# Does Economic Analysis Improve the Quality of Foreign Assistance?

Klaus Deininger, Lyn Squire, and Swati Basu

*October 27, 1998*

We would like to thank Bill Battaile, Shanta Devarajan, David Dollar, Christopher Kilby, Lant Pritchett, and participants at a World Bank workshop on Aid Effectiveness for detailed comments on an earlier draft. Support from the Bank's research support budget is gratefully acknowledged. The findings, interpretations, and conclusions expressed in this paper are those of the authors. They do not necessarily reflect the views of the World Bank, its Executive Directors, or the countries they represent.

**Does Economic Analysis Improve the Quality of Foreign Assistance?**

What is the benefit of the millions of dollars of foreign assistance provided to developing countries in the form of economic analysis and advice? This is a difficult question to answer for two reasons. First, in the absence of a clear market test, there is no simple way of valuing the contribution of analytical work; most of it is provided free of charge or else embedded in an investment or program loan. Second, the impact of economic analysis on some ultimate objective -- such as development or poverty reduction -- is almost impossible to identify because it is only one of many factors that determine outcomes.

Yet, the issue is of obvious importance. One would like to know whether past analytical work has generated measurable economic benefits that would justify its continued provision in an environment of increased scarcity of resources. Moreover, with increasing access of developing countries to non-concessional sources of finance, the comparative advantage of international institutions such as the World Bank will shift, it is claimed, towards the provision of analytical services (see for example Rodrik, 1995). But little empirical evidence exists that could be used to substantiate such a claim, or provide guidance on how and where to focus such services in the future.

The issue to be tackled in this paper faces the same problems that confront efforts to evaluate research more generally. The literature on this topic takes the view that spending on research is an investment that should produce measurable economic returns. This is the view adopted in this paper. The difficulty of course arises with the word "measurable". Traditionally, economists have relied on two approaches -- the production function approach and the consumer surplus approach (Averch, 1994). Both rely on market outcomes to provide measures of the benefits of research. Often, however, this is not possible. Research is frequently trying to create more basic knowledge which, like all public goods, proves difficult to value. This happens to be true in our case. Economic studies and policy advice creates and disseminates knowledge, the ultimate impact of which can be far-reaching but is usually difficult to trace.

An alternative approach is to focus more directly on the immediate outputs of research -- reports generated or articles published (Wise and Agranoff, 1991). For example, the National Science Foundation identified a set of indicators that included publications, citations, data archiving and sharing, software development and sharing, and education and pedagogy (National Science Foundation, 1995). The World Bank has also used a variant of this approach -- specifically, the number of World Bank citations appearing in university reading lists (World Bank, 1991). Apart from the incentive problems that these

measures may create. they fail to provide any measure of the value of research. In short, the available literature on the evaluation of research does not provide much guidance on how one might value the output from economic analysis and advice and hence how one might estimate the return to such activity.

This paper provides a partial answer to this question. It is partial in two respects. First, it does not encompass all the benefits of economic analysis and advice. Foreign assistance provided in the form of analysis and advice attempts, among other things, to: maintain a policy dialogue with the client country in order to provide analytical support for critical macroeconomic decisions; examine the steps needed to generate and maintain a policy and regulatory environment that promotes private sector activity; provide analytical inputs to the design of sectoral strategies; and build local capacity for policy analysis and discussion of key policy issues in the broader context of civil society. In addition, economic analysis and advice often provides a critical underpinning for the design of individual projects and lending programs. This last aspect is the focus of this paper. As such, our results will be a minimum estimate of the full benefits from this form of foreign assistance. And second, the only data set providing measures of resources spent on analytical services as well as measures of the benefits of lending is for the World Bank. We cannot therefore assess the value of analytical work and advice supplied by other agencies. Thus, the purpose of the paper is to assess the impact of the Economic and Sector work (ESW) provided by the World Bank on the quality of World Bank landing.

Although data availability forces us to focus on the World Bank's ESW, this is not insignificant either in coverage or quantity. ESW is undertaken on all developing countries which are members of the World Bank -- this amounts to well over 100 countries ranging from China to Chad. The work undertaken on each country is specific to that country but could include: a Country Economic Memorandum, a comprehensive account of economic performance and prospects, and/or more topic-focussed reports such as Poverty Assessments, Public Expenditure Reviews, Labor Market Studies, as well as a wide range of sectoral studies. In general, the more aggregate studies underpin the World Bank's policy advice while the sectoral studies -- Reviews of the Transport Sector, Health and Education Sector Reports, and so -- provide the foundation for lending operations in those sectors. Undertaking this range and magnitude of work is expensive -- since 1975 the World Bank has used about 22,000 staffweeks a year on ESW. This amounts to an annual expenditure of around US$60 million on economic analysis and advice. This almost certainly makes the World Bank one of the largest sources of such analysis for the devloping world.

Against this background, the paper addresses three issues:

- **Does ESW improve the quality (variously defined) of World Bank loans?**

We find that ESW has a significant positive impact on various measures of quality of World Bank projects. For example, a one-staffweek increase in the amount of time devoted to ESW before project initiation is associated with an increase in the economic rate of return (ERR) for an individual project of between 0.02 and 0.04 percentage points, translating into an increase of between US $12,000 and US $25,000 in the project's net present value (NPV) for a cost of no more than US $3,000 -- a dollar of ESW yields four to eight dollars in development impact. Indeed, since each staffweek of ESW benefits more than one project, this is an underestimate. Examining the impact of ESW for the entire lending program for a country, we find that a dollar of ESW yields between twelve and fifteen dollars of development impact. Even this is an underestimate because our calculation fails to capture any benefit ESW may have in terms of influencing policy formulation in developing countries, a key objective of ESW. We therefore conclude that there has been a high pay-off to ESW.

- **Has the World Bank underinvested in ESW relative to lending services?**

Assuming staff resources are fungible across time and between different uses, one would expect the marginal contribution of ESW to the quality of lending to be lower than that of lending services (preparation and supervision), because ESW can be expected to have benefits beyond its immediate impact on lending whereas lending services do not. Indeed, ESW is often undertaken to provide the basis for policy advice to governments and is not necessarily tightly linked to a particular project or lending program. We find, however, the reverse -- ESW has a systematically positive effect on the quality of the lending program, whereas neither preparation nor supervision turn out to be significant. This suggests that reallocation of staff time from lending services to ESW would have increased the quality of the lending program. Consistent with this, we find that task managers at the project level are able to reduce the time allocated to lending services by about 2.5 staffweeks for every staffweek expended on ESW. We infer that ESW helps staff to identify and support new investment options -- it expands the set of feasible projects -- and/or design better projects ex ante -- it improves the quality of projects already in the investment program. Preparation and supervision, on the other hand, can only improve the quality of a project (whether good or bad) ex post.

- **Is there a trade-off between lending quality and volume?**

Even though the preceding results suggest that higher levels of ESW will improve the quality of the lending program, it is still possible that, within a given resource envelope, shifting staff time from lending services to ESW will reduce the overall lending volume. Addressing this question, we find that this was indeed the case -- lending services were between 40 and 50 percent more effective in increasing total commitments than ESW. On the other hand, if disbursements -- resource transfer -- is the variable of interest, we find that managers could have increased both lending quality and disbursements by switching resources from lending services to ESW. This, together with our conclusion that there has been underinvestment in ESW from the standpoint of project quality, suggests that at least to some degree volume of commitments has been an additional objective guiding the disposition of staff resources. Our results provide some insight into this trade-off between quality and quantity -- the analysis suggests that on average a manager was indifferent between a decrease of US $2 million in the NPV of a lending program and an additional US $4 million of lending volume. If this estimate is broadly accurate, it suggests that managers were prepared to allow a substantial reduction in program quality in return for only a small increase in commitments relative to the average program size -- 2 percent.

To explore these questions we adopt the following strategy. We begin by setting out an "idealized" model of decision-making in which managers are assumed to be concerned exclusively with development impact and have sufficiently long planning horizons and sufficient information to be able to allocate resources broadly in line with their objective. Of course, in reality managers may be influenced by factors other than development impact, they may have relatively short time horizons, and the information relevant to successful decision-making may at best be only partial. The intention then is not to treat the "idealized" model as a description of reality, but to see how close reality approximates the ideal given all the factors that inevitably beset decision-making in the real world. Where we observe departures from this ideal, we try to explain the difference. Our model of decision-making operates at two levels. We envisage a "country manager" who makes a broad allocation of resources between lending services and ESW. Given that decision, the "task manager" for each project makes project-specific decisions with respect to the allocation of resources between preparation and supervision.

At the country level, decisions are assumed to be made by a country manager who can allocate staff time freely between lending services and ESW. In this model, therefore, ESW is *endogenous*. The country manager is assumed to maximize "development impact" subject to the overall availability of staff resources. ESW contributes to development impact in two ways -- through its contribution to the quality of

the countrywide lending program[1] and through its contribution to policy formulation more generally. Testing this model allows us to address all three questions posed above. The model also yields (implicitly) the shadow price of a unit of staff resources. This price enters the decision-making problem of the task manager responsible for a single project, the second level of decision-making that we examine in this paper. In this case, we assume that ESW is *exogenous* and that the manager's objective is focused more narrowly on maximizing the quality of the project. This model provides additional evidence on the first two questions posed above.

The paper is organized as follows. Section 1 places the study in the context of other work on foreign assistance and highlights how our data constitute an improvement over those used in earlier studies. Section 2 discusses the conceptual framework and estimating equations. The third section describes the data used in the study and provides descriptive statistics, and Section 4 presents our empirical results. Concluding remarks are presented in Section 5.

