Comments on
"Policy-Based Financing at the World Bank: Evolution, Performance, and Reform"

Comment by: Cheryl W. Gray

This chapter provides a useful, concise, and well-written summary of the evolution of the World Bank's approach to policy-based financing and methods to evaluate it. It shows the careful thinking undertaken by the World Bank as it has struggled to deliver effective support to countries, often in complex and difficult settings. As the chapter illustrates, policy-based financing has long been a major instrument of international development support, valued in the hundreds of billions of dollars annually across development agencies. Yet its breadth and complexity have made it exceedingly difficult to study, and evidence of its results remains elusive. This chapter helps to shed light on what is known, although for perhaps unavoidable reasons the picture is still incomplete.

1. World Bank Evaluation Architecture

The first point worth stressing is the practicality and value-added of the World Bank's evaluation architecture. Over more than three decades the World Bank has designed, implemented, and continuously improved a cohesive structure to document results from all its operations—both investment and policy-based operations—in a practical and cost-effective manner. The process begins with a self-evaluation by the task team, whose members know the operation best. That self-evaluation —which uses standard criteria applicable to all similar operations—is then reviewed and validated by the Independent Evaluation Group (IEG). The fact that 100% of self-evaluations are validated creates an incentive for task teams to report accurately while also producing a complete database of operation-level reviews across the institution. As noted in the paper, IEG follows up in some cases with more in-depth evaluations and/or broader sector or thematic evaluations, adding further context and evidence on results. The entire evaluation architecture is oriented toward documenting activities and outcomes, and it creates opportunities for learning through self-evaluation and analysis.

The evaluation systems of other multilateral development banks (MDBs) are similar, in part due to concerted efforts at harmonization across the MDBs. In my personal experience on both sides of this self-evaluation and validation system—as both an operational and an evaluation manager—the system is useful, practical and cost-effective.

While recognizing the positive aspects of the evaluation architecture noted above, it is also important to emphasize what the system does not do, especially in the complex area of policy-based financing. These types of routine World Bank evaluations are not in-depth impact evaluations with rigorous causal inference. They do not compare performance against counterfactuals to identify and measure cause and effect. Occasionally it is possible to apply impact evaluation techniques to specific interventions in specific settings, but this is not feasible across the board given the breadth and complexity of most World Bank operations, particularly policy-based financing. Rather, in the World Bank system, task teams and evaluators seek to define relevant objectives for World Bank operations and then determine whether those objectives were achieved during the life of the operation, with some implicit assumptions but no rigorous analysis of causation.

2. Evaluation of Policy-Based Financing

The chapter describes in detail the criteria for self-evaluation and validation of development policy financing (DPF), which is the World Bank's name for policy-based financing. These criteria have changed over time to reflect changes in the design of policy-based financing over time. When policy-based lending began in the World Bank in the 1990s, loans were disbursed in a series of tranches that were triggered by successful completion of policy reforms and institutional milestones. In contrast, the World Bank's current DPF approach provides all the financing up-front, upon completion of a small number of key prior actions. This differs from the policy lending of the European Union, for example, which has a performance element and disburses in part on the achievement of results.

The World Bank's approach thus puts a very heavy weight on a small number of upfront policy and institutional changes that it considers key to the country's success. While having a small number of prior actions simplifies the lending process and focuses the World Bank's oversight, it runs the risk that the assumptions regarding the impact of reforms may be wrong. Indeed, over time the consensus on what constitutes good policies has to some extent shifted, and it is likely that the prior actions and results indicators of many policy loans of the 1990s would now be seen as problematic by World Bank staff and evaluators. The World Bank's heavy reliance on selected prior actions calls for an equally determined analysis of the effect of these policy and institutional changes whenever possible. Yet in practice there are few opportunities for the World Bank to follow up with careful analysis of these effects after the completion of prior actions and DPF disbursement.

Recently the World Bank has moved from rating the relevance of the DPF operation's objectives (the standard approach in evaluations of investment operations) to rating the relevance of the prior actions, which are the only conditions for the operations that are directly within the World Bank's control. The World Bank is also putting greater weight on evaluating the relevance and quality of the operation's results indicators, World Bank performance, and the treatment of risk. These judgments are largely qualitative, and one person's judgment may differ from another's. To ensure these ratings are robust, it would be helpful to track whether guidelines, dialogue, and practice are converging in reasonably common standards across operations and over time.

An important aspect of DPF operations missed by the World Bank's evaluation approach is the impact of the resource transfers themselves, i.e., the impact of spending the hundreds of millions of dollars transferred to recipient countries through development policy financing. Indeed, some have argued that the increased availability of funds for governments to spend may be the biggest impact of policy-based lending in practice, greater than the support to policy and institutional reforms provided in the loans.

