state of the economy is fully determined by a random shock which follows a normal distribution with zero mean. Therefore the government is sure to be reinstated if the projects it accepts bring forth an outcome greater than *W*. But when the government refuses a project it is re-elected with probability $1 - \Phi(W)$.

Let us formalise the government's decision-making process when faced with a project. For the sake of simplicity, let us assume that the government always gets a personal gain, *G*, from a financed project. This personal gain is not contingent on the outcome of the project, it is for instance the amount the government can divert from the funds released for the financing of the project. The value of holding office for the government is given by *V*. Therefore the government's instantaneous welfare is given by *V* + *G* if it accepts the project and *V* if it rejects it. The acceptance or refusal of the project, however, affects the probability of re-election and therefore the next period's welfare. The government's anticipated next period utility is *V* if it keeps office and zero otherwise. It also has a discount rate δ. The expected anticipated utility is therefore:

$$\Pr(\text{winning election}) \cdot V = \begin{cases} P(\xi \cdot B - C \geq W) \cdot V & \textit{if} \quad \text{Accepts} \\ (1 - \Phi(W)) \cdot V & \textit{if} \quad \text{Rejects} \end{cases}$$

where ξ is equal to $\pi + \beta \cdot e$ or π^* depending on the government's screening capacity, and $P(\xi \cdot B - C \geq W)$ is a zero-one indicator. The government accepts (refuses) the project if

$$G + V + \delta \cdot P(\xi \cdot B - C \geq W) \cdot V \geq (<)\ V + \delta \cdot (1 - \Phi(W)) \cdot V \qquad (2)$$

Equation (2) governs the decision-making on the project. It defines the participation constraint of the government. Clearly, whenever the project is deemed "good," it has an expected return that permits re-election and the government approves of it. The government will still accept a bad project, $\xi \cdot B - C < W$, if the personal profit from accepting the project now is

[14] See also Seabright (1996) for a similar definition of accountability.

greater than the discounted value of holding office at *t = 1*, due to pure luck after refusing the project at *t = 0*, i.e. if $G > \delta \cdot (1-\Phi(W)) \cdot V$. Less accountable governments would display a higher propensity to accept bad projects. To simplify the analysis, let us assume

$$W^{**} = \Phi^{-1}\left(\frac{\delta \cdot V - G}{\delta \cdot V}\right) \leq \pi \cdot B - C \,. \qquad \text{(A3)}$$

A3 implies that all governments that willingly accept poorly designed projects just for the sake of their personal benefit, face a lack of accountability to the extent that they would have accepted any project anyway. This completes the description of the model.

4. Equilibrium

This section describes the equilibrium depending on the incentive systems that prevails in the aid agency.

4.1 Spend the Budget Incentive System

4.1.1. High Capacity Government

Suppose first that staff gets promoted whenever they sell a project to a borrowing country. Let a staff of ability β be matched with a country of accountability level *W*, which is also a perfect screener. With assumption A3, the government accepts the project if and only if the project's return is equal to, or greater than, the required population welfare level that ensures its re-election. Since it can perfectly assess the return of the project, the latter is accepted if and only if $(\pi + \beta \cdot e) \cdot B - C \geq W$. The government accepts the project only if the value added of the staff's effort, $\beta \cdot e \cdot B \geq W + C - \pi \cdot B$, exceeds the sum of the minimum welfare required by the population and the cost of a poorly designed project.

This equation also determines how much effort a staff needs to make to get a project accepted and secure promotion. Because effort is costly, the staff will provide the minimum input in terms of effort to ensure promotion. The profile of optimal efforts for staff depending on their ability and the government's accountability is[15]:

$$e_H^*(\beta,\, W) = \begin{cases} \dfrac{W + C - \pi \cdot B}{\beta \cdot B} & \text{if} \quad \pi \cdot B - C \leq W \leq W^*(\beta) \\ 0 & \text{Otherwise} \end{cases} \qquad (3)$$

[15] Let define $W^*(\beta)$ and $\beta^*(W)$ the maximum country accountability level for a staff of ability β to sell a project, and the minimum staff ability level for a country *W* to accept a project. They satisfy the equation $(\pi + \beta) \cdot B - C = W$ when solved with respect to β and *W*, respectively. They are defined at the maximum effort level of the staff.

where $W^*(\beta) = (\pi + \beta) \cdot B - C$ is the lowest level of accountability such that the staff of ability β will need to exert the maximum effort to get their project accepted. It is the staff's "Peter Principle" accountability ceiling, the accountability level above which the staff is incompetent to design a project that is satisfactory to the government.[16] If the staff meets a government whose accountability is greater that their Peter Principle ceiling, then even exerting the maximum effort is not enough for the staff to get her project through. In this case, the staff exerts no effort. As equation (3) clearly shows, it is more difficult for staff to sell a project to a country with a more demanding population. If the country is not accountable at all and can get away with a project value less than $\pi \cdot B - C$ the staff does not need to and thus makes no effort, whatever their ability. For more accountable countries, staff will need to exert some effort and the effort level increases linearly with the level of accountability. For countries with a very demanding population, $W > W^*(\beta)$, it is worthless for the staff to exert any effort because their project will be rejected.

From the country's perspective, all accepted projects will have a return that allows the government to exactly fulfil its commitment to reach the minimum welfare level. The country gets a project only if it is matched with staff whose ability is equal to or greater than $\beta^*(W)$ and therefore accepts a project with probability $1 - F(\beta^*(W))$.[17]

4.1.2. Low Capacity Government

Suppose now the government in the client country is not effective enough at project screening. The parameter of interest is then the level of trust it has in its counterpart aid agency, and the decision process is very simple. Indeed, the country will accept the project if and only if the expected return is such that $\pi^* \cdot B - C \geq W$, if the level of trust the government has in the aid agency leads to an excepted return greater than the required minimum welfare level. For a given trust level, the lower the government's accountability, the more likely is the acceptance of the project.

From the perspective of the staff who wants to ensure a promotion, all that matters is whether the trust level is high enough for the country to accept the project. In any event the staff's effort is zero. Indeed, if the government trusts the aid agency and systematically accepts project, there is no need to put any effort in the project since effort is costly and a promotion is guaranteed because the project will be accepted. In the same vein, there is no need to put effort in a project if the trust level is too low and the project is systematically rejected. Therefore, in case of a low capacity government the profile of effort is very simple, $e_L^*(\beta) \equiv 0 \ \forall \ \beta$, the staff always shirks and the country accepts the project depending on its beliefs only. From A2, because the staff shirks all projects have a very poor design and a negative economic return.

The discussion can be summarised in the following proposition:

Proposition 1*: When the incentive system of the international aid agency is such that the promotion of the staff is contingent only on the acceptance of a project by a client country, the following is observed:*

[16] See Peter and Hull (1969) or Fairburn and Malcomson (2001) for more on the Peter Principle.
[17] It is assumed that the matching between staff and countries is exogenous. However, one could easily imagine the aid agency implementing a sorting mechanism to maximise the number of accepted projects.

- ***If the country has a weak screening capacity:*** *the staff never puts effort in the design of the projects; all accepted projects have a negative return; for a given trust level, only countries with low accountability i.e.* $W \leq \pi^* \cdot B - C$ *accept projects.*
- ***If the country has a high screening capacity****: the staff exerts minimum effort to get projects accepted; projects are accepted only if their return is at least equal to the accountability level of the government; low accountability governments will knowingly accept projects with negative returns; the more accountable the country the higher is the return of the accepted project, and the higher the staff's effort.*

The profile of effort the staff exerts is given in figures 1 and 2 for high and low capacity governments, respectively. As shown in the figures, perfect screeners might very well cheer negative return projects because of their lack of accountability coupled with personal gains accruing from the acceptance of a project.

As is obvious from the above proposition, higher return projects are found in high capacity countries that are also more accountable to their population. Those governments refuse projects when not enough effort has been made to make them worthwhile. Low capacity countries, however, accept projects based on the trust level, and all accepted projects are of poor quality and end up being detrimental to the population. Indeed, the staff anticipate the likelihood of a government to accept (or refuse) the project, and have no incentive to put costly effort into the design of the project. This is a theoretical derivation of Burnside and Dollar (2000) result, where aid works in good environments but almost surely fails in bad ones. There is, however, a fundamental difference between the two results. For Burnside and Dollar (2000) all countries receive the same type of money which is subsequently wasted in 'bad' environments. On the contrary, in this paper 'bad' environments receive 'bad' money as a result of their low accountability and lack of capacity to filter aid money, coupled with absence of incentives for aid agency staff to propose 'good' aid in such environments.