## I. RELATION TO THE LITERATURE

Research related to foreign aid falls into three broad categories: game-theoretic models regarding the negotiation and implementation of conditionality and the credibility and commitment problems that arise if donor and/or recipient behave strategically (Mosley et al. 1992, Svensson, 1996); empirical models of how donors allocate aid across recipient countries and how aid may affect decision-making and economic performance of recipient countries (see Burnside and Dollar, 1996, Pack and Pack, 1990 and 1993, Feyzioglu et al. , 1995, Trumbull and Wall, 1994); and efforts, either descriptive or quantitative, to analyze what affects the performance of individual donor-financed operations and country-wide programs (see for example, Claessens, Pohl, and Quian, 1996; Arias, 1994; Cassen, 1994; Isham and Kaufman, 1995, Kilby, 1996, Lele and Nabi, 1992). This paper is most closely related to the third category. It attempts to identify the quantitative contribution of economic analysis to the quality of individual investment projects and the size of overall resource flows. Unfortunately, data availability limits the scope of the study to assistance provided by the World Bank.

The impact of economic analysis on project success has been investigated in two other studies with conflicting conclusions. Arias (1995) fails to find a significant effect of ESW (defined as the number of World Bank reports on the country published in the three years preceding project approval) on project quality. This finding is based on an analysis of all projects that were *evaluated* between 1991 and 1994.

[1] As explained below, the lending program is defined as the total number of projects in a country approved in a given year.

Analyzing projects *approved* between 1991 and 1994, and also defining ESW as the number of applicable World Bank reports issued in the three years before project approval, Schneider (1995) finds a significant quality-enhancing impact of ESW. These contrasting results are a further motivation for the present exercise.

While the basic issues to be investigated are the same, this paper differs from the above studies in two respects. First, we try to cast the problem in terms of a more explicit decision-making framework. And second, both data quality and coverage are significantly expanded. Using data on input measured in staff weeks of analytical services rather than the number of written reports allows us to reduce measurement errors that may have affected other studies. And relying on all available information in the World Bank's data bases facilitates an expansion of the domain of the study to projects that were approved after 1985.

## II. CONCEPTUAL FRAMEWORK

In this section we describe the processes governing the allocation of resources to different activities within the World Bank in the form of two decision-making models. This, in turn allows us to derive the equations for estimation and to specify empirically testable hypotheses. We start with the model of the country manager responsible for all activities in a given country and then turn to that of a task manager at the project level.

Model 1: Country Manager

We assume that the country manager's objective is to maximize a weighted function of lending program quality *Q* and policy impact *P* subject to the total staff time (T) available for Economic and Sector Work (ESW) and all lending services -- both investment lending (INV) and adjustment operations (ADJ). Formally, the country manager solves the following problem:

$$\text{Max} \quad L = U\,(Q\,(ESW, INV\ ..);P(ESW, ADJ\ ..)) - \lambda\,(ESW + INV + ADJ - T)$$

Solution of this problem will yield demand functions for ESW, INV, and ADJ contingent on the overall budget constraint as well as various exogenous factors -- for example, country-level policies and institutional factors. In principle, these demand functions, together with the exogenous constraints, could be used to solve for the optimum values of Q and P.

Empirically, however, it is impossible to distinguish between the parameters of the "production functions" for lending quality and policy impact (*Q* and *P*), and the relative weights that the decision-maker assigns to each of these outcomes. Nor do we have a variable to measure policy impact. To answer the

questions posed earlier, therefore, we estimate a reduced form equation for $Q^*$, the optimum value of lending quality. Using a linear approximation, we estimate the following equation.

$$Q^*_{jt} = \mu_0 + \mu_1 ESW_{jt} + \mu_2 INV_{jt} + \mu_3 POL_{jt} + \nu_{jt} \qquad (1)$$

where *j* and *t* stand for country and year respectively and we include variables to capture policies (POL) and may adjust for other unobservable country-specific factors by using dummies.

In estimating this equation one notes that while the observed value of ESW can be taken to be a reasonably accurate estimate of the decision-maker's choice, this is not so for INV. Instead, what is available in our data is the amount of resources that was actually spent on INV (preparation and supervision); this may differ from the country manager's plans if "emergencies" or external events cause management to allocate additional funds to redesigning the project during preparation or to restructuring the project during implementation. Furthermore, because task managers may have access to (negative) information about individual projects that is not available to the country manager, or because of a general desire not to be seen to cancel projects, more time may be systematically allocated to bad projects in practice than was intended by the country manager. To deal with this, we instrument for INV using anticipated project age, past supervision, and past preparation requirements as well as past project success in the country.

Equation 1 allows us to examine all three questions of interest. Regarding the first one, we expect that, if ESW contributes to the quality of the lending program, $\mu_1 > 0$ (Hypothesis 1). Second, resources allocated to ESW compete with resources for INV, both of which should in principle improve project quality[2]. ESW, however, also contributes to policy impact -- through its impact on economy-wide and sector-specific policies and institutions, ESW generates benefits beyond the projects supported by the World Bank. Therefore, if staff resources are allocated to maximize the contribution to development, the marginal contribution of ESW to the quality of the lending program plus its marginal contribution to policy impact should equal that of INV. As long as the impact of ESW on policy is positive, it follows the marginal contribution of ESW to project quality should be less than that of INV. That is, if the World Bank has systematically maximized development impact, we should expect $\mu_1 < \mu_2$ (Hypothesis 2). And finally, if the "idealized" model set out above is supported by the data, then we can conclude that managers have not compromised quality to increase the volume of lending. On the other hand, if the model is not

[2] As, according to our definition, the lending program consists of all the projects approved in a given year, this would imply some kind of intertemporal reallocation -- that is, higher expenditures on current ESW that would be rewarded by lower spending on INV in the future.

supported by the data, then such a trade-off may underlie the observed outcome. This issue -- our third question -- is taken up in Section 4.

Model 2: Task Manager

We now turn to the decision faced by a task manager of a single project. His or her decision differs from that of the country manager in three respects. First, ESW is taken as exogenous. That is, we can reasonably assume that decisions about ESW (macroeconomic analysis as well as sector-specific analysis) are taken by senior management rather than task managers associated with individual projects and, moreover, are generally taken well in advance of the assignment of individuals to the preparation of specific projects. Second, the task manager of an investment project has little control over economy-wide policies. We assume, therefore, that the manager at the project level focuses exclusively on project quality. To the extent that the project design does include some changes in policy, we assume that their primary impact is on the quality of the project itself. And third, we divide INV into preparation time (PRP) and supervision time (SUP) both of which are assumed to contribute to project quality. PRP and SUP are treated as substitutes -- a well prepared project reduces the need for subsequent supervision, and more supervision is required to rescue a badly prepared project -- so that the task manager chooses the levels of PRP and SUP to maximize the quality of his or her project subject to the given shadow price of staff time. In the event that this price is high, the project task manager will be obliged to cut corners and economize on staff input. In the event that the price is low, then the manager will be able to prepare and supervise the project more thoroughly.

In general then, the task manager's optimization problem is:

$$\text{Max} \quad L = Q(PRP, SPN; ESW, POL, C, S ....) - \lambda (PRP + SPN).$$

where $\lambda$ is the shadow value of time derived from the country manager's optimal solution, C and S represent country-specific and sector-specific factors respectively, and other variables are defined as above. Solution of this problem yields estimable demand functions for PRP and SPN in terms of the exogenous parameters. We use linear approximations and estimate the following two equations where both PRP and SPN are normalized by loan size:

$$PRP^*_{ijt} = \beta_0 + \beta_1 ESW_{jt} + \beta_2 PPRP_{jt} + \beta_3 PQUAL_{jt} + \beta_4 POL_{jt} + \beta_5 S_i + \nu_{jit} \qquad (2)$$

and

$$SPN^*_{ijt} = \gamma_0 + \gamma_1 ESW_{jt} + \gamma_2 LOAN_{jt} + \gamma_3 PQUAL_{jt} + \gamma_4 PSPN_{jt} + \gamma_5 POL_{jt} + \gamma_5 S_i + \nu_{jit} \qquad (3)$$

where in the preparation equation PPRP represents average past preparation requirements in the country and PQUAL is past quality of the country portfolio. In the supervision equation, LOAN is loan size, and PSPN is past supervision effort in the country and sector.

Equations 2 and 3 provide a partial test of the assumed decision-making model. The coefficients of greatest interest are $\beta_1$ and $\gamma_1$. If $\beta_1 < 0$, average preparation requirements are reduced by the availability of prior ESW, implying that the task manager does respond to the availability of ESW in a "rational" fashion consistent with the assumed decision-making process. Similarly, finding that $\gamma_1 < 0$ would provide empirical confirmation for the hypothesis that managers respond to the higher quality of projects at approval (because of more prior ESW) by reducing the allocation of resources to supervision. We can also use these equations to test for several other possibilities. In particular, if only $\beta_0 \neq 0$ or $\gamma_0 \neq 0$ and $\beta_i, \gamma_i = 0$ for all i, then preparation and supervision time are allocated merely on the basis of loan size. On the other hand, if only $\beta_2$ (or $\gamma_4$) are different from zero, preparation (supervision) would be determined solely on the basis of past experience. We expect that unfavorable policies would complicate both project preparation and implementation -- that is, $\beta_4 < 0$; $\gamma_5 < 0$. Similarly, higher past requirements for preparation and supervision at the country and sector level would indicate the absence of institutional and administrative infrastructure, thus increasing the need for staff input ($\beta_2 > 0$). Higher quality of the past country portfolio is expected to be associated with lower preparation requirements because existing infrastructure and talent can be drawn upon ($\beta_3 > 0$; $\gamma_3 > 0$).

