Overview

The Independent Evaluation Group's (IEG) Results and Performance of the World Bank Group (RAP) is a comprehensive assessment of World Bank Group performance, drawing on recent IEG evaluations. The report also examines how effectively the World Bank Group addressed current and emerging development challenges. This year's RAP focuses on gender integration in World Bank Group operations, building on previous examinations of World Bank Group approaches to risk management (RAP 2013) and the Millennium Development Goals (RAP 2014). This report also reviews how effectively the World Bank Group's portfolio and country programs deliver results, and its system for monitoring the implementation of IEG's recommendations.

"No country, community, or economy can achieve its potential or meet the challenges of the 21st century without the full and equal participation of women and men, girls and boys." This statement from the World Bank Group website's topic page on gender states the essential importance of gender for development. In line with that view, the World Bank Group made considerable progress in addressing gender issues during the past 15 years. Gender has been a prominent corporate objective since the first World Bank Group strategy, introduced in 2001. This year, a new World Bank Group gender strategy has been launched—the first joint World Bank–International Finance Corporation (IFC) strategy to focus on gender. It is an important step toward sharpening the corporate focus on gender and improving the approach to gender mainstreaming.

This report describes how mechanisms for integrating gender in projects and country strategies are working, and to what extent they provide meaningful information about progress and results on gender. The analysis aims to inform World Bank Group efforts to strengthen the approach to documenting, assessing, and evaluating results as part of the new strategy rollout.

Corporate commitments have helped broaden policy and strategy attention to gender, which increased gender uptake. However, the quality of gender integration was uneven. Corporate commitments on gender were agreed, and reflected in, the results framework of the 16th Replenishment of the International Development Association (IDA16) and IDA17, and in the World Bank Group Corporate Scorecard. The commitments were monitored through the introduction of gender flags, mandatory disaggregation of project beneficiaries, and inclusion of gender indicators in IFC's Development Outcome Tracking System. The gender mainstreaming strategy was successful in increasing gender uptake (the number and percentage of operations and country strategies that addressed gender issues at entry). This uptake is more notable in recent years.

At the same time, progress in including gender integration at entry (the increase in the rate of projects defined as "gender-informed") was not matched by similar attention to quality and depth of gender integration. Current guidelines refer to integrating gender when relevant, but the guidelines do not define relevance, resulting in variable practice. The portfolio review revealed that projects—and especially country strategies—do not clearly identify gender relevance, and therefore struggle to clearly articulate an explicit result chain and select appropriate indicators for measuring results.

Country strategies are required to integrate gender by corporate commitments, and they generally do this. Most country strategies identify gender as a cross-cutting theme, delegating the explanation of the rationale and results chain to individual projects within the country portfolio. This effectively dilutes the strategy focus on the country gender priorities that need to be addressed.

The monitoring and evaluation (M&E) frameworks of operations and country strategies do not adequately measure and report on gender results. The importance of tracking, reporting, and assessing gender results has become a higher organizational priority in the recent period, reflected in the corporate commitment to tracking female beneficiaries. An increasing number of projects report this indicator; several added it during implementation. When indicators were integrated at an early stage and were grounded in concrete actions and components, reporting on the indicators was substantial. This was true for both projects and country strategies—reporting on the indicators was typically sound when gender actions were identified and were supported by a relevant background diagnostic, and indicators were integrated into Country Assistance Strategy programs. Qualitative reporting of indicators is more frequent than quantitative reporting, but both have serious limitations, including poorly reported indicators, vague qualitative statements, incoherent reporting, and not reporting indicators at all.

However, weaknesses in M&E frameworks prevent meaningful tracking and assessment of projects' and country strategies' gender results. Nearly all country strategies reviewed incorporated gender in at least some dimension, but only a few had internally consistent background analysis, actions, and indicators and corresponding results measured and reported. At the project level, development outcome indicators reflecting gender are rare, and many person-level indicators are still not sex-disaggregated, even when meaningful disaggregation was technically feasible. Defining and counting female beneficiaries, though increasingly done, requires more than counting recipients or residents of the project area, and involves measuring both the direct and broader distributional impacts. For both country strategies and projects, indicators used are often inadequate to capture gender results since they are frequently narrow in scope and tend to measure outputs rather than outcomes.

The new World Bank Group gender strategy offers an opportunity to improve tracking and reporting of gender results. The introduction of the first World Bank Group gender strategy in fiscal year 2016 (FY16) is an opportunity to ensure that the mechanisms established to support gender integration in country strategies and projects are fine-tuned to generate and produce meaningful information and reporting. IEG's analysis shows this can be achieved by moving away from a purely mechanical observance of corporate mandates and a box-ticking culture to embrace a more consistent, robust approach that involves identifying priorities, articulating a results chain, selecting meaningful indicators, following up coherently, and monitoring and reporting the results achieved.

The performance of Bank lending, IFC Advisory Services, and Multilateral Investment Guarantee Agency (MIGA) guarantees remains stable, but the performance of IFC investments continues to decline. World Bank Group commitments peaked in FY10 after the global financial crisis, but lending tapered off through 2013. Commitments are now rising once again and have increased for two consecutive years. Commitments reached $60 billion in FY15. Weighted by commitment size, the performance of World Bank projects for the period FY12–14 exceeded FY17 corporate targets; measured by number (unweighted percentage), performance held steady, but was below the FY17 corporate target. The performance of Bank projects in East Asia and Pacific resisted the overall trend, declining in FY12–14 to a performance rating just above the Middle East and North Africa Region. Among the Global Practices, performance was particularly strong in Social Protection and Labor, and in Agriculture.

