A Rule for Private Sector Subsidies

Additionality and Competitive Allocation at the World Bank Group

ÁLVARO S. GONZÁLEZ

Abstract

The World Bank has yet to pin down a working definition for what it labels private sector development (PSD). This lack of clarity has permitted the World Bank to change what it treats as PSD and what PSD operations are expected to achieve. This paper defines a private sector development operation as one where public capital enters a market served by private suppliers to correct a market failure. For operations that fit that definition, it asks whether World Bank involvement is justified and whether it is additional to what would otherwise occur in the market. Applying this test, the paper identifies five approaches to private sector work between FY1975 and FY2025 and finds little evidence that changes in focus and strategy responded to worsening performance. It also finds that of the 156 operations classified by World Bank staff as PSD in FY24–25, only 29 meet the paper's definition, and 21 fail the additionality test, accounting for US$3.20 billion in total commitments. The paper recommends making the case for intervention an approval requirement under a World Bank Operational Policy, thereby establishing a common test that does not depend on which flavor of PSD is in fashion at the time.

POLICY PAPER 403 • SEPTEMBER 2026

A Rule for Private Sector Subsidies: Additionality and Competitive Allocation at the World Bank Group

Álvaro S. González
Center for Global Development

The author thanks Alan Gelb, Markus Goldstein, Charles Kenny, Ganesh Rasagam, Gaiv Tata and Chiaki Yamamoto for comments on earlier drafts. The author used Claude, developed by Anthropic, for copy editing, preparing figures, analyzing data under the author's direction, and citation checking. Three language models, ChatGPT, Claude, and Gemini, were used to code the 30-operation reliability subsample under the same written protocol. Section III reports the results and explains what each reader was given. All arguments, judgments, and remaining errors are the author's own.

This paper is a companion to Álvaro S. González. 2026. "The World Bank Cannot Stop Reorganizing Its Private Sector Work: The Missing Welfare Anchor." CGD Policy Paper 402. Washington, DC: Center for Global Development," which sets out the three-hurdle test this paper applies. The test is restated here in compressed form.

Álvaro S. González. 2026. "A Rule for Private Sector Subsidies: Additionality and Competitive Allocation at the World Bank Group." CGD Policy Paper 403. Washington, DC: Center for Global Development. https://www.cgdev.org/publication/rule-private-sector-subsidies-additionality-and-competitive-allocation-world-bank-group

CENTER FOR GLOBAL DEVELOPMENT
2055 L Street, NW Fifth Floor
Washington, DC 20036

1 Abbey Gardens
Great College Street
London
SW1P 3SE

www.cgdev.org

Contents

Executive summary **1**

Introduction **3**

I. What counts as private sector development **6**

The admissions test 10

What falls outside the definition—the exclusions 11

The common property: what operations share 12

II. The performance question **13**

III. Where operations fail the test **16**

What the appraisal documents record 17

IV. Three institutions, one obligation **26**

Institutional discipline: Gate and mechanism 28

Gate and enforcement 30

The fragile-state objection 30

V. Conclusion **32**

Figures

1. The private sector share of lending and the measured performance of private sector operations, FY1975–2025 15

Tables

1. The private sector development portfolio in the Bank's taxonomy, FY1975–2025 12
2. Portfolio composition and measured performance by vocabulary wave 15
3. Populations, frames, and samples used in this paper 16
4. Four readings of the 30-operation reliability subsample 23
5. The FY24–25 private sector development portfolio against the test 25

Executive summary

Some of the World Bank's lending is placed under the theme of private sector development (PSD). This paper asks what PSD is by asking what PSD work does. What PSD does is not a coherent set of activities. The condition for coherence is that an operation counted as private sector development is a Bank intervention that places public capital in a market where private suppliers already operate. If an operation does not do that, it is not a PSD operation.

Only two situations justify that kind of intervention—society gains more from an undersupplied activity than private suppliers gain, or the market will not provide financing or take the risk and lend, so the Bank must step in with its capital.