If the model posited above is appropriate, *Q* will depend on PRP, SPN, and ESW according to the following equation (where we again use a linear approximation):

$$Q_{ijt} = \mu_0 + \mu_1 ESW_{jt} + \mu_2 PRP_{jt} + \mu_3 SPN_{jt} + \mu_4 POL_{jt} + \mu_5 S_i + \nu_{jt} \qquad (4)$$

Estimation of this equation using the predicted values of PRP and SPN from equations 2 and 3 allows us to assess the marginal impact of ESW on project quality (hypothesis 1). In particular, if $\mu_1 > 0$, then we can conclude that ESW does improve the quality of projects. Indeed, if one uses a quantitative indicator for project quality that can be expressed in dollar terms, then we can measure the benefits of additional ESW and compare them with the nominal cost of staff input. In addition, we can utilize project level data to provide further evidence to investigate whether a reallocation of resources between ESW and preparation or supervision would improve project quality. Recall that because ESW makes a contribution to non-project objectives, we would expect $\mu_1 < \mu_2$ and $\mu_3$ (hypothesis 2). If this hypothesis is rejected we would have

further evidence that the country manager had failed to achieve the appropriate allocation of resources across the three activities from the standpoint of development impact.

## III. MEASUREMENT ISSUES AND DESCRIPTIVE STATISTICS

Before discussing our empirical results, we describe data definitions and sources, and present descriptive statistics. We begin with the variables used to measure lending services and ESW as well as project success. Then we discuss bivariate relationships between the variables that will be of importance in the subsequent discussion.

### Lending Services and ESW

*Lending services* are provided by the World Bank through projects that are associated with loans. In our analysis we consider a variety of project attributes such as the loan size, the amount disbursed, and the amount of resources spent in preparation and supervision of the project. We consider only loans that are either active or completed, thus discarding projects (about 10 percent of the total) that were dropped before becoming active. The amount of staff time (including long-term consultants) spent in preparation and supervision by project is available from the Bank's Management Information System for all years following 1975[3]. Unless otherwise noted, preparation and supervision variables in all the regressions are normalized by loan size -- that is, the shorthand "preparation" or "supervision" refers to staff-weeks of preparation or supervision time per million US $ lent.

*ESW* is defined as any use of resources by the Bank (excluding research and other activities that are neither country-specific nor project-specific[4]) that is not linked to the preparation or supervision of loans. Main activities included in this category are macro-economic and sector reports, public expenditure reviews, preparation of country assistance strategies, and donor coordination. We distinguish between general macroeconomic or multi-sectoral analysis and sector specific ESW, with the latter disaggregated into seven sectors -- agriculture and environment, education, health, finance, industry and energy, transport infrastructure and water, and public sector restructuring.[5] The variable for ESW is measured in weeks of staff time and is available on a consistent basis from 1985. As the variable on "staff cost" that is available

[3] The variables for time spent in preparation and supervision are based on entries from the Bank's time recording system where staff time devoted to specific activities is classified by project and by task code. The only three task codes encountered for lending projects were preparation, supervision, and preparation of the Implementation Completion Report (which we added to the time for supervision).

[4] There are about 26,000 staff weeks per annum of activities -- mainly research and a few administrative functions -- that are neither project-specific nor country-specific and that are therefore not included in our analysis.

[5] As the sectoral classification in the Bank's Management Information System is quite poor -- many sector specific pieces are classified as "general" -- we reclassified the "general" category based on the task title (for example, an education sector review would be placed into the "education" rather than the "general" category).

in the Bank data base does not add any information (its correlation with staff weeks is 0.82) and is available for a significantly lower number of observations, we measure ESW in staff weeks throughout.

We link ESW, defined at the country-level or sector-level, to specific lending projects by assuming that a project is affected by a simple average of the country and sector work undertaken during the four fiscal years preceding its Board approval. To account for the fact that macroeconomic work is likely to affect all projects in a given country while sector-specific ESW would have a more limited impact, the ESW variable is defined as the sum of all macroeconomic ESW for the country in which the project is implemented and all sector-specific ESW for the sector in which the project falls using the seven sectors mentioned above.

Table 1 illustrates that since 1975 the Bank has on average spent about 53,000 staff weeks (equivalent to about 1,260 staff-years) annually on supervision and preparation of loans, compared to about 22,000 staff weeks invested in country-specific macroeconomic (40 percent) or sector-related ESW (60 percent). Concerning the evolution of these activities one notes a considerable increase in the amount of resources devoted to supervision and sector-specific analysis, with much more modest increases in preparation of loans and in macroeconomic work. While total processing costs associated with lending stayed more or less constant at 3.2 staff weeks per million lent from 1977 to 1992, the composition changed considerably in favor of more supervision and less preparation; the share of time spent in project preparation dropped from 68 percent to 58 percent. The share of ESW in total staff weeks spent was consistently below 30 percent and, even though a significant increase in spending on ESW occurred between 1989/92 and 1993/95, in absolute terms the increase was less than the expansion of spending on lending during the same period (0.46 compared with 0.7 staff weeks per million lent), implying that in relative terms the importance of ESW actually decreased slightly.

A *breakdown by sector*, for the 1985-90 and the 1991-96 periods, highlights that the relative constancy of real commitments in the aggregate hides important intersectoral variations (Table 2). Regarding lending volume, considerably reduced levels of commitments in industry and energy and in agriculture (down $0.93 and $0.71 billion per annum respectively) are more than compensated by increases in lending for education, health, and public sector reform (up $0.92, $0.67, and $0.54 billion respectively). While the ranking of sectors was little affected -- the largest volume of lending is still for infrastructure, followed by industry and energy and by agriculture and environment -- the relative importance of education and health has more than doubled.

In 1985-90, direct lending costs, in staff-weeks per million lent, were above average in health (7.3), education and agriculture (4.4), and public sector reform (3.9), compared to below average costs in infrastructure (2.9), industry (2.3), and financial sector operations (2.0). Turning to intertemporal changes, the cost of lending -- normalized by total lending volume -- decreased in health and education and increased considerably in agriculture where, in contrast to industry and energy, the decrease in overall commitments was not matched by a concomitant reduction of staff weeks. While ESW increased in every sector, the increase was, in quantitative terms, largest in infrastructure where the amount of resources used for such services almost tripled. In the last period, the share of total sectoral resources (per million lent) spent on ESW ranged from a high of 28 percent in industry and energy to a low of about 16 percent in health and agriculture.

Turning from a sectoral to a *regional disaggregation* (Table 3), it emerges that the largest increase in commitments was registered by Eastern and Central Europe followed by East Asia and the Pacific (up $1.44 and $0.96 billion per year). By contrast, annual real lending to South Asia dropped by more than one billion, with smaller decreases in Latin America and Africa. The share of ESW in total use of staff time during 1991-95 was about 20 percent for almost all regions except Latin America (15 percent). This implies that staff time spent directly in both lending services and ESW varied widely across countries from a total of more than 10 staff weeks per million lent in Africa to less than 3 staff weeks in East Asia. Comparatively high costs of lending were also observed in South Asia (a sharp increase from the 1985-90 period) and in the Middle East and North Africa.

Measures for Project Success

Direct indicators regarding the success of individual lending projects are available from ex post evaluations conducted by the World Bank's Operations Evaluation Department (OED) and from the Annual Review of Portfolio Performance (ARPP). While the former are preferred insofar as they provide an assessment of the overall performance of a completed project, the larger availability of data for ARPP indicators leads us to use them as an alternative variable to check the robustness of our results although the corresponding regressions are not reported.

*OED indicators* include an estimate of the economic rate of return at completion, a binary measure (satisfactory and unsatisfactory) of the overall project outcome, an ordinal assessment (4 levels from highly satisfactory to highly unsatisfactory) of the project's sustainability, and its contribution to institutional development in the client country. While these indicators provide a succinct summary of overall project success, their disadvantage is that only completed projects undergo this evaluation. As information on

resources devoted to non-lending services is available on a consistent basis only after 1985, the number of projects for which both OED evaluations and data on non-lending services are available is substantially reduced. However, this drawback may be more than compensated for by the fact that the evaluation is undertaken by an independent unit, presumably with some minimum consistency across projects.

*The Annual Review of Portfolio Performance (ARPP)* rates projects annually according to a number of criteria relating to overall project performance (contribution to development objectives and implementation progress) as well as more specific "intermediate" characteristics (compliance with legal covenants, project management performance, availability of funds, procurement progress, training progress, technical assistance progress, studies progress, environmental aspects, fiscal performance, and gender aspects)[6] on an ordinal scale from 1 to 4. Analytical rigor and informational content of these measures are likely to be inferior to more precise indicators such as the economic rate of return. However, the much increased coverage partly compensates for this shortcoming, with a total of more than 5000 projects having been included in the ARPP at least once.[7]

Differences in average values for selected indicators of project success are illustrated in Table 4. Data on economic returns after completion are available for about 2000 projects in four sectors (agriculture and environment, industry and energy, infrastructure, and finance). Average ex post rates of return vary between 13.8 percent in agriculture and 20.9 percent in infrastructure and between 14.7 percent in Africa and 19.0 percent in South and East Asia. Ex post ratings concerning contribution to institutional development and sustainability are highest in infrastructure and education projects. Contemporaneous (ARPP) ratings, while characterized by limited variation across projects, are lowest in agriculture and in Africa, while being consistently high in East Asia. Table 5 provides simple correlations between these indicators, illustrating that ARPP indicators for achievement of development objectives and implementation progress, and OED indicators for project outcome, sustainability, and institutional development, as well as the ex post rate of return are significantly correlated with each other.