4.2. Wisely Spend the Budget Incentive System

As proposition 1 shows, two necessary ingredients have to be present for the country to receive good quality aid projects. First, the government must have the ability to screen and gauge the capacity of projects to bring about economic prosperity. Second, the government has to be accountable "enough" to its citizenry to resist bad projects just for its personal advantage. Whenever one of this ingredient is missing bad projects will be accepted. There is thus a sorting mechanism such that good governments, both politically and economically attract good aid, whereas weak or low-accountability countries attract bad aid. Can a remedy be found to ensure that aid works even in bad environments? Part of the answer lies in the re-design of the incentive systems for staff in international aid agencies.

Let us now turn to the evaluation of the second incentive system. In this system, the promotion of staff is contingent on the performance of the project designed for the client country. Let us suppose that the staff is promoted only if the return to the project is higher than a preset minimum level φ which for equity reasons is the same for all projects irrespective of the country or staff. It is assumed that the return of the project can be readily assessed by the agency.[18]

[18] For the World Bank, a project is satisfactory if its ERR is at least 10%.

4.2.1. High Capacity Government

What is the effort level exerted by staff when matched with a high capacity government. Because the government can assess the probability of success of the project, the staff is promoted if and only if the project satisfies $(\pi + \beta \cdot e) \cdot B - C \geq \text{Max}(\varphi, W)$. The project must pass the minimum standard test. Indeed, the project is not accepted by the country if it does not exceed its re-election threshold and even if it does, the staff is not promoted unless enough effort is exerted to meet the institution's requirement. It is straightforward to show that the effort profile in this incentive system is:

$$e_H^{**}(\beta) = \begin{cases} \dfrac{\varphi + C - \pi \cdot B}{\beta \cdot B} & \text{if} \quad W \leq \varphi \\ e_H^{*}(\beta) & \text{if} \quad W > \varphi \end{cases} \tag{4}$$

if $\beta \geq (\varphi + C - \pi \cdot B)/B = \beta^*(\varphi)$ and $e_H^{**}(\beta) \equiv 0$ for $\beta < \beta^*(\varphi)$. What are the effects introduced by this incentive system with respect to the first one. First of all there is no difference for countries whose accountability threshold is greater than the aid agency's standard. There is, however, a lower probability that a country will have a project because all staff whose ability is lower than $\beta^*(\varphi)$ will put zero effort in the design of their project. Such projects will only be accepted by countries whose accountability level is lower than $\pi \cdot B - C$; moreover, the staff member is not promoted. Such countries accept all projects in both systems, anyway. For countries between $\pi \cdot B - C$ and φ they refuse all projects when they are matched with staff with an ability lower than $\beta^*(\varphi)$. Hence, some projects with positive returns are lost. Indeed, beneficial matches between accountable countries ($W > 0$) and staff with ability $\beta^*(W) < \beta < \beta^*(\varphi)$, cannot materialise in this system. To compensate for this, failures due to non-accountability and bad incentive are avoided. Indeed, all countries such that $\pi \cdot B - C < W < 0$ that would have accepted negative return projects now do not have that option any more, all projects before them fare worse than they can accept due to the lack of design effort from the staff.

All staff with ability levels above the Peter Principle ceiling defined by the aid agency's standard, i.e. $\beta \geq \beta^*(\varphi)$, have the opportunity to be promoted. Whenever such a staff is matched with non accountable governments, their effort will compensate for that lack of accountability to generate beneficial projects. As a matter of fact, when the staff meet a government such that $W \leq \varphi$ then the effort level exerted is such that the agency's standard is just met and the project will have an economic return of φ.

4.2.2. Low Capacity Government

If the staff is matched with a low ability government that has a trust parameter π^* the profile of effort is much simpler and is given by:

$$e_L^{**}(\beta) = \begin{cases} \dfrac{\varphi + C - \pi \cdot B}{\beta \cdot B} & \text{if} \quad \pi^* \geq (W + C)/B \\ e_L^{*}(\beta) & \text{if} \quad \pi^* \leq (W + C)/B \end{cases} \tag{5}$$

if $\beta \geq \beta^*(\varphi)$ and $e_L^{**}(\beta) \equiv 0$ for lower ability staff. For low capacity government, for the same level of trust, things can only get better with respect to the first incentive system. Indeed, suppose they always accept a project because they find the aid agency trustworthy. Then if the country is matched with a staff with ability $\beta < \beta^*(\varphi)$ then it will still have a project with negative returns as before. On the other hand, if the staff's ability is $\beta \geq \beta^*(\varphi)$ the country will have a project that meets the agency's standard. However, because of its lack of capacity to screen projects the country will still not have a project that meets its accountability level when this one is greater than the standard set by the agency.

The following proposition summarises this section:

Proposition 2*: When the incentive system of the international aid agency is such that the promotion of the staff is contingent on both the acceptance of a project by a client country and the requirement of an economic return higher than the agency's standard φ, then the following is observed:*

- ***If the country has a weak screening capacity****: the project is accepted if the country's accountability satisfies $W \leq \pi^* \cdot B - C$; staff with ability β such that $\beta^*(\varphi) \geq \beta$ exert no effort in project's design, and propose negative economic return projects to the country; all staff who can meet the standard put the minimum effort required to exactly match it and propose projects with economic return φ irrespective of the country's level of accountability.*
- ***If the country has a high screening capacity****: projects are accepted only if their return is at least equal to the accountability level of the country; staff with ability β such $\beta^*(\varphi) \geq \beta$ exert no effort in the design of the project and propose negative economic return projects to the country, only governments with accountability lower than $\pi \cdot B - C$ accept projects; staff with ability β such that $\beta^*(W) > \beta \geq \beta^*(\varphi)$ cannot meet the country's standard will not exert any effort and the project will be rejected; staff with ability β such that $\beta \geq Max(\beta^*(\varphi), \beta^*(W))$ will exert the effort to just meet the minimum standard, the effort level is higher the more accountable the country.*

Proposition 2 shows that with this incentive system bad projects may still find their way to weak capacity countries that trust the international agency. However, with respect to the first incentive system in which they got bad projects with certainty, there is now a positive probability that will they get a beneficial project. The probability of accepting bad projects depends both on the distribution of staff ability and the standard of the agency. The more the distribution is skewed to the left, and the higher the standard, the greater the chances for bad projects being accepted. The international agency faces a trade-off in the choice of the optimal standard.

It is the binding side of the incentive system that drives effort. As long as the agency's standard is higher than the country's accountability level to its citizenry, the staff will exert enough effort to just reach the standard. For more demanding countries, it is the standard set by the country that determines the effort the staff has to exert in order for the project to be accepted. In any event, whether the project will attain any standard depends on the ability of the staff to exceed the Peter Principle ceiling defined by the government's accountability. Figure 3, shows the effort profiles under both incentive systems.

5. Econometric Analysis

This section explores the determinants of both the quality of the project design and project performance. It also aims at providing an accurate estimate of the impact of the quality of a project on its probability of success. Let quality and performance be determined by:

$$Q_{ij} = 1(f(X_i^q, Z_j^q) + u_{ij} > 0) = 1(Q_{ij}^* > 0), \text{ and} \quad (6)$$

$$P_{ij} = 1(\beta \cdot Q_{ij} + g(X_i^p, Z_j^p) + \varepsilon_{ij} > 0) = 1(P_{ij}^* > 0) \quad (7)$$

where 1(·) represents the indicator function, and P_{ij}^* and Q_{ij}^*, respectively, denote unobserved latent variables for performance and quality at entry of project *i* in country *j*. The superscripts *p* and *q* refer to performance and quality, respectively, and X_i, and Z_j represent project and country specific factors. The factors specific to the projects may include variables such as size, whether it is an adjustment or investment project, or more sophisticated variables, such as measures effort by both the donor and the recipient country at various stages of the project cycle. Country-specific factors will typically include institutional capacity, political, and economic variables.[19]

A natural continuous variable candidate for measuring performance is the economic rate of return (ERR) of the project. Unfortunately, projects in social sectors do not lend themselves easily to such measurement and have no ERR. They need to be evaluated using some other yardstick. The Operations Evaluation Department has developed performance indicators that proxy those quantitative measures when output is not directly observable. All projects are consistently evaluated with these indicators. Contrary to performance, there is no obvious continuous variable to quantify quality. A methodology for producing quality indicators has also been developed within the World Bank, and these indicators are used here. The analysis uses the indicator variables $P_{ij} = 1$ if the project is deemed to have performed satisfactorily by OED and $P_{ij} = 0$ otherwise. The quality at entry of the project is satisfactory when $Q_{ij} = 1$ and $Q_{ij} = 0$ otherwise.