Defining PSD operations this way presents a problem for the Bank. A PSD operation hopes to get firms to hire, invest, grow more and become more productive as well. The firm knows how much it would take for it to do what the PSD operation wants it to do. The Bank does not. Unless some arrangement obliges firms to reveal how much they would take to induce them to do these things, the Bank should expect to pay more than the minimum in every transaction. This is inefficient.

To confirm that definition fits the PSD label, a question is posed of any operation as it is being designed—is public money going into a market where private suppliers are already operating? A yes also creates an obligation to give an account of the value that public money will add once it enters a market with private suppliers already present.

When bringing that definition to bear against the Bank's PSD records, they return three findings. First, it is hard to know what is in the PSD box. Over a quarter of labeled projects—937 of 3,489—have that label without revealing which *kind* of PSD operation they are. In addition, work that obviously belongs to trade facilitation, like helping exporters or speeding up customs, is filed outside the private sector theme. Four careful readers (one human and three large language models), working from the same protocol detailing what a PSD operation should look like and reading the same 30 project documents, were asked to tell which projects count as PSD, and only agreed on nine. The other 21 depended on who (or what) was doing the reading. Admittedly, more rules could be written until the three models and the human agree—but then the agreement comes from the rules, not from anything in the project documents.

The second finding concerns the performance record. Here, the scoreboard was found to be broken. Over five decades the Bank adopted and dropped five different theories (orthodoxies)—directed credit, adjustment and privatization, the investment climate, jobs and inclusion, and the Cascade—on how it should intervene to help the private sector. The size of the PSD portfolio expanded and diminished wildly throughout that time—from 18 percent of lending up to 58 and back down to 17. The performance record, how well those projects were graded based on what they promised to get done, was about a 4 out of a possible 6, throughout. That is the same performance rating as everything else the Bank does. The grades bunched toward the middle—seven in ten recent projects get one of the two middle marks. A performance rating that reads the same no matter the project, across time says more about the rating system and less about the projects.

Each of those five orthodoxies was adopted, then left behind, with no signal from the scoreboard as to the results of the shifts that prompted either move—orthodoxies changed for other reasons. The project appraisal documents provide another result: take 300 finished projects, split them into those whose paperwork said what the public money was buying and those whose paperwork did not. Same grades. The scoreboard does not even reward showing one's work.

The third finding comes from reading the present portfolio at approval, when it is more informative, rather than after closure. Of 156 recently labeled PSD projects, 111 have an appraisal-stage document that the protocol can use to answer the questions applied to it. Of those, 29 put Bank money into markets where private sellers already operate. Twenty-one[1] of the 29 could not provide an answer to the definitional question. Most (17) either name no real market failure or name a market failure that country evidence could not support. Eight channel finance through an intermediary bank, making the intermediary responsible for the pricing decision of that financing.

The Bank's private sector work is adjacent to just about everything the institution does, so the natural suggestion is to track it everywhere—make it a cross-cutting theme, like gender. In the Bank, cross-cutting themes take on tags in the reporting system. And staff deliver what is counted by the corporate scorecard. In FY24–25, 43 of the 111 tagged operations for which an appraisal document is available are ordinary public work. The tag does not find private sector development work; the counting creates tagged work.[2]

If the aim is to discipline the Bank's PSD work, the point is to avoid the theme tag and create a gate, where the operation cannot proceed without clearance. A theme, on the other hand, records the claims made about the operation's intended outcome and has no authority to prevent an operation from proceeding. This paper uses the tag itself, but only to locate operations for review; the classification it applies to them is its own.

A gate is necessary, but not sufficient, to address the problems revealed by the PSD portfolio data. When project appraisals do not price the subsidy, identify a market failure that can be supported by country evidence, or correctly classify the operation, a reviewer can detect these omissions and prevent such operations from proceeding. But once teams supply this information and a diagnosis,

1 Why report "21" instead of a percentage? Because the percentage depends on what it is divided by, and the divisor moves based on where the field's boundary is drawn—21 is 19 percent of the 111, or 72 percent of the 29, and among the 10 operations that put most of their money into such markets, half fail. The count is based on what the Bank knew when it said yes to operation approval. Different readers get somewhat different counts—but every reader gets a failing majority.

