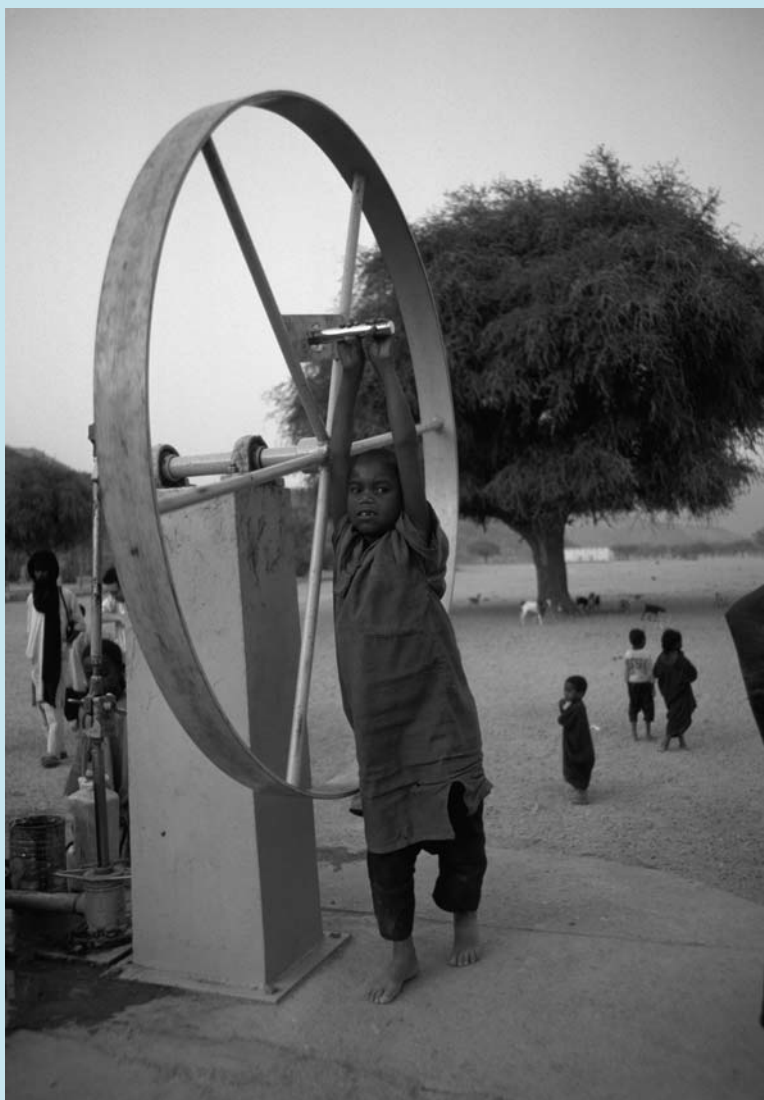


Chapter 2

Evaluation Highlights

- Project outcome ratings improved over the medium term.
- Attention is needed to address the growing lack of realism in assessments of ongoing project performance.
- Unsatisfactory projects in fiscal 2007 were hampered by overly ambitious objectives, complexity, and inadequate design.
- Many country programs produced satisfactory outcomes, especially those in large countries where many of the world's poor live.
- Overall, however, some 40 percent of country programs have been moderately unsatisfactory or worse in outcomes.



Tuareg child turns water pump wheel; photo ©Yann Arthus-Bertrand/Corbis, reproduced by permission.

Development Outcomes: Indicators of Performance

Over the past five years, the Bank's lending has risen in real terms, and IDA made record commitments of \$12 billion in fiscal 2007, alongside some \$13 billion of International Bank for Reconstruction and Development (IBRD) finance provided to middle-income countries (MICs). Indeed, with the successful 15th replenishment round for IDA resources (\$42 billion), the Bank plans to maintain total annual lending in the \$22 to \$26 billion range during fiscal 2008–10.