Because quality and performance are dichotomous variables, equations (6) and (7) will be estimated with the standard probit regression model. The problem with estimating equations (6) and (7) is that quality might be endogenous and the disturbances u_{ij} and ε_{ij} correlated. In this case, estimating the equations separately will produce biased estimates for equation (7). Not only would the impact of quality on the probability of success be biased, but also the coefficients attached to the other covariates may be adversely affected. The alternative is to simultaneously estimate the equations as a system. The appropriate model is then a recursive simultaneous bivariate probit because the quality at entry is present in the performance equation.[20] The errors

[19] Note that because of high correlation (pair-wise coefficient of correlation higher than 0.85) between three of the governance variables, namely government effectiveness, rule of law, and control of corruption, only government effectiveness will be retained in the econometric analysis.

[20] However, as suggested by Angrist (1991) and confirmed by Evans and Schwab (1995), and Evans, Farrely, and Montgomery (1999) in practice the average treatment effects obtained by the bivariate probit are close to the estimates that would result from two-stage least squares models.

are correlated and jointly distributed as a bivariate normal with $E(u_{ij}) = E(\varepsilon_{ij}) = 0$, $Var(u_{ij}) = Var(\varepsilon_{ij}) = 1$, and $Cov(u_{ij}, \varepsilon_{ij}) = \rho$.

For the bivariate probit model to be identified and effective, there must exist either valid exclusion or functional form restrictions, or one should assume equal selection on observed and unobserved variables, an alternative identification strategy suggested by Altonji, et al. (2000). The equal selection rule amounts to imposing a set of restrictions on the correlation coefficient of the errors and heavily hinges on the fact that observed exogenous variables have been randomly chosen from a large pool of candidates and had equal probability of being picked up than unobserved ones. This assumption does not apply well in our setting, for most of the data has been collected for the specific purpose of studying the effectiveness of the projects. On the other hand, functional form restrictions would imply for instance the imposition of *f* and *g* to be linear or polynomial in their arguments,[21] and this is not necessary here as showed by Altonji, et al. (2002). This paper uses exclusion restrictions for identification, it requires that there be at least one exogenous variable that impacts the probability of success only through its effect on quality. This variable enters the quality equation but is excluded from the performance equation (see Maddala 1983, p.122-123), or Bollen, et al. (1995). In other words, the following condition $(X_i^q, Z_j^q) \subset (X_i^p, Z_j^p)$ must be satisfied.

The paper uses two identifiers. The first one is the cumulative number of economic and sector work (ESW) deliveries for a specific country up to the date of entry of the project in the portfolio. This variable represents the stock of knowledge accumulated by the international aid agency on the country's economic and social environments. On purely theoretical grounds the number of ESWs should determine the quality of future projects by providing valuable analytical input for their design. There is no reason to expect it to influence the performance of any specific project during implementation. Deininger, et al. (1998) show that the number of ESWs has a strong positive impact on the quality of World Bank loans. The second identifier, the borrower's quality of preparation, is assessed by OED and quantifies the quality of the client country's involvement in the project at the preparation stage. It can also proxy the degree of the country's ownership and commitment for the project. The identifiers are valid to the extent that (a) they are determinants of the quality at entry of a project, but (b) are not correlated with its performance. The validity of these instruments will be assessed later in the analysis.

5.1 The Single-Equation Probit Models

Before turning to the full model, it is worthwhile to investigate the single equation models for both performance and quality at entry.

5.1.1. The Determinants of Performance

Table 5 reports separate estimates of the probit regressions. The project's quality at entry enters only in regression 4, presented in the last column of the left panel. Notice that voice and accountability and the regulatory quality are significant in none of the regressions that aim at

[21] A common practice is to assume linearity and consider $P_{ij}^* = \alpha_p + \beta \cdot Q_{ij} + X_i^{p\prime}\gamma_p + Z_j^{p\prime}\lambda_p + \varepsilon_{ij}$ and $Q_{ij}^* = \alpha_q + X_i^{q\prime}\gamma_q + Z_j^{q\prime}\lambda_q + u_{ij}$ as the system to be estimated with β being the primary parameter of interest.

explaining performance. This is all the more surprising for the regulatory quality that captures the quality of the economic environment is found highly and positively significant by most studies.

Regression 1 includes only a restricted set of explanatory variables: the governance indicators. Political stability and government effectiveness seem to be relevant indicators in explaining the projects' performance. However, with a pseudo R-square of only 3%, almost all of the variation in project performance remains unexplained. Therefore, there are certainly many variables omitted in this regression and performance can by no means be explained by country characteristics alone. In regression 2 we add loan characteristics and regional dummies as new covariates. Political stability loses its explanatory power with these additions, and the R-square goes now to 16%. In regression 3 the quality of the borrower country's involvement in the project and the supervision effort of the World Bank are further considered. Interestingly, adding these variables allows us to explain about 55% of performance variation. Government effectiveness is still significant, but only at 10%, and a marginal increase in the effectiveness of the government increases the probability of success by 8%. All variables concerning the borrower and agency's involvement are highly and positively significant. Finally, once the quality at entry of the project is controlled for, government effectiveness is no more significant in explaining performance. However, all variables related to the borrower and the World Bank's involvement are quite robust and remain highly significant, although with a slightly reduced impact. African projects are 6% less likely to perform well and the negative impact is strongly significant across all regressions, although the impact is gradually reduced as more covariates are controlled for.

Overall, a project with a satisfactory quality at entry has a huge 17.5% higher probability of performing well than a poorly designed project. None of the governance variables is significant contrary to DS (2000). The most powerful variable in explaining performance is the quality of the borrower's implementation. Projects perform poorly in Africa and adjustment loans fare worse than investment loans. A sizeable variation is explained by this regression. The predictive ability of the model is also fairly good, with an average predicted probability of success for satisfactory projects of 90% versus 28% for failed projects.

5.1.2. The Determinants of Quality

It is clear from the last section that an improvement in the quality of a project translates into much better prospects for good performance during implementation. It is therefore important to not only consider quality as endogenous, but also study its determinants for policy purposes. The right panel of table 5 presents regressions of projects' quality at entry. The last regression on which we focus here explains 27% of the variation in quality, most of which comes from introducing the quality of the borrower preparation. Consistent with the finding of Deininger, et al. (1998), the stock of ESWs has a positive impact on the quality of projects, with an additional piece of ESW increasing the chances of high quality projects in the future by almost 3%. Europe and Central Asia, along with Latin America and the Caribbean have a greater probability of receiving high quality projects. The African dummy has a negative sign, though insignificant. Strangely enough, adjustment loans have a better quality at entry but perform poorly as shown in the previous section, which might stem from their high complexity.

Voice and accountability and the quality of the regulatory framework have significant negative impacts, meaning that a more accountable country that conducts its policies in a sound macroeconomic environment is more likely to receive worse quality projects. From our theoretical model, even a highly accountable government is likely to accept bad quality projects as long as its screening capacity is low. The negative effect of a sound economic environment on a project's quality is much harder to rationalise and does not receive any obvious answer from our

model. Government effectiveness has a strong positive impact on the probability of having a project with a marginal effect of 20%. The most powerful explanatory variable is, however, the quality of the borrower preparation. Projects are thus more likely to be of good quality when the borrower's screening capacity is high and the donor manages to have the borrowing government fully "on board." A more committed borrower is more likely to make more effort in preparing the project which translates into higher quality.

5.2 The Bivariate Probit Specification

The standard probit specifications considered in the previous section ignore the potential endogeneity of the quality at entry. Let us now turn to the more complex two-equation models that address the endogeneity issue. Estimation results for the full bivariate probit models are presented in table 6. The associated marginal effects for model 6, the preferred specification, are in table 7. Model 1 includes only a restricted set of observables, namely the governance indicators. In model 2, the set of covariates is expanded to encompass all project characteristics except the Bank and client country effort-related variables. These latter are included in models 3 to 6. The quality at entry of the project enters the performance equation for specifications 5 and 6 only. Model 6 (resp. model 4), is obtained by adding to model 5 (resp. model 3) the Bank's supervisory performance in the project as assessed by OED.