2 A gate has a reviewer who is not on the team proposing the operation, and their standing does not depend on whether the operation gets approved. The test used to determine whether an operation should be held up is written somewhere that requires a World Bank Group's Executive Board vote to change, so the people under review cannot amend it.

the reviewer cannot determine whether what the project team—whose incentive is approval—provided is correct. A competitive allocation, however, can help price subsidies by eliciting the support firms will accept, and a bid at or near zero establishes that no subsidy was necessary. This is information the appraisal cannot reliably provide.

The border between the Bank, the International Finance Corporation (IFC), and the Multilateral Investment Guarantee Agency (MIGA) mandates which define which institution is better able to intervene in private markets when needed. The mandates also address concerns that all three institutions could inappropriately intervene where private sector actors may be able to act. However, the boundaries between the three may have been drawn in the wrong place or not clearly enough. A World Bank sovereign policy loan, an IFC investment, or a MIGA guarantee can all address the same binding constraint on investment, but choosing the appropriate instrument should be based on where the constraint is found and which channel would best change behaviors.

Introduction

Two similar operations passed the same World Bank bar to get approval, but they presented different levels of evidence to do that. Two private sector development operations for Türkiye, approved three months apart in FY2025 by the World Bank, and both reach firms through a domestic development bank. But that is where the similarities end. The first operation, a guarantee, accounts for what commercial financing would not have supplied with a market sounding and a published fee schedule. The second operation, a loan, provides no comparable evidence to support its intervention.

The €360 million credit guarantee to the Industrial Development Bank of Türkiye (TSKB) produced a published fee schedule, making the transaction's terms observable to anyone. To establish that the guarantee has additionality, the operation team conducted a market sounding. Twelve international commercial banks were contacted. Four responded with indicative participation levels and financing terms. Their responses confirmed the importance of the guarantee for TSKB's financing and helped set the operation at €600 million, with a partial credit guarantee of up to 60 percent. Each responding bank made its uptake conditional on supplemental credit risk mitigation for the portion of the exposure not covered by the guarantee. They also reported that such cover was only available to a limited extent from the private market. The operation team's review of TSKB's financial reports found no commercial hard currency financing with a tenor above five years in the preceding five years. Private commercial lenders had not supplied TSKB with ten-year financing during those five years. The absence of longer tenor is addressed with a Bank guarantee that secures repayment of a loan with a final maturity of ten years.

The second operation is an approximately US$500 million loan to the Development and Investment Bank of Türkiye (TKYB), intended to provide capital for lending to firms in earthquake-affected provinces. TKYB receives and uses the Bank's capital to extend sub-loans to target firms. The terms of those sub-loans are not disclosed in the appraisal. The interest rate charged by TKYB is also not disclosed in the appraisal and it is not observed in the market. It is assumed that firms will pay the officially reported average rate on commercial loans in Türkiye. It is not based on TKYB's sub-loan documentation or evidence of the rates TKYB charged or was willing to charge.

There are also two claims included in the appraisal that serve to justify the intervention. The first claim is that lending by private and state deposit-taking banks has not expanded into the region despite existing efforts to encourage expansion. The second claim is that the operation targets borrower types or loan categories that private banks have overlooked or declined to serve.
The appraisal also claims that maturity would not be extended funds are routed through other financial institutions, including non-bank intermediaries.

Financing is often needed after a major earthquake or other disaster because private credit frequently contracts following a large shock. The appraisal does not make that inference. Instead, it points to the absence of expanded lending by private and state deposit-taking banks in the earthquake-affected provinces, despite efforts by the Credit Guarantee Fund, a Treasury-backed credit guarantee scheme, to change bank behavior. The appraisal uses this absence of expanded lending as part of its case for Bank intervention.

Left unverified is the counterfactual; what commercial financing would have supplied had the Bank not provided its capital for the TKYB loan. The counterfactual is important to establish because when it is established, it avoids imposing costs of the lending that does not occur. Since the Bank lends its own capital, the borrower does not bear this cost. The borrower receives the loan, possibly on better terms. The cost also does not fall on the World Bank because its accounts record a normal loan. Instead, the cost instead falls on a third party outside the transaction, the firm that no private lender will serve and that otherwise might have received the Bank's capital. Public lending capacity exists to serve that unfinanced firm. If the operation did not provide additional value, value that markets would not have already provided, the same Bank capital could have been allocated to an operation that did. Committing World Bank capital to one transaction precludes its use for other lending with potentially greater value.