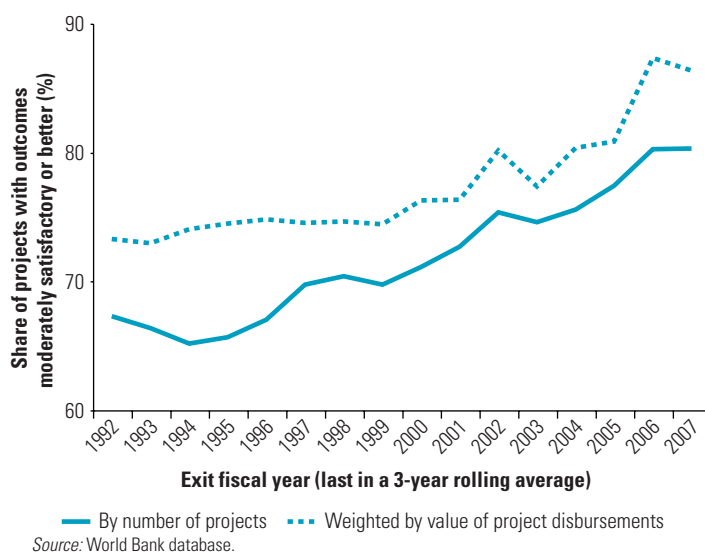
Lending volumes by themselves, of course, are not an indicator of development impact, and there have been many cautions for the Bank (and other development agencies) to take care to avoid a “lending culture” which places undue emphasis on volumes delivered rather than outcomes secured.¹ There are indicators of performance that can tell us about the Bank's development impact—notably IEG evaluations of Bank-supported projects and country programs. Overall trends and recent developments revealed in both of these metrics are presented below.

Measuring Project Performance: Trends from IEG Monitoring

How are Bank-supported projects performing in terms of achieving their development objectives? The performance of Bank projects in delivering development results has unquestionably improved over the medium term.² In the three years to end-fiscal 2007, 80 percent of projects were moderately satisfactory or better in delivering their targeted results, up from around 70 percent

at the start of the decade, as shown in *Lending outcomes have improved over the medium term.* figure 2.1.

Figure 2.1: Project Performance Has Improved over the Medium Term



Project ratings moved up the scale in recent years.

A comparison of five-year subperiods shows 78 percent of projects from fiscal 2003 to 2007 rated as moderately satisfactory or better in achieving their development results. This is a significant improvement from fiscal 1998–2002, in which 73 percent of projects had outcomes rated moderately satisfactory or better. Illustrations of development impact from some of these projects are in box 2.1. Moreover, a growing share of projects has outcomes that are sustainable.³ Almost 80 percent of projects evaluated in fiscal 2003–07 were rated as likely to be sustainable, up from 64 percent in the previous five years.⁴ A more detailed assessment of project performance is available in appendix A.

Against the backdrop of this overall improvement in outcomes, there are several features that illustrate changes in the performance of Bank-supported projects, as shown in table 2.1. There has been a steep decline in the share of projects rated unsatisfactory while, simultaneously, the share reaching the satisfactory standard held steady. In addition, the Bank and its clients have improved a number of projects, which in the past may have been rated moderately unsatisfactory, to reach the level of moderately satisfactory. As a result, a greater proportion of projects is being rated in the middle of the scale.

Outcomes of projects in two-thirds of sectors improved, while health and public sector governance worsened.

The Bank's lending is spread across a spectrum of sectors. It is encouraging that, as shown in figure 2.2, project performance in about two-thirds of sectors has improved over the medium term.⁵ The turnaround in the water supply and sanitation sector projects is especially dramatic: in fiscal 1998–2002 only a little more than 60 percent of projects (by value of disbursements) were moderately satisfactory or better, and this sector now comes close to leading the way in fiscal 2003–07, with over 90 percent of projects having moderately satisfactory or better outcomes. While there has been a creditable and significant uplift in economic policy project performance, the share of moderately satisfactory or better project outcomes in fiscal 2003–07 still trails other sectors by quite some margin. The largest declines in performance were in health, nutrition, and population (which, along with economic policy and the environment, were significantly below the Bank-wide average for the fiscal 2003–07 period) and in public sector governance.

Looking at the Regional aspects of project performance, figure 2.3 shows that the East Asia and Pacific and the Europe and Central Asia Regions led the way for the fiscal 2003–07 cohort. Their share of projects with moderately satisfactory or better outcome ratings significantly exceeded the Bank average of 83 percent (weighted by disbursement). A good illustration

Box 2.1: What Does a Satisfactory Project Look Like? Illustrations of Development Impact

There is no single metric that can be used across projects to assess satisfactory outcomes. Rather, development results span a range of different social and economic indicators depending on the sector and type of project. A project is rated satisfactory when the operation's objectives have been achieved with only minor shortcomings. This is illustrated with one type of development impact in Mozambique, where Bank-supported water and sanitation projects built 130 water source points, bringing drinking water to 62,000 people. And in Cambodia, a Bank-financed project brought clean water to 750,000 people in Phnom Penh.