Policy Variables

The literature has long emphasized that policies are important factors influencing project implementation as well as eventual development impact. It is therefore important to control for appropriate

[6] Not all of these are regularly rated, resulting in large variations in the number of observations available for different indicators.

[7] At the project level, the advantage of greater data availability is less impressive than might appear at first sight. While using the mean value of all available ratings for any given project to construct a project-specific performance indicator increases the number of observations, it lumps together projects at very different stages of their life cycle whereas use of the final rating (for completed projects) considerably reduces the number of observations. As discussed below, use of ARPP variable may be more meaningful at the level of the lending program where it enables us to compare the success of specific project-cohorts.

policy variables. Policies that have an important bearing on project success relate to fiscal policy, monetary policy, and external trade policy. These can be approximated by the public sector deficit; domestic inflation or the black market exchange rate; and openness, as in Burnside and Dollar (1996). To avoid problems of endogeneity, we use the value of policy variables at project initiation throughout. Note that policy variables are generally available only up to 1992, causing us to omit them in the instrumental equations for supervision.

Preliminary Statistical Analysis

Investigation of the bivariate relationship between OED success indicators for the lending program and our independent variables (Table 6) leads to three broad conclusions. First, the correlation between ESW (both macroeconomic and sector-specific) and development impact is significant and positive except in the case of project outcome. Second, higher levels of resources spent on supervision and preparation are negatively, rather than positively, related to project success. Such a phenomenon could arise from selection bias whereby bad projects are allocated more resources in order to improve their performance. Alternatively, resources spent in either preparation or supervision may be correlated with another omitted variable (from the above discussion one could suspect, for example, an Africa dummy) that is associated with lower project performance. And third, concerning the impact of policy variables, we find that openness is consistently associated with improved project success; that sustainability and contribution to institutional development are better in countries with low public sector deficits; that inflation affects negatively project outcome and sustainability; and that projects appear to perform better in all dimensions except economic rate of return in countries that invest a higher proportion of their own resources.

Table 7 uses descriptive statistics to summarize our hypotheses and to provide empirical support for the conjecture of a positive link between ESW and project success. Splitting the sample of completed projects into groups based on the overall project outcome, the variable used below in our analysis, one finds that the satisfactory group of projects benefited from considerably higher levels of ESW, but required significantly lower levels of preparation and supervision (per million lent) than the rest. Although key policies for the satisfactory projects were more favorable, with more open economies, lower public sector deficits, lower inflation and black market premia, and higher levels of investment than for the unsatisfactory group, the fact that unsatisfactory projects required 20 percent more preparation and 22 percent more supervision suggests that the ex-ante design of a project -- an activity to which ESW

presumably contributes -- is crucial. Similar results (not reported) are obtained if projects are split into top and bottom groups for the other indicators discussed earlier[8].

## IV. EMPIRICAL ANALYSIS OF THE IMPACT OF ESW

In this section we report the results of the reduced form regressions discussed in section two. The goal is to examine three issues, namely (i) whether, and if so by how much, ESW improves project quality; (ii) whether the allocation of staff resources across their three main uses -- ESW, preparation, and supervision -- has been consistent with the maximization of development impact; and (iii) whether the allocation of staff resources to ESW reduces the volume of lending, presumably by reducing the amount of staff time available for the more immediate tasks of project processing.

### Estimating the Country Manager Model

We first estimate the reduced form equation (1). As we are interested in the overall program outcome, we use the mean "satisfactory/unsatisfactory" rating for all projects (weighted by loan size)[9] as the dependent variable. Where included, INV is instrumented using past preparation and supervision requirements in the country, past project success, and the mean project age in a given country and year.

Results (Table 8, column 1) suggest that ESW has a significant positive impact on the quality of the lending program, thus supporting hypothesis 1. This effect remains significant if policy variables are added to the regression (column 2)[10]. Thus the use of better data provides empirical evidence in favor of a significant and positive impact of ESW on quality, in contrast to the findings of Arias (1995). Regarding the quantitative impact of ESW, we find that an additional 100 staff weeks would increase the probability of a program being satisfactory by between 12 and 20 percent. While other resources spent on lending -- supervision and preparation -- are insignificant, the point estimate is positive (column 3)[11]. And while the policy variables have the anticipated sign (and can be shown to be significant in regressions when ESW is excluded), they are not significant. Using other OED variables to represent project quality suggests that the results obtained are robust (not reported). While policy variables are rarely significant (only public sector surplus in the sustainability equation), ESW is highly significant throughout.

[8] Although similar descriptive results hold if ARPP indicators are used, we refrain from using them below because the only advantage of ARPP over OED ratings -- namely, that projects that are not yet completed are included -- is outweighed by the difficulty of appropriately aggregating ratings and by the fact that they may be subjective and prone to changes due to shifts in personnel.

[9] Results do not change much if the unweighted average outcome of the lending program is used.

[10] As our measure for ESW includes the years before project approval while policies here are measured at project approval, there is a potential endogeneity problem. Repeating the analysis with policies in earlier years reduces the number of observations but does not alter the substantive conclusions (results not reported).

[11] The point estimate for INV is negative, though insignificant, in a specification that excludes policies (not reported). This suggests that, with the instruments available, we were not able to fully resolve the endogeneity problem associated with project preparation and supervision.

The regressions using the ERR as the dependent variable have the added advantage of facilitating a simple cost-benefit analysis of assigning more resources to ESW. There are, however, three problems with using ERRs. First, ERRs are only available for certain sectors (see table 4) and -- even within these sectors -- ERRs have not been computed consistently for all projects. This implies that whatever estimate is available is unlikely to be representative for the entire lending program and problems of sample selection bias may be present. Second aggregation of ERRs over projects (even if the latter are weighted by loan size) is at best a rough approximation to the true net present value (NPV) of the lending program. And third, the number of observations is quite limited (305 and 164 without and with policy variables respectively). While this prevents us from reading too much into the figures obtained in the respective regressions, we note that ESW is significant and positive even if policy variables are included (Table 8, columns 3 and 4) and that introduction of region-dummies does not eliminate the significance of the coefficient on ESW. In fact, using the point estimate on the ESW variable in these regressions, we find that a one-staff week increase in ESW would increase the NPV of the average lending program (US $210 million) by between US $36,000 and US $44,000[12], similar to the higher end of the range of values (US $18,000 - US36,000) derived from the project level regression reported below. Compared with the average cost of one staff week (about US $3,000) as reported in Bank data, this suggests that an expansion in ESW would constitute an economically sound investment.

When interpreting this result, however, it is important to keep in mind the distinction between an improved NPV for the projects supported by the Bank and an improvement in the NPV of the government's investment program. If ESW enables Bank staff to identify and support new investment possibilities thereby expanding the range of investment options open to a country, or if it leads to an improvement in the ex-ante design of a project already in the investment program and hence its ultimate development impact, then ESW is yielding a genuine return to the recipient country. Alternatively, conducting more ESW may enable Bank staff to "cherry-pick" better projects from the existing portfolio, projects that would anyway have been undertaken by the country. In this event, ESW improves the performance of the Bank's lending portfolio but not necessarily the quality of the public investment portfolio in general. The second explanation, however, is not convincing for two reasons. First, in most cases countries are not in the position of having a well defined investment program from which Bank staff can pick and choose. Bank

[12] The transformation from ERR to NPV depends on the phasing of costs and benefits. Following Kilby (1996a), the calculation reported in the text assumes that project costs occur in the first period and benefits in the second. With this assumption and an average lending program of US$210 million, the increase in NPV lies between US$0.00019x210,000,000/1.1 and US $0.00023x210,000,000/1.1, where the discount rate is assumed to be 10

staff are usually heavily engaged in identifying and developing new investment opportunities. And second, in most cases projects are not close to being fully designed before Bank staff become involved. In fact, Bank staff usually play a major role in project design and conceptualization.

A more formal test of these alternative explanations is possible. If the second explanation -- cherry-picking -- underlies the observed impact of ESW on the NPV of the Bank-supported lending program, one would expect Bank portfolios supported by significant inputs of ESW to comprise exclusively "safe" projects. Cherry-picking implies that there would be few, if any, *dropped* projects -- either during preparation or after approval. Thus, the cherry-picking hypothesis leads us to expect an inverse relationship between the amount of ESW and the share of the program dropped. No such relationship is found in the data -- controlling for policies and level of GDP per capita, there is no statistically significant relationship between the share of the lending program dropped prior to the loan becoming effective and the amount of ESW, nor between the share of the portfolio canceled after approval and the amount of ESW (results not reported). This result suggests that, rather than enabling task managers to pick the plums out of a limited portfolio of investments, ESW contributes by either expanding the available investment opportunities or improving the ex-ante design of the available projects. Both avenues result in a genuine increase in the development impact of foreign assistance.