Corroboration is what makes the counterfactual checkable. Without it, the appraisal cannot establish whether the Bank's capital was needed in this transaction. Determining the standard for what is good or bad counterfactual is not about what corroboration entails. That standard is discussed in Section III.

Committing World Bank capital to an operation that does not provide additional value, while leaving the unfinanced firm untouched, is an opportunity cost not registered by the Bank. The World Bank's systems track completed transactions but do not capture the opportunity costs borne by firms excluded from participation in the operation. Those ignored firms, by definition, never complete a transaction. Greater diligence by operations staff will not register this opportunity cost.

The defect lies in the instrument and process used to prepare and approve private sector operations. The process of approving the second Türkiye PSD operation did not register that a counterfactual was left unverified.

The criterion for determining whether an operation qualifies as private sector development is the counterparty to the operation. The criterion has two elements: the identity of the party with which the Bank transacts, and the character of the market affected by the transaction. Operations outside private sector development involve a public counterparty—that is, the public sector. A private sector development operation, by contrast, uses public money in a market where private suppliers are already active. Public money might end up financing activity that private money would have financed. For this reason, the Bank must justify the use of public money in a market where suppliers are present and has an obligation to make sure that it does not displace them.

The Bank is obliged to explain what the operation will achieve with its public financing that would not have been achieved if the market is left to its own devices. The Bank will be responsible for addressing the obligation. The obligation also requires a statement, supported by evidence, that identifies what the market is not supplying and that public money will supply in its place. The statement must be provided before final approval of the operation—so there is a requirement on the timing to fulfill the Bank's obligation.

While it is important that an additionality obligation cover as many PSD operations as possible, four categories of financial sector work fall outside the scope of this paper: supervision of financial institutions, capital markets, payment and settlement, and financial stability. Each already has its own additionality practice. There is no need to include them here.

However, financial sector that lends to firms through intermediaries is within the scope of this paper. In fact, most of the operations examined in this paper take this form. These are included based on the presence of an incumbent supplier in the market where the operation intervenes. Their inclusion is not based on the category assigned to the operation. Operations classified under the public private partnerships sub-theme have this property and fall within the definition. Section III explains the application of the definition to those operations.

A harmonized framework on additionality for private sector operations was agreed in 2018. MIGA's framework applies only to non-sovereign operations. IFC also applies its own explicit and published additionality doctrine. The Bank was consulted on the harmonized framework and approved it, but Bank operations are sovereign operations and therefore fall outside its terms. But this is not exactly confirmed by the data. Within the World Bank Group, the Bank is the institution that holds the largest balance sheet that reaches firms. Yet the institution with the largest exposure to firms is the only one whose operations are excluded from the additionality standard it approved.

At a base level, there is nothing new proposed in this paper. The questions this paper applies to private sector operations are already present in the Bank's own policy framework. The Directive on

Investment Project Financing requires an economic analysis for every operation. That analysis must address (a) what the operation's expected contribution to development may be; (b) what the rationale for the public sector's intervention is, and (c) what could be the value added of the Bank operation.

That same Directive, however, does not require evidence in support of the answer to the second question. In the same way, it requires no evidence in support of the third, either. The two requirements exist without an evidentiary standard and without a consequence for the approval of the operation.

In comparison, the World Bank Group's peers had a framework to address these questions that predates the harmonized framework of 2018. The agreement establishing the European Bank for Reconstruction and Development dates from 1990. It prohibits financing where the applicant is able to obtain sufficient financing elsewhere on terms the bank considers reasonable. Financing available elsewhere on reasonable terms is the counterfactual later named in the 2018 harmonized framework. A comparable institution had therefore applied the same counterfactual for close to three decades before the framework adopted it.