Can governance indicators alone explain both quality and performance? The answer is given in column 1. Whereas the regulatory framework and the accountability of the government are significant in explaining neither performance nor quality, both political stability and the effectiveness of the government have strong positive impacts on quality and performance. The highly significant coefficient of correlation of 0.78, however, suggests that there is a strong correlation between unobserved variables. Moreover, this model has a very poor predictive power and always predicts satisfactory quality and performance. Regression 2 introduces the number of ESWs as an identifier and an expanded set of covariates such as regional dummies, loan size, among others. The positive significant impact of government effectiveness in explaining performance and quality is strengthened while political stability drops out. The regulatory framework is now negatively correlated with the quality of the project. Both quality and performance deteriorate with higher government accountability. The positive correlation among unobservables remains stable at high positive levels. The predictive power of this regression is slightly higher than that of the previous model but it still puts too heavy a weight on the both satisfactory quality and the outcome event.

The donor and borrower's effort-related variables are present in models 3 to 6. The borrower's involvement at the project's preparatory stage is our second identifier and does therefore not belong to the performance equation. The borrower quality of implementation and compliance are present in all models, whereas Bank's supervision is considered only in regressions 4 and 6. Regressions 3 and 4 are seemingly unrelated bivariate probits, and regressions 5 and 6 are recursive models whereby quality at entry is a right-hand side variable for performance as well. From regressions 4 and 6, it is clear than the quality of Bank supervision is quite important for the success of a project. Consistent with most of the findings in the literature a sound policy environment, proxied here by the quality of the regulatory framework, clearly increases the probability of success of a project as shown in regression 5. However, once the quality of Bank supervision is introduced in the model (regression 6), the economic environment matters much less for project performance. Supervision effort may thus be more important than the quality of the policy environment in explaining project's performance, as also pointed out by Kilby (2000).

Unlike donor's and borrower's effort variables, governance indicators are poor predictors of project performance. Indeed, none of the indicators can explain project performance once quality at entry is taken into account. However, government effectiveness is a powerful predictor of quality at entry, as it displays a stable, positive and highly significant estimate in all regressions. The more effective the government the higher the quality of the projects it receives. From the borrower's viewpoint, its involvement during the preparation of the project has a huge positive impact on quality. For performance, although borrower's compliance is quite important, it is the quality of implementation that matters most. Prior analytical work as given by the number of ESWs strongly influences the quality of the projects. Loan size matters neither for quality nor for performance. ECA, and LAC countries are more likely to receive high quality loans. Projects have a significant higher probability of failure in Africa. Quality was much lower during the 1980s, and IDA loans seem to have a better quality and higher probability of being satisfactory.

The validity of the identifiers will be assessed through two tests. First, they must be significant predictors of project's quality at entry. Second, they should be insignificant if included in the performance equation. The first test is easy and only involves checking whether ESWs and the quality of borrower preparation are jointly significant in the model. This is done through a simple Wald test, where the chi-squared test statistic with two degrees of freedom is equal to 261, making our identifiers jointly significant at any level. For the second test, as suggested by Bollen, et al. (1995) with *n* identifiers, one has to include *n* minus one of the identifiers in the performance equation in order to obtain a just identified model and then test for the (joint) significance of the identifiers. If the restrictions are valid, the identifiers should not be significant determinants of project performance. In this case *n* equals two so there two possible combinations, and both have been tested. The chi-squared test statistic with one degree of freedom is equal to 1.21, and 0.67 when borrower preparation and the number of ESW are included in the performance equation. The null that these are not significant determinants cannot be rejected. One can now be confident that the exclusion restrictions are valid and therefore expect the estimates of the bivariate probit model to be robust.

Typically, when the exclusion restrictions are not strong enough, the bivariate model performs very poorly. A fairly good signal of this is given by the variance of the estimates in the two-equation model, which are usually large relative to their counterpart in the simple probit regressions. In this model, however, comparing standard errors in tables 5 and 6, the coefficients of the bivariate probit display a precision similar to those in the single equation models. The bivariate probit therefore performs quite well in this instance. Also, note that the standard errors have been adjusted for within-country clustering.

5.2.1. Marginal Effects

While estimates tell us whether the variables are significant determinants of performance or quality, they do not tell us how important a determinant a variable is. Marginal effects provide us with this piece of information by computing the impact of a marginal change in a variable of interest on the probability of satisfactory quality and/or performance. Table 7 presents the marginal impacts of the exogenous variables on selected joint, marginal, and conditional probabilities. Columns 4 and 5, which represent the marginal effect on the probability of having a performing project or a project with satisfactory quality at entry, respectively, are the sum of columns 1 and 2, and 1 and 3, respectively.[22] Column 6 gives marginal effects on the probability that the project's outcome will be satisfactory conditional on having a good quality at entry.

[22] This is simply $Prob(x=1) = Prob(x = 1, y = 0)+Prob(x = 1, y = 1)$ where the last two terms are the joint probabilities.

By far, the most powerful variable in determining success is the quality at entry of the project. A good quality project has a 25.8% higher probability of performing well during implementation than a poorly designed project. The fact that the bivariate estimate is substantially higher than the univariate estimate (17.5%) hints to a causal relationship between quality and performance. There is thus more than a mere statistical association between these variables; their association ought to be structural. A quick comparison between the marginal effects predicted by the single equation model and this recursive one shows that except for quality at entry, all marginal effects are similar in importance across the models. The borrower's quality of implementation, and supervision by the Bank also have an important impact on project performance, but this impact is slightly reduced with respect to the single probit model. The borrower's involvement during the preparation stage of the project has a tremendous impact on the quality at entry with the marginal impact of 48 %, as in the probit case. The effectiveness of the government is the second most important variable in explaining the quality at entry of projects. As in the single equation model, a marginal increase in effectiveness increases the probability of having a good quality project by 20.5%. An additional piece of ESW increases by 2.7% the chances for high quality projects in the future. A marginal improvement in the government's accountability or the country's regulatory environment would decrease the probability of satisfactory quality for the project by 5 and 13%, respectively.

Coming to joint probabilities, better supervision decrease the probability of failure of a satisfactory quality project by 13.3% and at the same time it increases the probability of success of a poorly designed project by 2.9%. As shown by Kilby (2000) but unlike DS (2000), timely supervision is very important to keep good projects on track, and it might help salvage badly designed projects, although it is much less effective in this respect. The quality of implementation has the same impact on the marginal probabilities as supervision, but with a higher intensity. Interestingly, marginal improvements in the soundness of the regulatory environment increase the chances of a bad project ending up with a satisfactory rating by 13.4%, and significantly decreases the probability that a well-designed project fails by 6%. Africa has a 5.6% less chance for project success, but it also has a 6% lower probability of receiving a well-designed project.

6. Conclusion

This paper tackles the recently hotly debated issue of the quality of foreign aid. It empirically demonstrates that the quality of aid has a tremendous impact on its effectiveness. The paper also shows that design is an essential component of aid quality. The design of aid is an upstream process undertaken by aid agencies. Therefore, the role and impact of donors in determining the quality of the assistance they provide goes well beyond selectivity. Aid is not donor-neutral and its quality is not exclusively determined by the characteristics of its recipients.

The paper establishes both theoretically and empirically that the quality of aid is endogenous to the incentive system that prevails in the aid agency and the capacity and accountability of the recipient country. On the agency side, the higher the effort by the agency staff to design a project the better is its quality and probability of having a positive development impact. Recipient governments may, on one hand, have strong incentives to accept projects because projects bring personal benefits. On the other hand they might be deterred from welcoming bad projects if they are accountable "enough" to their citizenry in the event of project failure. On the agency side, unless there exists a minimum standard for project's quality, the staff exert more effort on project's design only when they are matched with a highly accountable perfect screener. When the incentive system in the agency leans towards a culture of "pushing money," where only the number of accepted projects matters, then all low-capacity governments will receive poorly designed projects. High capacity governments with low accountability are in

this instance willing to accept bad projects for their own benefit to the detriment of their population. Highly accountable and capable governments will refuse all bad aid projects.

This confirms Burnside and Dollar (2000) celebrated result that "aid works in good environments only" holds but hides a serious endogeneity issue. In effect, it is the combination of perverse incentives on both sides, and recipient's lack of capacity and accountability that channels bad aid into bad environments. Aid agencies have the capacity to direct good aid even into bad environments by adopting an incentive system that only rewards good projects.