Regarding a possible trade-off between ESW and INV as stated in hypothesis 2, we find that, contrary to expectation, the marginal contribution of ESW exceeds that of INV. Indeed, the marginal contribution of INV, while positive, is not significantly different from zero (table 8, columns 3 and 6). This suggests that it would have been possible to shift resources from INV to ESW and increase program quality. And since ESW also -- we assume -- contributes to policy impact, we conclude that the observed allocation of staff time between these two uses has not been optimal from the point of view of maximizing development impact. In other words, our "ideal" model of decision-making at the level of the country manager may not be a good approximation to reality.

Of course, the representative manager may not have been concerned exclusively with development impact as we have assumed. Other objectives are clearly possible. For example, managers may have given weight to resource flows or lending volume. This could reflect a concern with the overall level of financial transfer (a development concern) or, if volume of commitments was the basis for rewards within the

---

percent. A more realistic assumption regarding the phasing of costs and benefits would tend to increase the calculated NPV. We therefore treat the value reported in the text as a conservative estimate.

institution, a career concern[13]. Either way, this suggests that the objective function outlined in Section 2 needs to be expanded to include a third element -- lending volume -- in addition to the quality of lending and policy impact. Corresponding to equation 1, we now have an additional reduced form equation for the optimal lending volume consistent with our revised objective function for the country manager. Again assuming a linear approximation, the new equation to be estimated is:

$$V^{*}_{jt} = \lambda_0 + \lambda_1 ESW_{jt} + \lambda_2 INV_{jt} + \lambda_3 POL_{jt} + \lambda_4 ADJ_{jt} + \nu_{jt} \qquad (5)$$

where as before, ESW represents time spent on ESW, INV is staff time devoted to lending services excluding adjustment operations, ADJ is the staff time for adjustment operations, and POL represents policy variables. Focusing on the difference between $\lambda_2$ and $\lambda_1$, we can examine whether a reallocation of staff resources in favor of ESW would have a negative impact on lending volume (hypothesis 3). If $\lambda_2 < \lambda_1$, managers would have been operating within the frontier -- that is, they could have increased both quality and volume of the lending program by shifting resources out of INV into ESW. If, on the other hand, $\lambda_2 \geq \lambda_1$, reallocating staff from lending services to ESW -- regardless of its impact on quality -- would have come at the cost of decreased lending volume. This would imply a tradeoff between quality and quantity, indicating that mangers did indeed place some value on volume of lending as well as quality.

We use as dependent variable both the volume of commitments and the volume of actual disbursements for a given lending program. The former would be the better variable if the representative manager believed that approval of a large lending program by the Executive Board was a key factor in promotions and career development. The latter is more appropriate if the manager was interested in achieving a large resource transfer.

Our results (Table 9, column 1) support the hypothesis of a significant difference between the effect of ESW and INV on the volume of commitments -- INV appears to be between 40 and 50 percent more effective in increasing commitments than is ESW. If the observed relationship is, indeed, the outcome of an optimization process on the part of the decision-maker, one can assess the trade-off between volume and quality quantitatively. Ignoring for the moment the possibility that ESW has a positive policy impact, at the optimum the decision-maker will be indifferent to a marginal shift of staff time between INV and ESW. Using the estimated results for equations 1 and 5, we calculate that the average manager seems to be indifferent between a one-percent decrease in the program's ERR (or a US $2 million decrease in the

[13] Claessens, Pohl, and Quian (1996) find a significant impact of task managers on project quality but note that such quality of task managers is not correlated to promotion decisions within the Bank.

program's NPV) and an additional US $4 million of commitments (compared with an average lending program of $210 million). This suggests the presence of a strong trade-off -- World Bank managers are willing to accept a moderately large decline in the quality of the lending program in return for only a modest increase in the level of commitments. Of course, when making this decision, the country manager should also take account of the contribution of ESW to policy, an impact that is additional to its immediate effect on the quality of the lending program. If this is true, and the contribution of ESW to policy impact is quantitatively important, then our unfavorable interpretation of managerial behavior is strengthened.

The picture with respect to disbursements is very different. Here the evidence suggests that mangers were operating within their frontier -- resources could have been switched from INV to ESW to the benefit of program quality and the volume of disbursements. In this case, mangers could have achieved a one-percentage point increase in the program's ERR (an increase in the NPV of about US $2 million) and increased disbursements by almost US $7 million. The different outcomes are consistent with the view that managers did give weight to the level of commitments, even though this reduced the quality of the lending program and reduced the level of disbursements.

Estimating the Task Manager Model

We turn now to the decision-making problem faced by the task manager of a single project. Recall that in this case ESW is treated as an exogenous variable and the manager is assumed to maximize project quality given the (shadow) price of staff inputs.

*Project Preparation*

As discussed in Section 2, the assumed decision-making process yields demand functions for preparation and supervision contingent on existing ESW. We examine the results for preparation first. According to the model, a higher level of ESW is expected to reduce the need for preparation resources, because, as a result of ESW, the basic project rationale and design will be on a much firmer footing. Using total staff time spent on preparation (normalized by loan size) as the dependent variable and estimating equation (2) confirms this notion. In particular we find that the representative manager clearly and significantly reduces the amount of resources used to prepare any given project as ESW increases (Table 10). This result is consistent with the view that managers see a benefit from ESW and adjust their resource allocations accordingly. Regarding the magnitude of this impact we find that, at the mean level of ESW across all projects (136 staffweeks), one additional staff week of ESW causes the task manager to reduce the time allocated to preparation by 0.015 staff weeks per million lent or 1.05 staff weeks for the mean $70

million loan (table 10, column 3). We also note that at empirically observed input levels the impact of ESW on preparation is concave -- that is, it is subject to slowly decreasing returns (columns 1,2, and 3). Returns from ESW start declining at almost 1,000 staffweeks on average per project, a value that is well beyond the mean (136 staffweeks). As we are dealing with individual projects, we add dummies to control for unobservable sector-specific or country-specific characteristics. Their inclusion leaves the statistical results unchanged.

*Project Supervision*

According to our decision-making model, ESW should -- by improving the design of projects -- enable managers to reduce the amount of time allocated to supervision and thus have a negative effect on resource requirements for supervision (equation 3). Table 11 provides partial support that this is indeed the case; higher levels of ESW are clearly associated with lower inputs into supervision except when preparation is included as an explanatory variable, a point we return to below (column 3). The other regressions reported in Table 11 support the notion that -- whether consciously or not -- task managers recognize the benefits from ESW and, in line with the proposed maximization model, adjust the allocation of staff time to supervision accordingly. The quantitative magnitude of the effect of ESW on supervision is lower than in the case of preparation; a one staff-week increase in ESW is associated with a 0.009 staff week decrease in supervision per million lent, equivalent to a reduction of supervision requirements by 0.64 staff weeks for the mean US $70 million project (column 1). Furthermore, it is much cheaper to implement projects in countries with an appropriate policy environment -- a 3 percentage point decrease in the public sector deficit, for example, reduces the amount of supervision needed by 0.27 staffweeks per million lent or 56.7 staffweeks for the mean lending program of US $210 million (column 2).

Taken together with the results for project preparation, this implies that assigning one more staff week to ESW would lead task managers to reduce the time allocated to project preparation and supervision per project by about 1.7 staffweeks over the entire life of the project. Moreover, ESW actually benefits more than one project. If we assume that macroeconomic work benefits all projects in the lending program (3 on average), and that sectoral analysis only benefits projects in that sector, we calculate that each staffweek of ESW benefits 1.5 projects on average[14]. Thus the saving in preparation and supervision for the entire lending program generated by one staffweek of ESW amounts to 2.5 staffweeks. A focus on lending volume that comes at the cost of project quality (as evidenced at the country level) does, therefore, seem to increase the costs associated with administering World Bank loans.

*Project Quality*

The above evidence implies that task managers recognize the positive impact of ESW on lending quality. This hypothesis can be tested directly by estimating equation 4 at the project level for a wide range of independent variables and specifications. The variables of greatest interest are the project outcome rating (satisfactory-unsatisfactory) and the ERR (estimated after project completion).[15] From the probit equation of project outcome, one notes that ESW does significantly increase the probability of a satisfactory project outcome, even if other policy variables are included (Table 12, columns 1,2, and 3). An increase of 100 staffweeks allocated to ESW increases the probability of a successful project by between 9 and 13 percent, comparable to the estimate obtained at the program level. In line with the literature we find that better policies -- lower levels of inflation and a more open economy -- also contribute significantly to greater project success and are quantitatively quite important.

Turning to the ERR, one has to bear in mind the fact that ERRs are not available for projects in the social sectors, and that a project's benefit to the borrowing country is determined by the size of the loan as well as the ERR, so the net present value (NPV) would be a more appropriate measure. Unfortunately, the cost and benefit streams underlying the ERR calculation are available only for very few projects so we cannot reproduce NPVs. That said, our results suggest that ESW causes a considerable increase in the ERR for a typical project (Table 12, columns 4, 5, and 6), regardless of whether the specification includes policy variables, instrumented preparation and supervision, country and/or sector dummies or not (hypothesis 1). A one staffweek increase in ESW before project initiation increases a project's rate of return by between 0.02 and 0.04 percentage points.

For the mean loan of US $70 million in the sample of completed projects, this result translates into an increase in NPV of between US $12,000 and US $25,000.[16] And since each staffweek of ESW includes macroeconomic and sectoral analysis, it should, as noted above, benefit more than one project in the lending program. Using the earlier calculation -- that one staffweek of ESW benefits 1.5 projects -- implies an increase in NPV for each additional staffweek of ESW of between US $18,000 and US $36,000, a result that is very similar to the one obtained when the lending program was the unit of analysis. To interpret the estimated benefits from ESW, it is worth noting that, in our sample, the reported cost of a staff

[14] For the average investment project, the ratio of sector-specific to macroeconomic ESW is about 3:1.