This re-think of private sector work is needed now more than ever before because developing economies face jobs, productivity, and growth challenges which are expected to be addressed to a greater extent by the private sector. The contraction now under way in aid budgets has no precedent in the post-war period: a US$8.9 billion, 23 percent reduction in US development-related allocations (Mitchell and Hughes 2026), a planned reduction of UK official development assistance (ODA) to 0.3 percent of gross national income (GNI), and a decline in aggregate ODA, reported by the Organisation for Economic Co-operation and Development (OECD), from US$215.1 billion in 2024 to US$165.5 billion in 2025, at 2024 prices and exchange rates (OECD Development Assistance Committee April 2026).

Over the next decade, 1.2 billion young people in developing economies will reach working age. Current projections put the number of jobs those economies will produce over the same period to be about 420 million. The private sector is expected to be the largest employer, by far, of all those jobs but it is projected to fall short. The World Bank Group has been given the mandate to support private sector development to help firms and markets close that jobs gap. At the moment, the Bank Group and its private sector development practices are not yet ready to use the public resources they hold to take on a task of this size.

I. What counts as private sector development

The World Bank uses theme code 120000 to group its private sector development work. The code includes work in Enterprise and Sector Development, Public Private Partnerships, Innovation, Technology and Green Competitiveness, Business Enabling Environment, Maximizing Finance for Development, Regional Integration, and Market Competition. These tags are not mutually exclusive, and a single operation sometimes receives more than one. Bank staff assign the theme code. Because there is some leeway and discretion in that assignment, a theme code does not indicate whether an operation is, in fact, private sector development work.

This paper uses the theme code only to search and identify operations classified as private sector development. The paper defines private sector development work independently of the tag so its classification need not coincide with how staff classified it.

Section III uses theme code 120000 to identify a cross-section of operations from fiscal years 2024 and 2025. Section II uses the Bank's earlier Major theme 4, which the Bank calls Financial and Private Sector Development. Major theme 4 places financial sector lending and private sector development lending in a single category where they cannot be distinguished. Neither the old nor new classification system identifies the full population of private sector development operations. Up to a point in the approval process, staff have nearly full discretion to assign and reclassify operations under different theme codes. The analytical consequences of allowing such discretion in defining the population of operations under private sector development are discussed in Section V.

In 2016, the Independent Evaluation Group (IEG) reviewed the World Bank's private sector work and compiled a list of activities that it classified as private sector development. The list included innovation and entrepreneurship, fragile-state engagement, business-regulation reform, transformational engagements, industrial competitiveness, targeted support to small and medium enterprises (SMEs), financial inclusion, housing finance, capital-market development, and public-private partnerships. The IEG most likely recognized that these activities do not have that much in common but anyway cast a wide net and presented the list as its best attempt to describe the Bank's private sector work. Trying to devise a clear definition of private sector development by reviewing this list is an insurmountable challenge and probably not called for in this IEG evaluation. But it suggests that the ambiguous boundaries of what is and is not a private sector development operation cannot be specified using IEG's own classification of PSD.

To identify if the boundaries across PSD operations are as porous as the IEG list seems to suggest, this paper undertook a classification exercise where it tested whether four independent readers, using a shared protocol, would agree on which operations were and were not PSD operations. The four readers reviewed the same thirty appraisal documents—the author and three large language models Gemini, Claude, and ChatGPT. All four agreed on nine operations, admitting three and excluding six as PSD operations. They were divided on the remaining twenty-one. Table 4 identifies each reading, and Section III reports the agreement statistics.

A controlled test is used to examine the effect of this definitional ambiguity on how the Bank classifies its own operations. Four readers reviewed the same thirty appraisal documents using the same written protocol to classify the operations as either PSD operations or not. That protocol provided guidance on the dimensions and characteristics that a private sector development operation must have. The thirty appraisal documents were taken from a pool tagged with the private sector development theme code.

The author was one of the readers and Gemini, Claude, and ChatGPT each conducted one of the other three readings. When all four readers finished, there was agreement on nine operations; three operations were unanimously classified as PSD operations and six were unanimously excluded. That left twenty-one operations where they disagreed. Table 4 shows the details on the four readings, and Section III reports the agreement statistics in detail.