This result has strong implications for the use of selectivity as a yardstick for allocating aid to recipient countries. Selectivity is becoming the cornerstone of foreign assistance among many bilateral and multilateral donors. As a matter of fact, both IDA allocations and the recently developed aid strategy of the United States administration, the Millennium Challenge Account, are based on this paradigm. Selectivity is grounded on efficiency arguments. For aid to be effective, it should target countries able to translate a dollar of aid into economic growth or poverty reduction. There are potential large gains in targeting aid to appropriate countries as shown by Collier and Dollar (2002). However, the risks involved are also commensurate to those gains. Indeed, for a selective aid strategy, only countries are those that carry out good economic policies in a democratic environment are aid-deserving. However, plenty of evidence points to the existence of a high correlation between income per capita and "good" environments. A selective allocation of aid would thus exclude the poorest countries from the aid sphere, and it is exactly there that 'good' aid can have the greatest impact. Because donors can affect the quality of aid, they must select for both the countries they assist and the quality of the projects they finance.

References

Adams, C. S. and S. A. O'Connell. (1999). 'Aid, taxation and development in sub-saharan Africa.' *Economics and Politics*, vol. 11(3), pp. 225-53.

Altonji, G. A., Elder, T. E. and Taber, C. R. (2000). 'Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools.' NBER Working Paper 7831. National Bureau of Economic Research, Cambridge, Mass.

Altonji, G. A., Elder, T. E. and Taber, C. R. (2002). 'An evaluation of instrumental variable strategies for estimating the effects of catholic schools.' NBER Working Paper 9358. National Bureau of Economic Research, Cambridge, Mass.

Angrist, J. D. (1991). 'Instrumental variables estimation of average treatment effects in econometrics and epidemiology.' NBER Technical Working Paper 115. National Bureau of Economic Research, Cambridge, Mass.

Azam, J. P. and Laffont, J. J. (2003). 'Contracting for aid.' *Journal of Development Economics* vol. 70, pp. 25-58.

Bollen, K. A., Guilkey, D. K. and Mroz, T. A. (1995). 'Binary outcomes and endogenous explanatory variables: tests and solutions with an application to the demand for contraceptive use in Tunisia.' *Demography* vol. 32(1), pp. 111-31.

Boone, P. (1996). 'Politics and the effectiveness of foreign aid.' *European Economic Review* vol. 40: 289-329.

Burnside, C. and Dollar, D. (2000). 'Aid, policies, and growth.' *American Economic Review* vol. 90(4), (September), pp. 847-68.

Collier, P., and Dollar, D. (2002). 'Aid allocation and poverty reduction.' *European Economic Review* vol. 46(8), pp. 1475-1500

Deininger, K., Squire, L. and Basu, S. (1998). 'Does economic analysis improve the quality of foreign assistance?' *World Bank Economic Review* vol. 12(3), pp. 385-418.

Dewatripont, M., Jewitt, I. and Tirole, J. (1999). 'The economics of career concerns, Part II: application to missions and accountability of government agencies.' *Review of Economic Studies* vol. 66, pp. 199-217.

Dollar, D., and Svensson, J. (2000). 'What explains the success or failure of structural adjustment programmes?' **ECONOMIC JOURNAL** vol. 110, pp. 894-917.

Easterly, W. (2002). 'Evaluating aid performance of donors.' Center for Global Development, October 2002.

Easterly, W., . Levine, R and Roodman, D. (forthcoming). 'New data, new doubts: revisiting 'Aid, policies, and growth'.' *American Economic Review*.

Evans, W. N. and Schwab, R. M. (1995). 'Finishing high school and starting college: do catholic schools make a difference?' *Quarterly Journal of Economics* vol. 110, pp. 941-74.

Evans, W. N., Farrely, M. C. and Montgomery, E. (1999). 'Do workplace smoking bans reduce smoking?' *American Economic Review* vol. 89(4), pp. 728-47.

Fairburn, J. A., and Malcomson, J. M. (2001). 'Performance, promotion, and the Peter principle.' *Review of Economic Studies* vol. 68, pp. 45-66.

Ferejohn, J. (1986). 'Incumbent performance and electoral control.' *Public Choice* vol. 50, pp. 5-25

Hansen, H., and Tarp, F. (2000). 'Aid effectiveness disputed.' *Journal of International Development* vol. 12, pp. 375-98.

Hansen, H., and Tarp, F. (2001). 'Aid and growth regressions.' *Journal of Development Economics* vol. 62 (4), pp. 547-70.

Isham, J., and Kaufmann, D. (1999). 'The forgotten rationale for policy reform: the productivity of investment projects.' *Quarterly Journal of Economics* vol. 114, pp. 149-84.

Kilby, C. (2000). 'Supervision and Performance: The case of World Bank projects.' *Journal of Development Economics* vol. 62, pp. 233-59.

Kaufmann, D., Kraay, A. and Mastruzzi, M. (2003). 'Governance matters III: governance indicators for 1996-2002.' Policy Research Paper 3106.World Bank, Washington, D.C.

Kaufmann, D., Kraay, A. and Zoido-Lobaton, P. (1999). 'Governance matters.' Policy Research Paper 2196.World Bank, Washington, D.C.

Kaufmann, D., and Wang, Y. (1995). 'Macroeconomic policies and project performance in the social sectors: a model of human capital production and evidence from the LDCs.' *World Development* vol. 23(5), pp. 751-65.

Kosack, S. (2003). 'Effective aid: How democracy allows development aid to improve the quality of life.' *World Development* vol. 31(1), pp. 1-22.

Llavador, H. G., and Roemer, J. E. (2001). 'An equal-opportunity approach to the allocation of international aid.' *Journal of Development Economics* vol. 64, pp. 147-71.

Maddala, G. S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge, U.K.: Cambridge University Press.

Mosley, P. (1985). 'The political economy of foreign aid: a model of the market for a public good.' *Economic Development and Cultural Change* vol. 33(2), pp. 373-93.

Mosley, P. (1996). 'The failure of aid and adjustment policies in Sub-Saharan Africa: counter-examples and policy proposals.' *Journal of African Economies* vol. 5(3), pp. 406-43.

Peter, L. J., and Hull, R. (1969). *The Peter Principle: Why things always go wrong*. New York: William Morrow and Company Inc.

Rao, J. M. (1994). 'Judging givers: equity and scale in aid allocation.' *World Development* vol. 22(10), pp. 1579-84.

Roodman, D. 2003. 'An index of donor aid performance.' Center for Global Development, Washington, D.C. Processed.

Sah, R. K., and Stiglitz, J. E. (1986). 'The architecture of economic systems: hierarchies and polyarchies.' *American Economic Review* vol. 76(4), pp. 716-27.

Seabright, P. (1996). 'Accountability and decentralisation in government: an incomplete contracts approach.' *European Economic Review*, vol. 40, pp. 61-89.

Svensson, J. (1999). 'Aid, growth, and democracy.' *Economics and Politics* vol. 11(3), pp. 275-97.

Svensson, J. (2000). 'When is foreign aid policy credible? Aid dependence and conditionality.' *Journal of Development Economics* vol. 61, pp. 61-84.

Svensson, J. (2003). 'Why conditional aid does not work and what can be done about it?' *Journal of Development Economics* vol. 70, pp. 381-402.

Wapenhans, W. 1992. *Effective implementation: key to development impact*. Report of the World Bank's Portfolio Management Task Force. Washington, D.C.: World Bank.

White, H., and Woestman, L. (1994). 'The quality of aid: measuring trends in donor performance.' *Development and Change* vol. 25(3), pp. 527-54.