[15] We also utilized ARPP ratings (either the average over the whole project live or the final rating for a completed project) to construct a large number of quality indicators that were then used as dependent variables. As the results from these regressions do not add any substantively new insights we do not report them separately.

[16] Again following Kilby (1996a), with an average project cost of US$70 million ,a staff week increase in ESW before project initiation increases net present value of an average project by between US $0. 0002x70,000,000/1. 1 and US $0.0004x70,000,000/1.1, where the discount rate is assumed to be 10 percent..

week is slightly below US $3,000.[17] Thus, if the sample of projects available here is any guide, putting additional resources into ESW would be an economically attractive option. The fact that ESW is similarly significant in a probit regression of project outcomes suggests that this result is not driven by the sectoral limitations of the availability of ERRs.

Since equation 4 captures our central result, we subjected it to three tests of robustness. In the first, we changed the specification -- from linear to log-linear. With or without policy variables, the coefficient on ESW in the log-linear specification with either project outcome rating or ERR as the dependent variable was positive and significant at least at the 5 percent level (results not reported). In the second, we took advantage of new information that became available late in our study. The results reported thus far cover projects that were evaluated before July, 1996. Having completed the estimation for projects up to this date, information on additional projects -- those evaluated up to February, 1998 -- became available. Armed with this additional information, we reran equation 4. The results (not reported) show that with or without policy variables, the coefficient on ESW remains positive and highly significant in the equation with project outcome rating as the dependent variable. In the equation with ERR as the dependent variable the coefficient is positive and significant at the 5 percent level without policies and at the 10 percent level with policies.

The third test entails expanding the righthand side variables to include some others that have been found in the literature to have an impact on proejct outcomes These include indices of country risk, institutional quality and bureaucratic delays (Knach and Keefer, 1997), and the number of revolts, coups, and assassinations to capture political stability (King and Levine, 1996). The results (not reported) for both the program-level and project level regressions reveal that the coefficient on ESW never changes sign and remains significant in the majority of cases at the 10 percent level. Thus, the results are robust to these tests of specification, sample selection, and choice of independent variables.

Although the task manager does not have control over the time allocated to ESW, we can still use equation 4 to test whether the country manager achieved the right allocation between lending services and ESW by exploring whether a shift in resources from preparation and supervision to ESW would have resulted in a better project outcome (hypothesis 2). Table 12 (columns 3 and 6) indicates that the marginal contribution of resources spent in supervision (instrumented) is positive but not statistically significant, while the contribution of resources spent on preparation is also positive but insignificant in the equation for

[17] Although this should include travel cost, the latter may not have been reported consistently in our data sources. Even a generous adjustment for this will, however, not change the substantive results,

the rate of return, and negative and marginally significant in the equation for project outcome. We interpret this last result as an indication that, with the instruments available, we were not able to control for of the endogeneity of these to variables, and not as an indication of a negative impact of preparation on project success. The general result is consistent with our results at the program level -- an increase in the allocation of resources to ESW would have been associated with an overall improvement in project quality and presumably with a positive impact on policy.

As we saw earlier, there is a negative raw correlation between project quality and the amount of resources used in supervision (or preparation). We also note that supervision requirements are highly correlated with resource use in preparation (preparation alone explains about 55 percent of the variation in allocations to supervision). This suggests that managers do not behave as though preparation and supervision are substitutes as we have assumed -- to the contrary, projects that consume substantial amounts of resources during preparation are generally associated with more than proportional allocations to supervision. This may reflect a different decision-making model. Instead of a simultaneous decision-making process undertaken by a single task manager, we may be observing a sequential one undertaken by two or more managers. In the first step, staff time is allocated to preparation given the availability of ESW to achieve maximum project quality at time of Board approval, reflecting the fact that the project and its quality is subject to considerable scrutiny at this time and staff are rewarded for performing well. In the second step, supervision resources are allocated -- often by a different task manager -- based on the project's quality at approval to maximize project quality by the time project implementation is completed. In this variant, the task manager who is preparing an ill-conceived project is obliged to allocate more time to preparation than would be the case for a project with a better conceptual foundation. But he or she is not willing to allocate enough time to reduce future supervision requirements, because someone else will be responsible for project implementation. Hence, both preparation and supervision tend to be larger than normal for such projects. This suggests that the reduction in preparation that occurs when more ESW has been undertaken will carry over to a reduction in supervision as well.

Our interpretation of this is that, while there are certainly cases -- for example, in the development and piloting of untested or relatively complex approaches -- where higher spending on preparation (and supervision) may be justified, high preparation costs may also be an indication of a failure to have carried out the necessary analytical background work or to have engaged in a constructive process of consensus building with the main stakeholders involved, and as a consequence a lack of government commitment. Experience as well as our results suggest that in such situations it might be advisable to first prepare the

ground with ESW before moving ahead with a lengthy and prolonged process of project preparation. Or, where this has not been done, the expected high costs of supervision should be borne in mind when decisions are being made on whether to drop or to continue preparation of projects that have incurred large costs during preparation.

The fact that, even if instrumented, project preparation and supervision are never significantly different from zero at the 5 percent level of significance suggests that, at the program as well as the project level, higher spending on preparation or supervision does not automatically increase quality. This failure to identify a positive contribution of preparation and supervision on project success does not mean that, at any given point in time in a project's life, higher spending on supervision will not increase the probability of an improvement in project quality as demonstrated by Kilby (1995). The difference is that Kilby looks at changes between successive ARPP ratings -- thus eliminating project-specific fixed effects -- while our focus is on the quality of the project as a whole. We therefore conclude that greater attention to ESW -- in the form of elaborating the broader context, clarifying the rationale for a specific type of intervention, and assessing its feasibility and economic desirability against a broader set of potential alternatives -- could have resulted in the design of better projects, and thus less need for spending time on preparation or supervision.

## V. CONCLUSION

At the start of this paper we posed three questions. In these concluding remarks, we report our answers to each question. We also try to explain why we get the results that we do in terms of the behavior of managers and point to where changes are needed if the pursuit of development impact is to be our exclusive goal.

In answer to the first question, we find that ESW improves the quality of World Bank projects whether we look at lending programs for a whole country or at individual projects. One telling statistic in support of this observation is that spending an additional dollar on ESW yields an increase in the development impact of the average lending program for a country in the order of twelve to fifteen dollars. The second question we posed had to do with the allocation of resources across different activities. Our results at both the program and the project level indicate clear evidence of underinvestment in ESW. And finally, we looked at the impact of a reallocation of resources in favor of ESW on lending volumes. Here we find that time allocated to preparation and supervision is between 40 and 50 percent more effective in increasing commitments than ESW, but that resources could have been switched from preparation and supervision to ESW to the benefit of both program quality and the level of disbursements.

How do we explain these results? Identifying the behavioral relationships underlying any set of data is a difficult task, and in this case we can only point to some possibilities which are at least consistent with the data. We note two points. First, there is evidence that broad allocations of staff resources across different activities have been influenced by a concern with commitments -- our representative country manager appears to be indifferent between a one-percentage point decrease in the program's ERR and an additional $4 million of lending volume. As a result, country managers have chosen to allocate resources in favor of preparation at the expense of ESW. Second, there is evidence to suggest that task managers clearly perceive the benefit of ESW and respond by reducing the amount of time allocated to preparation and supervision. Our results suggest that managers reduce preparation and supervision by about 2.5 staffweeks if one more week is allocated to ESW. Together, these results are consistent with the view that task managers factor *availability of ESW* into their decisions regarding the allocation of resources at the project level at least to some extent, but that country managers systematically *underinvest in ESW* at the program level. The compensating behavior of task managers -- partial maximization in the small -- is not able to offset the fact that country managers have failed to maximize at the global level.

There is good news and bad news in these results. The good news is that our analysis points to areas of improvement which could yield substantial benefits in terms of the development impact of our lending and the total resource transfer that the World Bank provides to developing countries. The bad news is that our analysis of behavior suggests that in the past managers have had a strong interest in the level of commitments. This suggests a need for action in two areas. First, to continue to strengthen the reward system such that managers will pay even more attention to development impact. And second, to educate managers to the importance of ESW in improving project quality.

**References**


Averch, Harvey (1994), "Economic Approaches to the Evaluation of Research", Evaluation Review, Vol. 18, No. 1.

Arias, Omar (1994), "Results of the Country Econometric Analysis for the Annual Review of Evaluation Results 1994 for the World Bank, Operations Evaluations Department," World Bank, mimeo.

Burnside, Craig, and David Dollar (1996), "Aid, Policies and Growth" World Bank, mimeo.

Cassen, Robert and Associates (1994), Does Aid Work? Report to an intergovernmental Task Force, Oxford, Clarendon Press.

Feyzioglu, Tarhan, Vinaya Swaroop, and Min Zhu (1995), "The Impact of Foreign Aid on Public Spending," World Bank, mimeo.

Isham, Jonathan, and Daniel Kaufman (1995), "The Forgotten Rationale for Policy Reform," Policy Research Working Paper 1549, the World Bank.

Isham, Jonathan, Deepa Narayan, and Lant Pritchett (1995), "Does Participation Improve Performance? Establishing causality with Subjective Data," *The World Bank Economic Review* 9 (2):175-200.