In the energy sector, a Bank electrification project in the Lao

People's Democratic Republic provided main-power-grid electricity to 51,000 households in 721 rural villages, while another 6,000 households gained access to off-grid electricity through solar and hydro systems. Having electricity greatly increased productivity of new small-scale, home-based businesses, and enabled children to study at night.

Finally, another type of development impact is demonstrated in Mali, where the Bank-supported Grassroots Initiatives to Fight Poverty and Hunger Project helped more than 6,000 children to attend school, as well as created small-scale health centers in 19 villages.

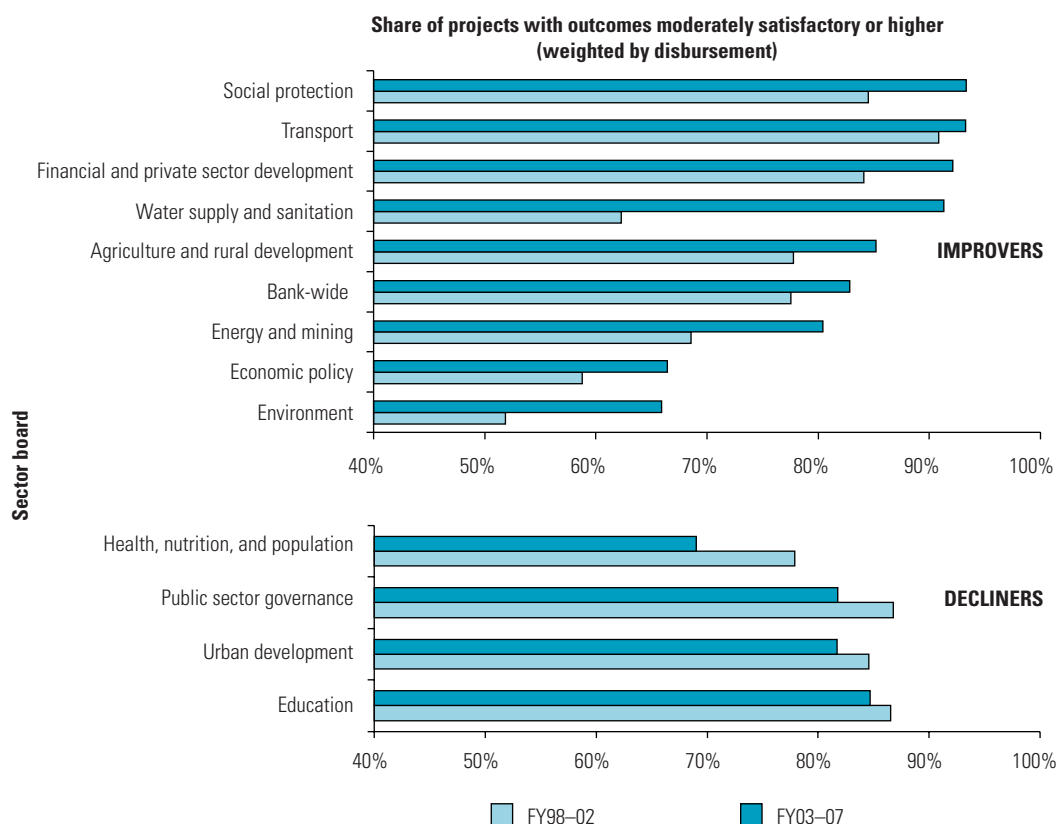
Table 2.1: Distribution of Project Ratings Moved Up the Scale in FY03–07

IEG rating of projects meeting development objectives	Percentage of Projects in Rating Category		
	FY98–02	FY03–07	Percentage point change
Highly satisfactory	6%	4%	-2
Satisfactory	45%	45%	0
Moderately satisfactory	21%	29%	+8
Moderately unsatisfactory	9%	10%	+1
Unsatisfactory	16%	10%	-6
Highly unsatisfactory	2%	1%	-1

Source: World Bank database.

of such successful performance can be seen in an agricultural development project in China's Anning Valley. The Bank's \$120 million of support supplied more reliable irrigation, introduced