World Bank and International Monetary Fund. 2003. 'Statement of the Development Committee.' Sixty-sixth meeting of the development committee on September 28, 2002. Washington, D.C. Statement by Antonio Fozio

Table 1: Governance Indicators by Region

	Africa	*East Asia & The Pacific (EAP)*	*Europe & Central Asia (ECA)*	*Latin America and The Caribbean (LAC)*	*Middle-East & North Africa (MENA)*	*South Asia (SAR)*	*Full Sample*
Governance Indicators							
Voice & Accountability	-0.53	-0.29	-0.03	0.40	-0.73	-0.53	-0.28
Political Stability	-0.39	-0.01	0.07	0.02	-0.51	-0.31	-0.19
Government Effectiveness	-0.59	-0.28	-0.21	-0.13	-0.18	-0.13	-0.25
Regulatory Quality	-0.47	-0.44	-0.16	0.31	-0.26	-0.14	-0.19
Rule of Law	-0.60	-0.42	-0.21	-0.11	-0.11	-0.31	-0.29
Control of Corruption	-0.56	-0.42	-0.26	-0.13	-0.25	-0.24	-0.31

Table 2: Features of Successful and Unsuccessful Projects

	Successful	*Unsuccessful*
Country Governance Characteristics		
Average Voice and Accountability	–0.27	–0.4
Average Political Stability	–0.25	–0.5
Average Government Effectiveness	–0.23	–0.44
Average Regulatory Quality	–0.09	–0.26
Average Rule of Law	–0.31	–0.47
Average Control of Corruption	–0.37	–0.52
Project Characteristics		
Loan Size ($Million)	99.4	98.0
Preparation Resources ($'000)	337	343
Supervision Resources ($'000)	373	423

Table 3: Project Ratings by Region

	Africa	*East Asia & The Pacific*	*Europe & Central Asia*	*Latin America and The Caribbean*	*Middle-East & North Africa*	*South Asia*	*Full Sample*
Number of Countries	*42*	*15*	*29*	*30*	*9*	*7*	*132*
Number of Projects	*491*	*265*	*281*	*319*	*125*	*216*	*1697*
Average Size of Projects	*42.9*	*149.7*	*111.1*	*118.4*	*78.3*	*132.1*	*99.1*
Project Performance (%)							
Highly Satisfactory	1.8	7.2	10.3	8.8	4.8	5.5	6.1
Satisfactory	59.7	73.6	71.5	72.1	63.2	66.7	67.3
Unsatisfactory	35.4	17.3	17.1	18.1	27.2	25.0	24.4
Highly Unsatisfactory	3.1	1.9	1.1	0.9	4.8	2.8	2.2
Project Quality at Entry (%)							
Highly Satisfactory	2.9	6.5	13.9	10.1	6.4	6.0	7.2
Satisfactory	55.4	69.8	68.7	65.8	66.4	54.6	62.5
Unsatisfactory	40.1	22.6	15.6	23.2	24.8	37.5	28.7
Highly Unsatisfactory	1.6	1.1	1.8	0.9	2.4	1.9	1.5

Table 4: Correlation between Quality at Entry and Performance

	Project Quality at Entry			
Project Performance	Highly Unsatisfactory	Unsatisfactory	Satisfactory	Highly Satisfactory
Highly Unsatisfactory	***13***	*20*	*5*	*0*
Unsatisfactory	*12*	***282***	*117*	*3*
Satisfactory	*1*	*185*	***889***	*67*
Highly Satisfactory	*0*	*0*	*50*	***53***

Table 5: Project Performance and Quality Univariate Probit Regressions

Regression No	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Number of Observations	1687	1683	1683	1683	1687	1683	1683
Number of Countries	126	124	124	124	126	124	124
	Project Performance				***Project Quality at Entry***		
Project Quality at Entry				0.735*** (0.108)			
				[0.175]			
Voice and Accountability	-0.089 (0.093)	-0.085 (0.075)	-0.033 (0.114)	-0.025 (0.116)	-0.092 (0.102)	-0.217*** (0.078)	-0.160** (0.081)
	[-0.029]	[-0.026]	[-0.007]	[-0.005]	[-0.032]	[-0.072]	[-0.051]
Political Stability	0.159** (0.077)	0.072 (0.081)	0.006 (0.093)	0.027 (0.100)	0.180** (0.088)	0.118 (0.078)	-0.007 (0.082)
	[0.051]	[0.022]	[0.001]	[0.006]	[0.062]	[0.039]	[-0.002]
Government Effectiveness	0.436** (0.177)	0.525*** (0.184)	0.399* (0.221)	0.227 (0.227)	0.471*** (0.163)	0.707*** (0.165)	0.650*** (0.179)
	[0.141]	[0.162]	[0.084]	[0.046]	[0.163]	[0.236]	[0.205]
Regulatory Quality	0.046 (0.162)	0.014 (0.169)	0.106 (0.203)	0.256 (0.209)	-0.204 (0.170)	-0.388** (0.154)	-0.431*** (0.154)
	[0.015]	[0.004]	[0.022]	[0.052]	[-0.070]	[-0.130]	[-0.136]
Bank Quality of Supervision			0.981*** (0.097)	0.808*** (0.102)			
			[0.207]	[0.164]			
Borrower Quality of Implementation			1.401*** (0.121)	1.262*** (0.129)			
			[0.296]	[0.256]			
Borrower Quality of Compliance			0.404*** (0.110)	0.399*** (0.107)			
			[0.085]	[0.081]			
Borrower Quality of Preparation							1.531*** (0.105)
							[0.484]
ESW Deliveries at Project Entry						0.098*** (0.035)	0.089** (0.040)
						[0.033]	[0.028]
Log of Loan Size ($ Millions)		-0.042 (0.051)	-0.039 (0.063)	-0.014 (0.067)		-0.089* (0.054)	-0.091 (0.060)
		[-0.013]	[-0.008]	[-0.003]		[-0.030]	[-0.029]
Preparation Resources (% of Loan Size)		-0.067* (0.040)	-0.094** (0.044)	-0.091* (0.048)		-0.045 (0.036)	-0.057 (0.044)
		[-0.021]	[-0.020]	[-0.018]		[-0.015]	[-0.018]
Cancelled Loan Amount (% of Loan Size)		-2.351*** (0.218)	-1.654*** (0.278)	-1.540*** (0.282)			
		[-0.724]	[-0.350]	[-0.312]			
IDA Loan		0.175* (0.103)	0.226* (0.134)	0.261* (0.143)		0.143 (0.112)	0.179* (0.108)
		[0.054]	[0.048]	[0.053]		[0.048]	[0.056]
Adjustment Loan		0.021 (0.129)	-0.275* (0.160)	-0.382** (0.159)		0.400*** (0.137)	0.306** (0.150)
		[0.006]	[-0.064]	[-0.090]		[0.121]	[0.089]
Eighties		-0.323*** (0.090)	0.050 (0.110)	0.095 (0.118)		-0.356*** (0.106)	-0.275*** (0.105)
		[-0.105]	[0.010]	[0.019]		[-0.125]	[-0.091]
Africa		-0.482*** (0.108)	-0.280** (0.133)	-0.279* (0.153)		-0.261** (0.108)	-0.179 (0.127)
		[-0.158]	[-0.063]	[-0.060]		[-0.090]	[-0.058]
Europe & Central Asia		0.135 (0.146)	0.069 (0.219)	-0.012 (0.242)		0.401*** (0.159)	0.383*** (0.147)
		[0.040]	[0.014]	[-0.002]		[0.122]	[0.110]
Latin America & Caribbean		0.109 (0.178)	0.069 (0.207)	-0.005 (0.230)		0.471*** (0.134)	0.423*** (0.147)
		[0.033]	[0.014]	[-0.001]		[0.142]	[0.121]
Constant	0.809*** (0.062)	1.491*** (0.274)	-5.975*** (0.513)	-5.701*** (0.562)	0.675*** (0.064)	0.923*** (0.268)	-3.487*** (0.415)
Pseudo R-squared	0.03	0.16	0.55	0.57	0.02	0.08	0.27
Log Likelihood	-941.45	-814.31	-441.58	-417.43	-1008.41	-943.13	-756.18

Robust standard errors in parentheses. Marginal Effects in Brackets and computed at means. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 6: Project Performance and Quality (Recursive) Bivariate Probit Regressions