Keefer, and Knack

Kemp, Murray C., and Kar-yiu Wong (1993), "Paradoxes associated with the administration of foreign aid," *Journal of Development Economics* 42:197-204.

King, and Levine

Kilby, Christopher (1995), "Supervision and Performance: The Case of World Bank Projects," Center Working Paper 9545, Tilburg University.

Kilby, Christopher (1995), "World Bank-Borrower Relations and Project Supervision," Working Paper # 32, Vassar College Poughkeepsie, New York.

Kilby, Christopher (1994), "Risk Management: An Econometric Investigation of Project-level Factors," Background Paper for the Annual Review of Evaluation Results 1994, World Bank, mimeo.

Lele, Uma and Nabi, I. Eds (1992), Aid to African Agriculture: Lessons from tow decades of donor's experience, Baltimore, Johns Hopkins University Press.

Mosley, Paul, Jane Harrigan, and John Toye (1995), Aid and power: The World Bank and policy-based lending, New York Routledge.

National Science Foundation (1995), Report of NSF Stakeholders' Panel in Linguistics.

Pack, Howard and Janet Rothenberg Pack (1990), "Is Foreign Aid Fungible? The Case of Indonesia," *The Economic Journal* 100:188-194.

Pack, Howard and Janet Rothenberg Pack(1993), "Foreign Aid and the Question of Fungibility," *Review of Economics and Statistics*: 258-265.

Rodrik, Dani (1995), Why is there multilateral lending? NBER Working Paper 5160, Cambridge, MA

Schneider, R (1995), “Economic Sector Work and Results on the Ground ,” World Bank, mimeo.

Svensson (1996), When is foreign aid policy credible? Aid Dependence and Conditionality, World Bank, Policy Research Department, mimeo.

Trumbull, William N. , and Howard J. Wall, (1994), “Estimating Aid-Allocation Criteria with Panel Data,” *The Economic Journal* 104:876-882.

Wise, Lois Rescasino, and Robert Agranoff (1991), “Organisational Characteristcs and Productivity Measurement in Research Organisations”, Public Productivity and Management Reveiw, vol. XV, no. 1.

World Bank (1991), Report on the World Bank Research Program, World Bank, mimeo.

Table 1: Evolution of lending and non-lending services

| | Commitments (annual; 1987 1987 US $) | Lending (Staff Weeks) | | | Non-Lending (Staff Weeks) | | |
|---|---|---|---|---|---|---|---|
| | | Preparation | Supervision | SW/Mn lent | Macro | Sector | SW/Mn lent |
| 1975-76 | 12.14 | 24265 | 11326 | 2.93 | | | |
| 1977-80 | 14.28 | 29267 | 16443 | 3.20 | | | |
| 1981-84 | 16.03 | 28657 | 21359 | 3.12 | | | |
| 1985-88 | 17.32 | 31819 | 22783 | 3.15 | 6917 | 8604 | 0.90 |
| 1989-92 | 19.09 | 36039 | 25172 | 3.21 | 7152 | 11819 | 0.99 |
| 1993-95 | 18.26 | 36702 | 34621 | 3.91 | 9310 | 17225 | 1.45 |
| TOTAL | 16.47 | 31512 | 22359 | 3.27 | 9143 | 13336 | 1.36 |

Source: Own calculations based on Bank MIS data as described in the text.

Table 2: Resources spent on lending and non-lending services by sector

| | ***1985-1990 Period*** | | | | |
|---|---|---|---|---|---|
| | Annual Commitments | Lending | | Non-lending | |
| | billions 87 US $ | SW spent | SW/Mn lent | SW spent | SW/Mn lent |
| Agric. & Env't | 3.81 | 16853 | 4.42 | 2535 | 0.67 |
| Pub Sector Reform | 0.25 | 972 | 3.89 | 255 | 1.02 |
| Education | 0.92 | 4092 | 4.45 | 964 | 1.05 |
| Health | 0.39 | 2815 | 7.28 | 664 | 1.72 |
| Industr/Energy | 4.37 | 10230 | 2.34 | 2828 | 0.65 |
| Infrastructure | 4.22 | 12083 | 2.86 | 1170 | 0.28 |
| Finance | 2.08 | 4052 | 1.95 | 1045 | 0.50 |
| Total | 16.04 | 51097 | 3.19 | 9460 | 0.59 |
| | ***1991-1995 Period*** | | | | |
| | Annual Commitments | Lending | | Non-lending | |
| | billions 87 US $ | SW spent | SW/Mn lent | SW spent | SW/Mn lent |
| Agric. & Env't | 3.10 | 18694 | 6.02 | 3688 | 1.19 |
| Pub Sector Reform | 0.79 | 3282 | 4.14 | 696 | 0.88 |
| Education | 1.84 | 7488 | 4.07 | 1792 | 0.97 |
| Health | 1.06 | 6543 | 6.19 | 1220 | 1.15 |
| Industr/Energy | 3.43 | 9576 | 2.79 | 3728 | 1.09 |
| Infrastructure | 4.39 | 14491 | 3.30 | 3117 | 0.71 |
| Finance | 1.48 | 3675 | 2.48 | 1080 | 0.73 |
| Total | 16.09 | 63749 | 3.96 | 15320 | 0.95 |

Source: Own calculations based on Bank MIS data as described in the text.

Table 3: Resources spent on lending and non-lending services by region

| | ***1985-1990 Period*** | | | | |
|---|---|---|---|---|---|
| | Annual Commitments | Lending | | Non-lending | |
| | billions 87 US $ | SW spent | SW/Mn lent | SW spent | SW/Mn lent |
| Africa region | 2.81 | 18519 | 6.58 | 3270 | 1.16 |
| East Asia/Pac | 3.60 | 8785 | 2.44 | 1982 | 0.55 |
| MENA | 1.25 | 5103 | 4.09 | 880 | 0.71 |
| LAC | 4.99 | 10007 | 2.00 | 1542 | 0.31 |
| ECA | 1.54 | 3623 | 2.35 | 471 | 0.31 |
| South Asia | 3.77 | 9261 | 2.45 | 1314 | 0.35 |
| Total | 17.97 | 55298 | 3.08 | 9460 | 0.53 |
| | ***1991-1995 Period*** | | | | |
| | Annual Commitments | Lending | | Non-lending | |
| | billions 87 US $ | SW spent | SW/Mn lent | SW spent | SW/Mn lent |
| Africa region | 2.55 | 21152 | 8.29 | 4975 | 1.95 |
| East Asia/Pacific | 4.56 | 10531 | 2.31 | 2508 | 0.55 |
| MENA | 1.25 | 4864 | 3.88 | 1260 | 1.01 |
| LAC | 4.60 | 12322 | 2.68 | 2322 | 0.51 |
| ECA | 2.98 | 9026 | 3.03 | 1983 | 0.67 |
| South Asia | 2.58 | 11184 | 4.33 | 2272 | 0.88 |
| Total | 18.52 | 69079 | 3.73 | 15320 | 0.83 |

Source: Own calculations based on Bank MIS data as described in the text.

Table 4: Indicators of project success by sector and region

| | Mean values | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | All | Agr. Env. | Pub Sect. | Educ. | Health | Ind. Energy | Infrastr. | Finance |
| ***OED INDICATORS*** | | | | | | | | |
| ERR at completion | 16.81 | 13.83 | | | | 15.45 | 20.88 | 17.14 |
| Project Outcome (Sat/Unsat)1 | 71.71 | 62.52 | 48.00 | 82.33 | 60.32 | 74.69 | 77.97 | 75.08 |
| Inst. Development | 2.69 | 2.44 | 2.38 | 3.19 | 2.59 | 2.87 | 2.78 | 2.76 |
| Sustainability | 2.40 | 2.34 | 2.16 | 2.53 | 2.22 | 2.50 | 2.38 | 2.54 |
| ***ARPP INDICATORS*** | | | | | | | | |
| Development Objectives | 3.39 | 3.25 | 3.21 | 3.44 | 3.25 | 3.49 | 3.50 | 3.42 |
| Implementation Progress | 3.10 | 2.98 | 3.07 | 3.17 | 3.01 | 3.16 | 3.14 | 3.13 |

| | Mean values | | | | | | | No of obs. |
|---|---|---|---|---|---|---|---|---|
| | All | Africa | EAP | MENA | LAC | ECA | SAS | Total |
| ***OED INDICATORS*** | | | | | | | | |
| ERR at completion | 16.81 | 14.73 | 18.95 | 17.82 | 15.72 | 16.95 | 18.97 | 1959 |
| Project Outcome (Sat/Unsat)1 | 71.71 | 61.82 | 82.83 | 77.61 | 68.86 | 75.49 | 76.98 | 3956 |
| Inst. Development | 2.69 | 2.32 | 3.22 | 2.94 | 2.67 | 2.94 | 2.69 | 2057 |
| Sustainability | 2.40 | 2.18 | 2.79 | 2.57 | 2.37 | 2.60 | 2.29 | 1992 |
| ***ARPP INDICATORS*** | | | | | | | | |
| Development Objectives | 3.39 | 3.25 | 3.56 | 3.47 | 3.40 | 3.36 | 3.42 | 5004 |
| Implementation Progress | 3.10 | 3.04 | 3.30 | 3.10 | 3.07 | 3.04 | 3.07 | 5064 |

Source: Own calculations based on Bank MIS data as described in the text.

[1] The figure indicates the share of projects that are rated "satisfactory".