Measuring the impact of the resource transfers would require knowledge of how those funds are actually spent, and this is not straightforward. DPF transfers are wholly fungible and are likely in practice to fund the "marginal expenditure" in the public budget—i.e., expenditure that could not otherwise be funded. This marginal expenditure could be in any sector—including infrastructure, social programs, defense, agriculture, and enterprise subsidies. It is by no means a foregone conclusion—nor is it even likely—that such marginal spending will be in the sectors with the policy reforms supported by the DPF. Although determining the marginal expenditure is likely to be difficult in practice, it might be possible to get a broad sense of the overall patterns of public spending with and without the extra resources transferred through the DPF. Trying to assess such changes in public spending in at least a few large DPF operations would be a worthwhile evaluative exercise for the World Bank.

Finally, the focus of most evaluations on individual DPF operations fails to capture the overall distribution of World Bank support and resource transfers among borrowing countries, although larger thematic evaluations may help to capture this dimension. Given the political incentives facing both borrower and donor governments, as well as bureaucratic incentives within the World Bank and other development institutions, it is not surprising that many resource flows go to middle-income countries—where it is easier to lend and where absorption power and interest rates are typically higher—than to low-income countries where the need may be greater and access to alternative financing sources more limited. This is particularly true in multilateral development banks, whose income and credit ratings are dependent in part on loan proceeds, in contrast to bilateral or multilateral donors (such as the EU) whose resources come wholly from governments.

3. Findings on Development Policy Financing Effectiveness

The chapter reviews the data on the results of DPF operations over time and highlights several academic studies and thematic evaluations that have tried to draw further conclusions from these data. In addition to the inherent limitations on results measurement noted above, a few points stand out. First, there is a high prevalence of "moderately satisfactory" ratings for outcomes and World Bank performance. The difference between "moderately satisfactory" and "satisfactory" development policy financing—like the difference between "moderately satisfactory" and "moderately unsatisfactory" development policy financing—is based on the qualitative judgments of the validators, and this runs the risks of inconsistency noted earlier.

Second, thematic evaluations emphasize the prevalence and salience of DPF prior actions related to public financial management (PFM). Managing public finances is indeed an important and powerful responsibility of government that strongly influences the distribution of resources and effectiveness of public programs. It is an area that the World Bank has been able to focus on and influence relatively effectively through its operations. PFM has technocratic aspects—e.g., budgeting processes, computer systems, and auditing—to which the World Bank can bring needed expertise and resources. Other areas of governance reform, such as election systems, public employment, or direct anticorruption efforts, may be as (or even more) important for development outcomes but have been more difficult for the World Bank to address in the political environments in which it works. These sensitivities put limitations on what kinds of prior actions are feasible in DPF operations, which might also limit their potential development impact.

Third, the chapter also notes the value of the granularity gained through more in-depth evaluations of particular operations through project performance assessment reports (PPARs). Yet it is unclear to what extent operational World Bank staff read PPARs and integrate the findings into their operational work. Continued efforts to increase PPAR accessibility and impact, through both greater outreach and continued experimentation with content and process, would help to increase their usefulness. For example, in some instances comparative analyses of several similar operations may offer insights that reviews of individual projects do not. Doing some evaluations jointly with outside experts, other development organizations, or operational staff might also increase their ownership and visibility.

Finally, one of the cited academic studies concludes that the level of macroeconomic stability is positively associated with the success of DPF operations.[1] As noted in the chapter, it is not possible to untangle causation, i.e., whether the World Bank's operation influenced the country's policies or good policies made it possible for the World Bank to lend. The fact that government ownership is also key to achieving outcomes and that "the World Bank's policy lending is significantly and positively correlated with the quality of social policies and institutions"[2] both reinforce the overwhelming importance of committed country counterparts.

In sum, the evidence strongly supports the finding that enlightened leadership, pro-development policies, and effective World Bank support go hand in hand, which is probably as much as can confidently be claimed. Attributing positive causal impact to the DPF operations themselves is not likely to be supported by the evaluation techniques available. But it is an important finding in today's

[1] Lodewijk Smets and Stephen Knack. 2014. World Bank Lending and the Quality of Economic Policy. *World Bank Policy Research Paper.* No. 6924.

[2] Željko Bogetić and Lodewijk Smets. 2017. Association of World Bank Policy Lending with Social Development Policies and Institutions. *World Bank Policy Research Paper.* No. 8263.

world that the World Bank can contribute to development by recognizing and supporting committed and effective leaders without having to prove that its actions led to that commitment.