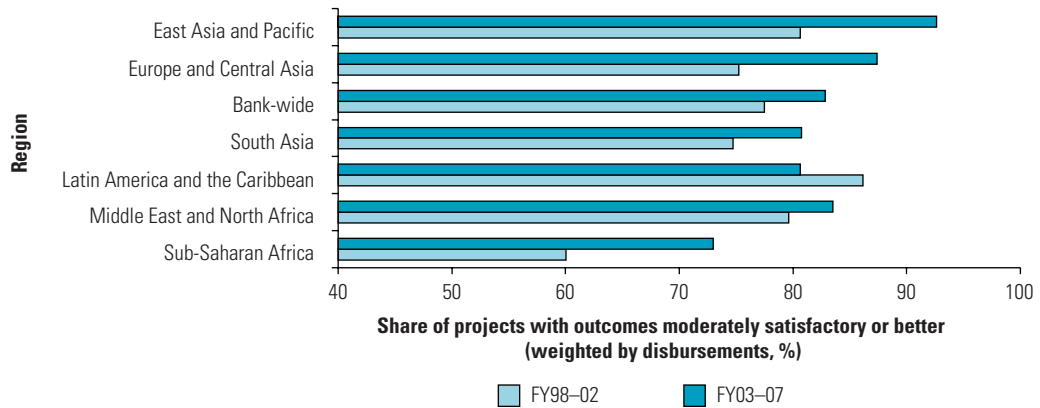
new plant varieties and cultivation techniques, and added value to production through processing. Local farmers in the project's area enjoyed a tripling in their per capita income; the propor-

Figure 2.2: Trends in Sectoral Performance

Source: World Bank database.

Note: The Sector Board classification applies to the whole project and enables outcomes to be matched to it.

Figure 2.3: Africa's Projects Have Improved Substantially but Still Lag Behind Other Regions



Source: World Bank database.

tion of the area's families living in poverty fell from one-third to one-tenth; and the project generated a 27 percent rate of return on its investment.

In these Regions, MICs make up a large share of the client base, and their relatively strong institutional capacity is conducive to stronger project implementation. Nonetheless, the Bank could well be faced with increasing demands from these clients to take on projects that are complex or located within a country's lagging or remote subregions, under difficult operating conditions. These are challenges for the Bank to embrace, but expectations for the success of projects must be adjusted to recognize concomitant risks.

Africa's performance improved the most of any Region but still lags behind other parts of the Bank. In the Africa Region, the share of projects with moderately satisfactory or better outcomes improved the most of any Region over fiscal

2003-07. Again, it should be acknowledged that difficult operating conditions can make it difficult to secure success—and that may at least partly explain why project performance, overall, in Africa continues to lag behind all other Regions. But the Bank has to find ways to deliver outcomes in adversity, as a central part of its mission. Indeed, this can be done, as illustrated in the Bank's \$220 million investment to assist war-affected households in

northern Ethiopia. By removing landmines, building new homes, and equipping households with basic goods and agricultural inputs—including seeds and fertilizer—some 67,000 families returned home, rebuilt their lives, and generated new economic activity, which delivered a 50 percent return on the project's investment.

Project Outcomes in Fiscal 2007 Data

Project development outcome ratings fluctuate, sometimes significantly, from year to year. Differences are sometimes a result of the vicissitudes of a particular cohort of projects rather than a substantive change in underlying institutional or partner performance. So while a single year's data should not be taken in isolation, the most recent data are worth examining to see whether any patterns or signals emerge. There are two movements in fiscal 2007 that warrant the attention of Bank management.⁶

First, increasingly the Bank is too optimistic in its own assessment of ongoing project performance. Of the 45 projects that exited in fiscal 2007 and were rated as moderately unsatisfactory or worse by IEG, over two-thirds—32 projects in total—had been reported by the Bank as moderately satisfactory or better just before they were closed. This failure to identify problem projects early impairs real-time managing for results, since overly optimistic ongoing ratings mean management is

less likely to take remedial action. Most Regions were too optimistic in their ongoing reporting of project ratings for fiscal 2007, with almost a fifth of all projects—and in Africa close to a third—downgraded during the process of self-evaluation and independent evaluation.⁷ Comparatively, these “disconnect” rates in previous years were much lower—less than 1 in 10 were downgraded in fiscal 2005 and 2006, as shown in table 2.2. Reducing this disconnect—by more carefully and accurately identifying underperforming projects during implementation—is needed to focus attention on project supervision and to improve performance. Of course, to subsequently improve development outcomes will then require project-specific actions on the ground.