Regression No	(1)		(2)		(3)		(4)		(5)		(6)	
Number of Observations	1687		1683		1683		1683		1683		1683	
Number of Countries	126		124		124		124		124		124	
	Performance	Quality	Performance	Quality	Performance	Quality	Performance	Quality	Performance	Quality	Performance	Quality
Project Quality at Entry									1.329***		1.017***	
									(0.169)		(0.181)	
Voice and Accountability	-0.100	-0.104	-0.132*	-0.223***	-0.069	-0.168**	-0.062	-0.163**	0.011	-0.158*	-0.005	-0.158**
	(0.091)	(0.102)	(0.069)	(0.076)	(0.098)	(0.082)	(0.108)	(0.082)	(0.106)	(0.081)	(0.113)	(0.081)
Political Stability	0.161**	0.184**	0.106	0.121	0.080	-0.002	0.040	-0.004	0.043	-0.005	0.014	-0.007
	(0.077)	(0.089)	(0.079)	(0.080)	(0.085)	(0.083)	(0.092)	(0.083)	(0.090)	(0.081)	(0.098)	(0.081)
Government Effectiveness	0.440**	0.472***	0.547***	0.697***	0.224	0.676***	0.383*	0.662***	-0.033	0.642***	0.168	0.648***
	(0.174)	(0.158)	(0.189)	(0.161)	(0.199)	(0.180)	(0.221)	(0.180)	(0.206)	(0.177)	(0.228)	(0.178)
Regulatory Quality	0.065	-0.202	0.016	-0.375**	0.194	-0.439***	0.141	-0.434***	0.369*	-0.430***	0.282	-0.433***
	(0.164)	(0.172)	(0.168)	(0.154)	(0.193)	(0.155)	(0.201)	(0.154)	(0.201)	(0.154)	(0.209)	(0.154)
Bank Quality of Supervision							0.864***				0.789***	
							(0.098)				(0.100)	
Borrower Quality of Implementation					1.483***		1.340***		1.312***		1.209***	
					(0.115)		(0.121)		(0.130)		(0.135)	
Borrower Quality of Compliance					0.452***		0.382***		0.470***		0.399***	
					(0.092)		(0.105)		(0.095)		(0.105)	
Borrower Quality of Preparation						1.426***		1.477***		1.533***		1.533***
						(0.101)		(0.102)		(0.105)		(0.105)
ESW Deliveries at Project Entry				0.085***		0.090**		0.092**		0.085**		0.086**
				(0.026)		(0.038)		(0.038)		(0.041)		(0.040)
Log of Loan Size ($ Millions)			-0.045	-0.091*	0.013	-0.094	-0.033	-0.092	0.047	-0.090	-0.007	-0.091
			(0.054)	(0.052)	(0.065)	(0.060)	(0.064)	(0.060)	(0.068)	(0.060)	(0.067)	(0.060)
Preparation Resources (% of Loan Size)			-0.082*	-0.051	-0.119***	-0.056	-0.099**	-0.056	-0.107**	-0.057	-0.084*	-0.058
			(0.042)	(0.035)	(0.041)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)	(0.048)	(0.044)
Supervision Resources (% of Loan Size)			0.015		0.045				0.052			
			(0.025)		(0.038)				(0.039)			
Cancelled Loan Amount (% of Size)			-1.519***		-1.362***		-1.548***		-1.345***		-1.521***	
			(0.203)		(0.255)		(0.272)		(0.267)		(0.279)	
IDA Loan			0.240**	0.141	0.302**	0.171	0.258*	0.174	0.282*	0.183*	0.250*	0.182*
			(0.105)	(0.110)	(0.136)	(0.107)	(0.134)	(0.107)	(0.146)	(0.108)	(0.143)	(0.108)
Adjustment Loan			0.086	0.407***	-0.105	0.305**	-0.275*	0.303**	-0.296*	0.314**	-0.410**	0.311**
			(0.124)	(0.132)	(0.156)	(0.149)	(0.158)	(0.150)	(0.160)	(0.150)	(0.162)	(0.150)
Eighties			-0.313***	-0.364***	-0.034	-0.292***	0.023	-0.283***	0.115	-0.267**	0.125	-0.271**
			(0.085)	(0.105)	(0.098)	(0.105)	(0.110)	(0.105)	(0.111)	(0.105)	(0.121)	(0.105)
Africa			-0.475***	-0.251**	-0.359***	-0.175	-0.306**	-0.173	-0.295*	-0.187	-0.257*	-0.184
			(0.105)	(0.108)	(0.132)	(0.128)	(0.137)	(0.128)	(0.154)	(0.127)	(0.153)	(0.127)
Europe & Central Asia			0.137	0.396**	0.063	0.383***	0.073	0.381***	-0.099	0.388***	-0.043	0.386***
			(0.141)	(0.156)	(0.213)	(0.146)	(0.220)	(0.146)	(0.245)	(0.148)	(0.245)	(0.148)
Latin America & Caribbean			0.173	0.459***	0.137	0.431***	0.084	0.426***	-0.039	0.423***	-0.039	0.424***
			(0.180)	(0.128)	(0.211)	(0.147)	(0.211)	(0.147)	(0.232)	(0.145)	(0.229)	(0.146)
Constant	0.813***	0.674***	1.331***	0.937***	-3.969***	-3.165***	-5.456***	-3.324***	-4.590***	-3.501***	-5.755***	-3.494***
	(0.062)	(0.063)	(0.294)	(0.261)	(0.434)	(0.400)	(0.544)	(0.405)	(0.477)	(0.413)	(0.563)	(0.414)
Correlation Coefficient (ρ)	0.78***		0.73***		0.42***		0.32***		-0.26**		-0.20*	
	(0.024)		(0.031)		(0.054)		(0.058)		(0.105)		(0.111)	
Log Likelihood	-1716.38		-1590.53		-1228.11		-1184.16		-1205.13		-1172.38	

Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Marginal Effects of Project Performance and Quality (Recursive) Bivariate Probit Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	Prob(perf=1, qae=1)	Prob(perf=1, qae=0)	Prob(perf=0, qae=1)	Prob(perf=1)	Prob(qae=1)	Prob(perf=1| qae=1)
Project Quality at Entry	0.208***	0.049***	-0.208***	0.258***	(no effect)	0.276***
	(0.056)	(0.008)	(0.057)	(0.062)		(0.075)
Voice and Accountability	-0.046	0.045*	-0.004	-0.001	-0.050**	-0.004
	(0.029)	(0.023)	(0.02)	(0.024)	(0.025)	(0.026)
Political Stability	0.0005	0.0024	-0.002	0.003	-0.002	0.003
	(0.027)	(0.023)	(0.018)	(0.022)	(0.024)	(0.024)
Government Effectiveness	0.214***	-0.178***	-0.009	0.034	0.205***	0.05
	(0.060)	(0.051)	(0.039)	(0.047)	(0.058)	(0.05)
Regulatory Quality	-0.076	0.134***	-0.061*	0.058	-0.137***	0.055
	(0.053)	(0.043)	(0.031)	(0.038)	(0.047)	(0.04)
Bank Quality of Supervision	0.136***	0.0289***	-0.133***	0.162***	(no effect)	0.178***
	(0.020)	(0.008)	(0.019)	(0.022)		(0.026)
Borrower Quality of Implementation	0.205***	0.044***	-0.205***	0.249***	(no effect)	0.272***
	(0.019)	(0.012)	(0.019)	(0.024)		(0.026)
Borrower Quality of Compliance	0.068***	0.015***	-0.067***	0.082***	(no effect)	0.089***
	(0.018)	(0.005)	(0.018)	(0.021)		(0.023)
Borrower Quality of Preparation	0.439***	-0.439***	0.046***	(no effect)	0.485***	0.029
	(0.030)	(0.030)	(0.01)		(0.032)	(0.018)
ESW Deliveries at Project Entry	0.025**	-0.025**	0.0026**	(no effect)	0.027**	0.0016
	(0.010)	(0.010)	(0.001)		(0.011)	(0.0012)
Log of Loan Size ($ Millions)	-0.027*	0.026*	-0.001	-0.001	-0.029*	-0.003
	(0.016)	(0.014)	(0.011)	(0.013)	(0.015)	(0.014)
Preparation Resources (% of Loan Size)	-0.031**	0.013	0.012	-0.017*	-0.018	-0.02*
	(0.012)	(0.013)	(0.008)	(0.01)	(0.013)	(0.011)
Cancelled Loan Amount (% of Size)	-0.257***	-0.056***	0.257***	-0.313***	(no effect)	-0.342***
	(0.047)	(0.016)	(0.047)	(0.054)		(0.062)
IDA Loan	0.094***	-0.043	-0.036	0.051*	0.057*	0.0594*
	(0.033)	(0.029)	(0.023)	(0.028)	(0.03)	(0.031)
Adjustment Loan	-0.0036	-0.095***	0.094**	-0.098**	0.091**	-0.099**
	(0.046)	(0.032)	(0.038)	(0.042)	(0.037)	(0.045)
Eighties	-0.062**	0.086***	-0.028	0.025	-0.09***	0.022
	(0.031)	(0.029)	(0.019)	(0.023)	(0.03)	(0.026)
Africa	-0.099**	0.0425	0.039	-0.056*	-0.06*	-0.065*
	(0.04)	(0.034)	(0.027)	(0.033)	(0.037)	(0.037)
Europe & Central Asia	0.093**	-0.102***	0.018	-0.009	0.11***	-0.003
	(0.045)	(0.032)	(0.035)	(0.039)	(0.035)	(0.041)
Latin America & Caribbean	0.103**	-0.111***	0.018	-0.008	0.122***	-0.001
	(0.044)	(0.032)	(0.033)	(0.038)	(0.034)	(0.04)

Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

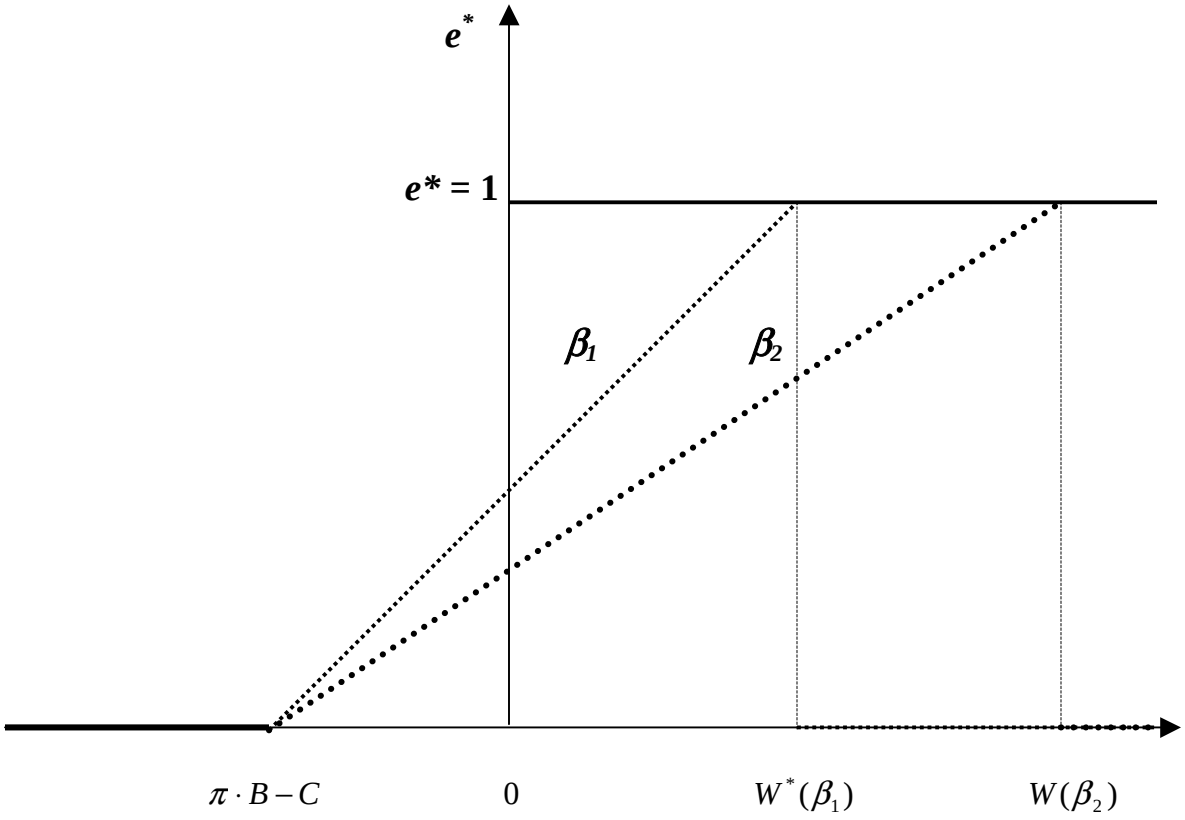

Fig.2: High Capacity Government

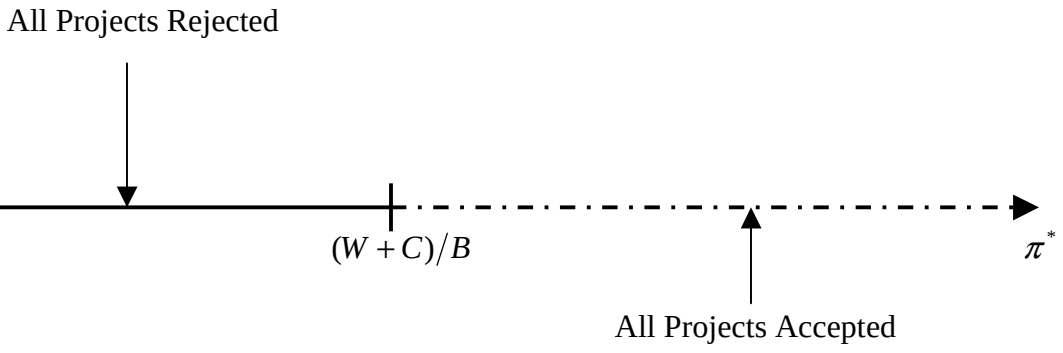

Fig. 3: Low Capacity Government

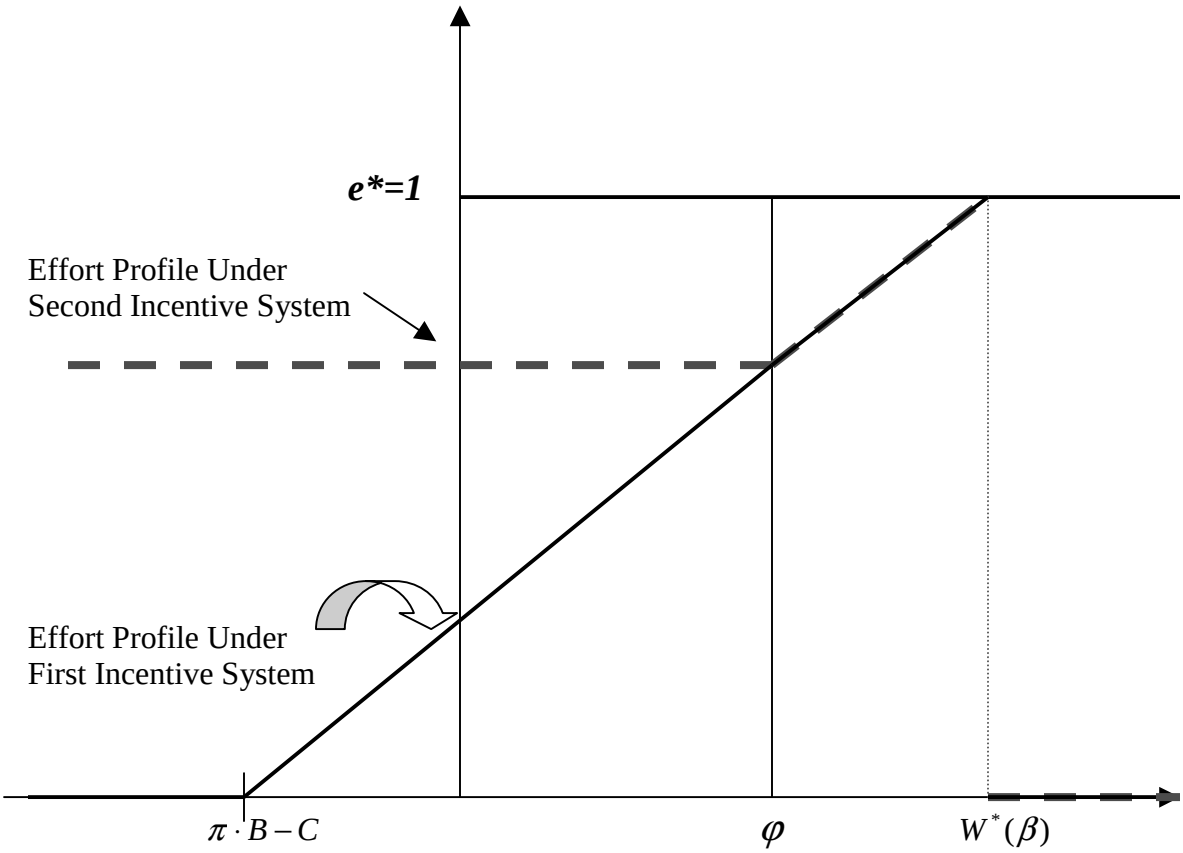

Fig.4: High Capacity Government