IFC advisory and MIGA guarantee products continued to perform well, but IFC investment lending continued the downward trend first reported in 2013. Falling equity success, influenced by ongoing fallout from the financial crisis and global economic slowdown, affected investment success. Investment project performance improved in IDA and blend countries, but continued to dip significantly otherwise, reflecting poor investment outcomes and work quality, particularly in the manufacturing and services industry group, and in the Europe and Central Asia and East Asia and Pacific Regions.

Mid-course corrections matter more than project size for successful performance in World Bank projects; for IFC projects, size matters for real sector projects, but less than do other risk factors. Using statistical analysis, IEG found that initial commitment size is not a key element of success for World Bank investment lending projects, but the change in project size during the project cycle is a significant correlate of a project outcome rating. Additional financing is typically introduced for what are deemed well-performing projects, which increases their size, but does not cause their success. Performance can improve by paying more attention to certain factors at entry, such as risk assessment, relative design complexity, and clear objectives. Currently, quality at entry is not systematically rated at the beginning of the project.

The analysis also found that higher outcome ratings were associated with projects in countries with higher Country Policy and Institutional Assessment ratings, when controlling for country- and project-specific characteristics, Global Practice, and Region. Lower project outcome ratings were associated with task manager turnover, higher supervision cost, and whether the project was ever labeled a problem project. Related analysis suggests that early and candid assessment of project implementation performance is important because lack of corrective actions or untimely restructuring were the key reasons for poor project outcomes. Projects in countries with greater gender equality, more effective government functions, or more stable rule of law were also associated with higher outcome ratings.

For IFC projects, IEG found that project size was a significant correlate of development results for real sector projects, but not for banking projects. However, for real sector projects, the association of commitment size with development success diminished as other risk factors were added to the model. For these projects, external project risks (such as management quality, market conditions, investment climate, and internal controllable risk factors in IFC's work quality) are more significantly correlated with development outcomes.

Using the identified factors associated with development outcomes, analysis to predict the performance of IFC projects revealed that recently committed IFC projects are likely to perform worse than recently evaluated projects, despite a larger concentration of recent IFC commitments in less risky countries.

Outcomes of country programs for International Bank for Reconstruction and Development (IBRD) and IDA have improved during the past three years and remained stable in Fragile and Conflict States with a higher success rate than for the World Bank Group average. Improvement is led mostly by the European and Central Asia and Latin America and Caribbean Regions. The performance of the World Bank Group in designing and implementing country programs deteriorated slightly overall, especially in Latin America and Caribbean. It improved in all other regions including Africa region or remained stable.

The Management Action Record (MAR) process was successful in creating a formal,

transparent, and well-understood structure within the organization for reporting about progress made to address recommendations in IEG evaluations. IEG evaluations make recommendations to improve the development effectiveness of the World Bank Group. IEG and management then monitor the implementation of actions associated with those recommendations to promote accountability and generate knowledge about where improvements are and are not made. The World Bank Group's Boards of Executive Directors can use the MAR as a tool to hold World Bank Group management accountable for actions to which it committed.

Between FY12 and FY14, IEG produced 25 corporate, sector, and thematic evaluations, resulting in 170 recommendations being tracked using the MAR. This year, as in previous years, IEG found that implementation of those recommendations improves over time. IEG rates implementation of just over 80 percent of the recommendations substantial or better by year four.

However, M&E recommendations have eluded meaningful response. Implementation progress is even across all major recommendations categories except for M&E quality. For M&E, World Bank Group management generally agreed with IEG's recommendations, but implementation was difficult because of issues with data collection, assessment methodologies, and the time required for outcomes to materialize. Management acknowledged these difficulties and rated implementation substantial for only half of M&E-related recommendations in the fourth year of implementation, which is well below average.

The MAR could be an even more effective tool if it were less formalistic and more purpose driven, and by integrating active, deliberative, and ongoing dialogue throughout the process.

The MAR is a useful accountability tool, but the process requires further reform to make it an effective tool for learning. Interviews with selected IEG and World Bank Group managers and staff involved with the six evaluations entering their fourth year of follow-up revealed that the evaluations themselves have more influence than the recommendations alone. Many managers and staff interviewed considered the MAR follow-up to be a static, bureaucratic accounting exercise that resulted in little deep reflection on progress. Recommendations are tracked even when they may have lost their relevance as the operational environment and strategic priorities evolved.

The review identified three major factors that contributed to an evaluation's influence. Timely evaluations that generated findings and recommendations aligned with ongoing strategic priorities and operational programs tended to have relatively strong adoption of recommendations. World Bank Group managers and staff also said they were more likely to take an evaluation and its recommendations seriously if they considered its analysis to be of high quality and the evaluation team technically credible. Management also cited the value of early and frequent engagement with evaluation teams as a factor in their receptiveness to findings and recommendations.

Still, it was also noted that for some evaluations that address difficult or cross-cutting issues with unclear ownership, early engagement with the right stakeholders may not be achieved. Such evaluations may be among the most influential in the longer run, but the influence may take a longer time, and avenues of influence other than what the MAR can offer may be required. A statement in the Independent Panel's report to the Committee on Development Effectiveness stressed the importance of IEG's strategic engagement and a close but uncompromised relationship with management and staff.

Further reforms of the MAR process should seek to encourage earlier and deeper engagement between evaluators, management, and topical stakeholders.