Second, in fiscal 2007 there was a significant fall to 76 percent in the share of projects rated moderately satisfactory or better, from 83 percent in fiscal 2006.⁸ This absolute level of performance still meets the 75 percent target set by the Bank a decade ago, although it is lower than the 80 percent benchmark noted by the Bank’s own Quality Assurance Group (QAG) in recent annual reviews. If a 75 percent threshold for project portfolio performance is considered reasonable, given the risks of the “development business,” then this one-year drop, by itself, is not too worrisome. But certainly vigilance is needed to ensure that this drop does not foreshadow a persistent decline.

What has caused the drop in ratings in fiscal 2007 and produced 45 projects with outcomes moderately unsatisfactory or worse? Sometimes a difference in the performance of the Bank’s

portfolio from one year to the next can be influenced by a change in the composition of projects being evaluated—for example, if there is a larger share of projects in challenging sectors or countries. But the change in portfolio composition in fiscal 2007—related to Region, sector, instrument, lending arm, and other factors—does not explain the fall in ratings. Even if the fiscal 2007 cohort had maintained the same composition of lending to conflict-affected and postconflict countries as the fiscal 2004–06 cohorts, the result would not be materially affected. Appendix A shows the impact of various changes in the fiscal 2007 composition in more detail.

Another possible explanation is that the drop in measured project performance is due to methodological changes in the way projects were evaluated in fiscal 2007. Some changes in methods were introduced by IEG and the Bank together in fiscal 2007 to strengthen the robustness of project ratings and to cover new elements of project design. In the near term they may have introduced some element of discontinuity in the data series between fiscal 2007 and earlier years.⁹ It is estimated that the influence of methodology changes has been small—accounting for around 1 percentage point of the fall.

The final possibility is that in the fiscal 2007 cohort, there was simply a greater occurrence of five key factors influencing weak outcomes. First, poor or overly complex project design has been a problem in more than half of these underper-

Ongoing self-assessments of project performance are increasingly overly optimistic.

Project outcomes deteriorated in fiscal 2007.

Table 2.2: Disconnect between the Bank’s Self-Ratings and IEG Ratings Increased Dramatically in FY07

Source of rating	Share of projects rated moderately satisfactory or better		
	FY05	FY06	FY07
Bank’s final Implementation Status and Results report (ISR)	87.9%	91.1%	93.0%
Bank’s Implementation Completion and Results report (ICR)	85.7%	89.9%	83.3%
IEG’s ICR review	80.8%	82.6%	75.8%
<i>Difference between ISR and ICR review ratings—“disconnect”</i>	<i>–7.1%</i>	<i>–8.5%</i>	<i>–17.2%</i>

Source: World Bank database.

Five factors have influenced weak project outcomes. forming loans, a finding also made by the World Bank's QAG (World Bank 2008a). For instance, two projects failed to recognize the importance of an appropriate legal and regulatory framework as a precondition to a privatization process. Several health projects failed to ensure a heightened focus on those interventions that would yield the greatest impact, leading, for instance, to inadequate targeting of the poor, the absence of a cost effective package of health services, or inadequate funding for behavior change interventions to prevent HIV/AIDS transmission among high-risk groups.

Second, overambition was a weakness. While project objectives were almost always relevant, a majority were too far-reaching. Sometimes this was in terms of assessing political commitment and the feasibility of certain reforms. Other times it was in assessing government effectiveness and capacity, or in requiring coordination across several ministries or cumbersome financial management procedures that were not manageable by the parties involved. IEG's evaluation of public sector reform (2008c) shows several examples where these factors led to unsatisfactory outcomes.

Third, delays in implementation caused difficulties, because circumstances changed and project design or implementation could not respond. About one-fifth of underperforming projects suffered from this problem.

Fourth, a majority of the unsatisfactory projects had a weak results framework with poor or no baseline data, making it difficult to assess the outcomes of the project, and outcomes were often not well linked to inputs and outputs.

Finally, various gaps in the Bank's own performance contributed to a lack of success. For example, despite being flagged by the QAG for poor quality at the outset, three projects were not reassessed or redesigned. The quality of the Bank's supervision was rated as moderately unsatisfactory or worse in two-thirds of all underperforming