Table 5: Simple correlations between different indicators of project success

| | | ***OED INDICATORS*** | | | ***ARPP INDICATORS*** | |
|---|---|---|---|---|---|---|
| | Proj. outcome (Sat/Unsat) | ERR at evaluation | Project Sustainability | Institutional Development | Development Objectives | Implementatio progress |
| Project outcome | 1 | | | | | |
| | 3957 | | | | | |
| ERR at evaluation | 0.3807*** | 1 | | | | |
| | 1376 | 1376 | | | | |
| Sustainability | 0.6381*** | 0.3107*** | 1 | | | |
| | 2024 | 842 | 2031 | | | |
| Institutional Development | 0.4829*** | 0.1477*** | 0.5084 | 1 | | |
| | 1956 | 799 | 1951 | 1968 | | |
| Development Objectives | 0.2035*** | 0.1522*** | 0.1791*** | 0.1355*** | 1 | |
| | 2885 | 1371 | 2024 | 1960 | 4806 | |
| Implementation Progress | 0.2134*** | 0.1741*** | 0.2447** | 0.2184*** | 0.5719*** | 1 |
| | 2928 | 1375 | 2028 | 1965 | 4805 | 4855 |

Note: ***denotes significance at the 1% level,** at the 5 %level and * at the 10% level throughout
Source: Own calculations based on Bank MIS data as described in the text.

Table 6: Correlations between program-level OED indicators and policy variables

| | Economic Rate of Return | Project Outcome | Project Sustainability | Insitutional Development |
|---|---|---|---|---|
| Analytical services (ESW) | 0.1538*** | 0.0715 | 0.1331** | 0.1237*** |
| | 305 | 347 | 601 | 595 |
| ... macro only | 0.13368** | 0.05633 | 0.1269*** | 0.1102*** |
| | 305 | 347 | 601 | 595 |
| ... sectoral only | 0.1525*** | 0.07473 | 0.1432*** | 0.1215*** |
| | 305 | 347 | 601 | 595 |
| Preparation | -0.1139*** | -0.0507* | -0.0499* | -0.0214 |
| | 1299 | 1080 | 1500 | 1474 |
| Supervision | -0.1218*** | -0.0751** | -0.0903*** | -0.0309 |
| | 1299 | 1080 | 1500 | 1474 |
| Openness | 0.0495* | 0.0914*** | 0.1303*** | 0.0978*** |
| | 1225 | 1008 | 1391 | 1368 |
| Public sector surplus | 0.0361 | 0.0488 | 0.0993*** | 0.0830** |
| | 706 | 587 | 894 | 881 |
| Inflation | -0.0438 | -0.0955*** | -0.0575** | -0.0241 |
| | 1095 | 927 | 1287 | 1269 |
| Investment | -0.01121 | 0.1174*** | 0.1999*** | 0.1554*** |
| | 1225 | 1008 | 1391 | 1368 |

Source: Own calculations based on Bank MIS data as described in the text.

Note: Lower number of observations is due to the use of portfolio-level data.

Table 7: Input and policy indicators for satisfactory and unsatisfactory projects

| | Project Outcome | |
|---|---|---|
| | Sat. | Unsat. |
| SW spent in ESW1 | 88.32 | 74.33 |
| SW per $m spent in preparation | 4.88 | 5.84 |
| SW per $m spent in supervision | 4.75 | 5.82 |
| Openness | 54.11 | 48.53 |
| Public sector surplus | -6.09 | -6.57 |
| Inflation | 23.56 | 49.69 |
| Black market premium | 28.37 | 35.16 |
| Investment | 14.97 | 12.71 |

Source: Own calculations based on Bank MIS data as described in the text.

Table 8: Impact of ESW on project outcome and economic rate of return at the program level

| Dependent var. | Project outcome (Sat/Unsat) | | | | | |
|---|---|---|---|---|---|---|
| ESW | 0.12104** | 0.1760** | 0.1773* | 1.933*** | 2.240** | 2.1605* |
| | (0.0529) | (0.0753) | (0.1008) | -0.7132 | (0.9319) | (1.243) |
| Openness | | 0.36271 | 0.3778 | | -2.262 | 0.6309 |
| | | (0.3394) | (0.4178) | | (4.647) | (4.778) |
| Inflation | | -0.2451 | -0.3372** | | -0.0778 | -0.7225 |
| | | (0.1403) | (0.01717) | | (0.5883) | (0.5187) |
| Public Sector Surplus | | 0.8937 | 0.8037 | | -2.917 | -37.215 |
| | | (1.538) | (1.707) | | (2.964) | (0.31215) |
| Lending | | | 0.0054 | | | 0.0709 |
| | | | (0.0221) | | | (0.2227) |
| Log-likelihood/$R^2$ | -224.99 | -136.52 | -98.96 | 0.02 | 0.05 | 0.12 |
| No of observations | 347 | 215 | 157 | 305 | 196 | 164 |

Table 9: Determinants of lending volume (disbursements or commitments in millions of US $)

| | Commitments | Disbursements |
|---|---|---|
| ESW | 0.1899*** | 0.2118*** |
| | (0.039) | (0.04465) |
| Investment lending | 0.2709*** | 0.0727** |
| (prep. and supervision) | (0.029) | (0.0314) |
| Adjustment lending | 0.1401*** | 0.1432*** |
| (prep. and supervision) | (0.0338) | (0.0324) |
| No of observations | 406 | 302 |
| Adj R-Sq | 0.9090 | 0.7248 |

Note: Coefficients of policy variables and country dummies included but not reported

Table 10: Determinants of Projects' Preparation Requirements

| Dependent Variable: Preparation per million lent | | | | |
|---|---|---|---|---|
| ESW | -1.1784*** | -1.1907*** | -2.0273*** | -0.2847*** |
| | (0.143) | (0.222) | (0.171) | (0.076) |
| $ESW^2$ | 0.0011*** | 0.0012*** | 0.0019*** | |
| | (0.000) | (0.000) | (0.000) | |
| Public Sector Surplus | | -5.2725* | | |
| | | (2.876) | | |
| Inflation | | -0.0408 | | |
| | | (0.040) | | |
| Openness | | 0.9206* | | |
| | | (0.485) | | |
| Country performance | | | | 65.5442 |
| (relative to region) | | | | (51.208) |
| Past proj. success[1] | -40.8768*** | -39.9245* | -60.7319*** | -42.7433 |
| | (15.185) | (22.790) | (14.018) | (32.492) |
| Past preparation[1] | 55.9327*** | 52.0307*** | | 67.2006*** |
| | (3.753) | (5.789) | | (6.239) |
| Adj R-Squared | 0.1984 | 0.1606 | 0.1115 | 0.2224 |
| No of observations | 2355 | 1239 | 3194 | 737 |

[1] At the country level

Coefficients and Standard Errors multiplied by 100

Sector dummies included but not reported

Table 11: Determinants of Projects' Supervision Requirements

| Dependent Variable: Supervision per million lent | | | |
|---|---|---|---|
| ESW | -1.5812*** | -1.0446*** | -0.2464 |
| | (0.300) | (0.320) | (0.250) |
| $ESW^2$ | 0.0024*** | 0.0014** | 0.0004 |
| | (0.001) | (0.001) | (0.001) |
| Public Sector Surplus | | -9.0527*** | -7.6614*** |
| | | (2.872) | (2.222) |
| Inflation | | 0.0769 | 0.0781 |
| | | (0.062) | (0.048) |
| Openness | | -1.2403** | -0.8083* |
| | | (0.543) | (0.420) |
| Loan Size | -2.0057*** | -1.8145*** | -0.7596*** |
| | (0.145) | (0.151) | (0.123) |
| Preparation | | | 63.4220*** |
| | | | (2.410) |
| Past proj. success[1] | -73.1693*** | -60.5990** | -57.4011*** |
| | (19.044) | (23.937) | (18.519) |
| Adj R-Squared | 0.2134 | 0.228 | 0.538 |
| No of observations | 1639 | 1045 | 1045 |

[1] At the country level

Sector dummies included but not reported

Table 12: Impact of ESW on project outcome and economic rate of return

| Dep. Var | Project Outcome | | | | Rate of Return | |
|---|---|---|---|---|---|---|
| ESW | 0.0973** | 0.0900** | 0.1302** | 2.3737* | 4.2285*** | 2.3439* |
| | (0.387) | (0.407) | (0.063) | (1.4047) | (1.6638) | (1.3990) |
| Public sector surplus | | 0.7192 | 0.8727 | | 14.9735 | 42.0955 |
| | | (0.9632) | (1.365) | | (53.6540) | (43.8989) |
| Inflation | | -0.0371* | -0.06002** | | 1.9165 | 4.90144 |
| | | (.0214) | (.0214) | | (10.2482) | (5.798) |
| Opennness | | 0.7391*** | 0.4729* | | -3.4949 | -1.0494 |
| | | (0.018) | (0.2564) | | (21.8876) | (5.241) |
| Preparation (instr) | | | -0.5611* | | | 42.0955 |
| | | | (0.3192) | | | (139.257) |
| Supervision (instr) | | | 0.02102 | | | 5.1640 |
| | | | (0.2649) | | | (103.772) |
| No of projects | 1324 | 873 | 643 | 481 | 302 | 200 |
| R-Squared/LL | -825.56 | -534.62 | -394.045 | 0.146 | 0.142 | 0.1109 |

Notes: Coefficients and standard error multiplied by 100
Sector dummies included but not reported