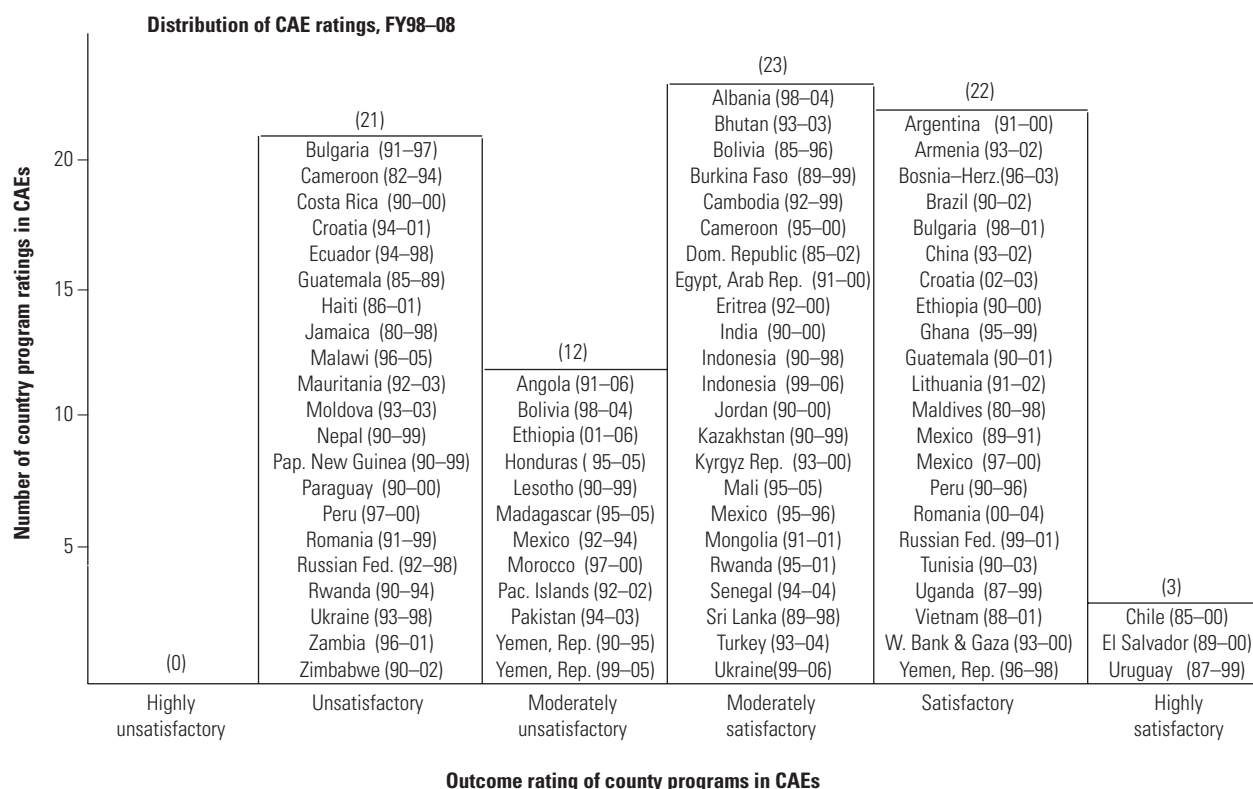
projects (and many such projects were not identified as problems in ongoing status reports). And Bank overall performance—as distinct from borrower performance or the effects of uncontrollable events—was ranked moderately unsatisfactory or worse in two-thirds of these problem projects, compared with only about one-fifth of the full sample. All of this points to a challenge in re-emphasizing a proactive quality control in management's attention to ongoing project performance.

How It Adds Up: Outcomes of Bank Country Programs

While Bank-supported projects have largely yielded positive development outcomes, country program outcomes—measured against their own objectives, which typically include growth, poverty reduction, and the environment—are far less satisfactory. Some previous assessments suggest that synergies between and among the Bank's lending, knowledge services, and dialogue are not fully exploited or that the projects are not always directly relevant to the country's core development challenges (World Bank 2008a; IEG 2006c).

The lower country-program ratings could also be influenced by other factors. The number and scale of the factors that affect broader country-level objectives is typically much greater than at the smaller-scale project level. Hence, outcomes measured at this level are more likely to be affected by pressures outside the control of development partners. It is also possible that objectives may be relatively more ambitious at the program rather than project level (and therefore less frequently attained) or indeed that it is more difficult to establish the connection between Bank inputs and outcomes at the program level (as opposed to at the project level).

Over the past 10 years, evaluations of 81 Bank country programs—incorporating projects, policy and technical advice, and other types of assistance—show that three-fifths of them were moderately satisfactory or better in meeting their development objectives. Looking at specific

Figure 2.4: CAEs Show Three-Fifths with Outcomes Moderately Satisfactory or Better

Source: IEG Country Assistance Evaluations.

Note: CAEs assess country program performance over a long period. Years in parentheses indicate the period that Bank country programs are assessed.

grades on IEG's ratings scale, as shown in figure 2.4 and table 2.3, the Bank succeeded in supporting satisfactory outcomes in 30 percent of evaluated programs; a further 30 percent of

country programs were rated moderately satisfactory. But the remaining 40 percent of programs—concentrated in countries that are smaller or have extensive poverty, such as

Table 2.3: Summary of CAE Ratings, FY98–08

	MICs	LICs	All countries
Highly satisfactory	6%	0%	4%
Satisfactory	33%	18%	27%
Moderately satisfactory	25%	33%	28%
Moderately unsatisfactory	13%	18%	15%
Unsatisfactory	23%	30%	26%
Highly unsatisfactory	0%	0%	0%
<i>Satisfactory and highly satisfactory</i>	<i>39%</i>	<i>18%</i>	<i>31%</i>
<i>Moderately satisfactory or better</i>	<i>64%</i>	<i>51%</i>	<i>59%</i>

Source: IEG Country Assistance Evaluations.

Box 2.2: What Does a Satisfactory Country Program Look Like?

No single template exists for a satisfactory Bank-supported country program, since such programs vary greatly depending on the country's institutional capacity, stage of development, and particular development needs. The characteristics of the Indonesia country program illustrate some features of good outcomes found in several cases.

During fiscal 1999–2006, the Indonesia country program was largely successful. Bank support provided help to national authorities in securing economic recovery from the late 1990s economic crisis. By 2004, per capita income achieved precrisis levels, inflation fell to under 7 percent, and public debt was cut from

nearly 100 percent of gross domestic product in 1999, to under 50 percent by 2005. These gains largely managed to turn around the effects of the crisis—the percentage of Indonesians living on less than one dollar per day was 16 percent in 2005, compared with 23 percent in 1999. The fall in income poverty and improvement in human development indicators were assisted by the Bank's work in community-driven development through the Kecamatan Development Program and by reconstruction following the 2004 tsunami. These achievements outweighed some areas in which outcomes from the Bank's program were less satisfactory—including in decentralization and fighting against corruption.

Outcomes of country programs are less satisfactory than project outcomes. Malawi—were moderately unsatisfactory or worse in meeting their stated development objectives.

Very few country programs are producing best-practice results—indeed, of 36 programs rated since fiscal 2002, not one has been highly satisfactory. And among the 14 low-income countries (LICs) rated since fiscal 2002, not a single program has been rated satisfactory or better. At the same time, no program has ever been rated highly unsatisfactory.

Many programs in the largest countries that house the majority of the world's poor—including Brazil, China, India, Indonesia, and Mexico—were rated moderately satisfactory or better in their development outcomes. Box 2.2 shows such outcomes from the Indonesia country program. When country programs are weighted by the number of people in poverty (living on less than \$2 per day), 86 percent of country programs are given a satisfactory rating. Even when excluding

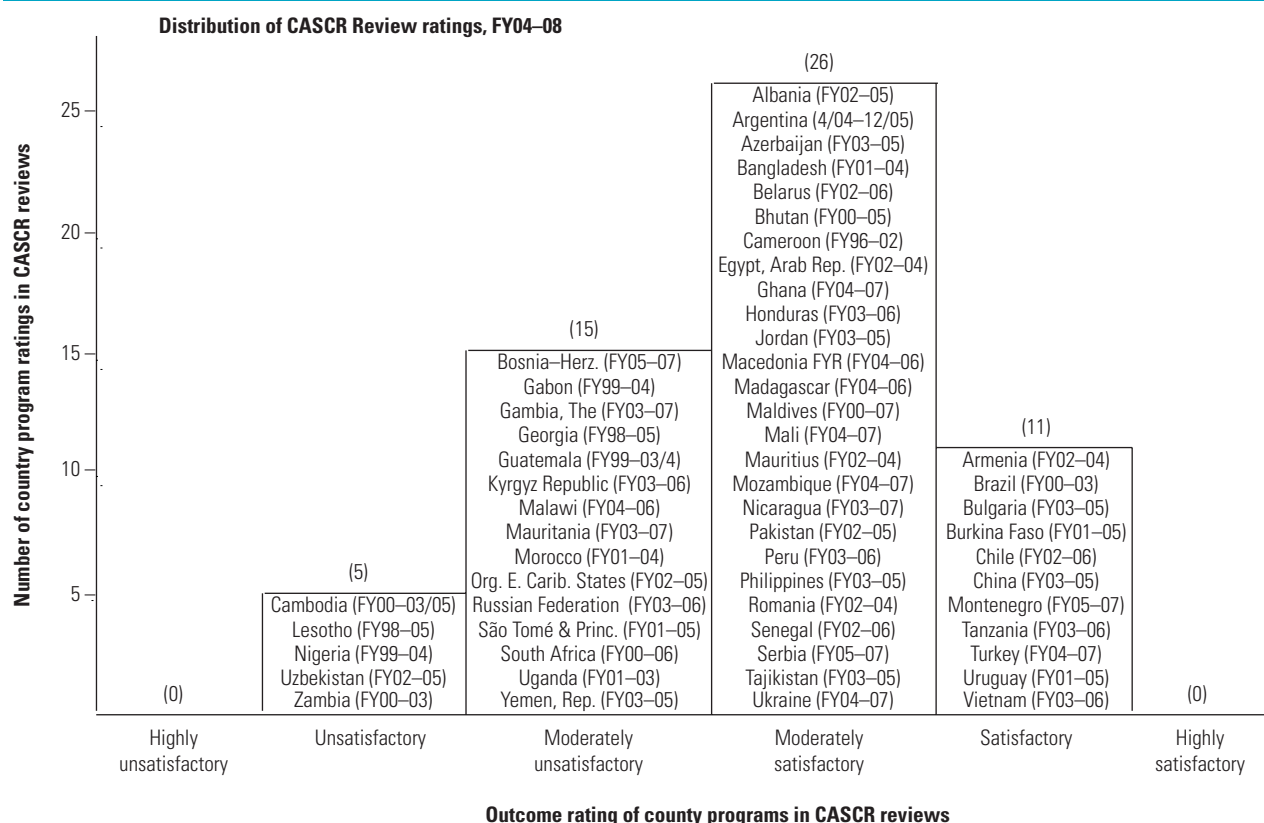
Few country programs have produced best-practice results. China and India, both above the line, the poverty-weighted satisfactory share is nearly two-thirds.

There are also significant differences among subgroups of countries, as shown in table 2.3. Bank programs in MICs far outperform those in LICs. While IEG rated 39 percent of programs in MICs

satisfactory or better, this was true of only 18 percent in LICs. Using a lower threshold, the difference is smaller but still substantial: 64 percent of ratings in MIC programs were moderately satisfactory or better, compared with 51 percent of ratings for LIC programs.

A less in-depth but more up-to-date snapshot of country program outcomes is provided by IEG reviews of Country Assistance Strategy Completion Reports (CASCRs). CASCR reviews differ from Country Assistance Evaluations (CAEs) in that they are a review of the Bank's own self-evaluation of its country assistance program and cover a shorter time period, typically three to five years, as compared with a decade or more in CAEs. CASCR review data, available since fiscal 2004, present a slightly more favorable assessment. Two-thirds of all programs reviewed were rated moderately satisfactory or better, as shown in figure 2.5 and table 2.4.¹⁰ Here too, none of the country programs were rated highly unsatisfactory.

Consistent with the CAEs, higher ratings are almost entirely due to better outcomes in MICs than in LICs. More than three-fourths of the MICs had moderately satisfactory or better country outcomes, compared with half of LICs reviewed. And in only one MIC was Bank assistance rated unsatisfactory, as compared with four LICs, even though there were many more MICs in the sample.

Figure 2.5: CASCR Reviews Indicate That Bank Programs in MICs Outperform Those in LICs

Source: IEG reviews of Country Assistance Strategy Completion reports.

Note: CASCR Reviews cover a shorter period than CAEs. Years in parentheses indicate the period that Bank country programs are assessed.

Table 2.4: Summary of CASCR Review Ratings, FY03–08

	MICs	LICs	All countries
Highly satisfactory	0%	0%	0%
Satisfactory	22%	13%	18%
Moderately satisfactory	54%	39%	48%
Moderately unsatisfactory	22%	30%	26%
Unsatisfactory	3%	17%	8%
Highly unsatisfactory	0%	0%	0%
<i>Satisfactory and highly satisfactory</i>	<i>22%</i>	<i>13%</i>	<i>18%</i>
<i>Moderately satisfactory or better</i>	<i>76%</i>	<i>52%</i>	<i>66%</i>

Source: IEG reviews of Country Assistance Strategy Completion Reports.