The root cause of the disagreements was next. All four readers received the same instructions, so differences between their classifications do not reflect differences in the information or guidance they received. Differences arose on issues where the protocol was silent. For example, the protocol did not specify how budget support operations should be treated, and it did not define what "substantially" private sector development meant. Each reader had to resolve those questions independently. After the readings were completed, an addendum, providing more detailed instructions to resolve those issues, did close both gaps in the protocol. However, the addendum did not alter the rules under which the original readings had been conducted.

What sets all PSD operations apart is that they enter markets where private firms are already present. The public sector is the only counterparty to lending done in other parts of the Bank. Loans to governments are given on concessional terms (interest rates at or near zero, long maturities and grace periods). Those same terms would not be available from a commercial lender, so it is unlikely that the Bank's loan to governments displaces private lending.

The companion paper concludes that the Bank is necessary and adds value when the operation passes a three-hurdle test. The first hurdle is surmounted when an operation identifies a binding market failure (justification), passing the second hurdle requires a channel that removes that failure and not merely compensates for it (effectiveness), and the third hurdle is overcome if the operation demonstrates that private capital would not have financed the activity anyway (financial additionality). Failure at any one of the three hurdles is sufficient for the operation to fail the test.

The common property and the three-hurdle test provide a way to determine what properties an operation must have to count under PSD. Each question corresponds to one of three properties, taken up in the same order below. These questions are:

- Is somebody already supplying the relevant good, service, finance, or risk cover in the market concerned?
- What will public money do that the private sector is unwilling or unable to do?
- Who gets to keep the profits of the operation?

When the Bank lends in a market with other suppliers, it is impossible to know whether the firm that takes Bank money would have accepted the private lender's terms if the Bank had not entered the market. The operation will record a financed firm, a new investment, or more jobs. But with only that record, it remains unclear whether the Bank was needed or not. The issue of additionality must be resolved before approval because once the operation is done, it is difficult to show whether public money replaced private finance or not.

If the Bank operation competes directly with a private supplier, it is likely to displace it. A Bank operation that offers what no other business offers competes with no one, and there is no displacement.

The threat of displacement is based on the product, not the degree of concessionality. Where commercial lenders offer nothing beyond five-year tenor, they will not be displaced by a Bank operation that offers 15-year tenor. These are different products; the same five-year loan priced 200 basis points below the market rate is the same product at a subsidized price.

The second property asks under what conditions can public money enter a market already served by private suppliers. A Bank operation is allowed to trespass these boundaries when there is undersupply of some good. Under this condition, what a firm is doing is worth more to society than to the firms doing it. Jobs to disadvantaged young people, especially adult ex-combatants in conflict-affected areas, may be a good example of this where the benefit a firm gets from hiring an additional worker is much less than what society gains from an ex-combatant no longer being idle and susceptible to taking up arms once again. The difference between society's benefit from that hire and the firm's private returns from the additional worker's value is called the externality wedge. Narrowing that wedge is what public money pays for and that is why a Bank operation is justified to enter a market where firms are already there.

A coordination failure arises when each firm's gains depend on other firms' actions. Firms need to coordinate their actions to benefit. For example, a supplier will not build a plant without buyers, while buyers will not enter without a supplier. Either everyone participates or no one does, and both outcomes can persist. A PSD intervention would shift the economy from one outcome to the other. It is singled out here because this type of market failure does not cleanly fit the proposed test.

Whether a PSD intervention can address an identified market failure depends first on how easily evidence of that failure can be checked. Some market failures show up in data anyone can examine: small firms are turned down for loans while larger firms in the same sector receive them; a few firms hold most of the market with almost no new entry; property rights cannot be enforced. The companion paper calls these first-tier failures, and a project appraisal that records them has enough evidence to support the diagnosis.

Other market failures are real, but it is harder to pin down the evidence. Coordination failures and capability externalities, for example, depend on what would have happened if the Bank operation did not take place. No one can observe that. For these reasons, for slippery cases, the intervention is presumed unwarranted. The companion paper does not rule out Bank interventions to address this slippery kind of market failure, but it asks for outside validation before they clear the hurdle. Most project appraisals describe a symptom—an access gap, weak linkages, insufficient integration—restating what a failure would look like without saying what the failure is. When this is the case, it does not clear the hurdle.

An externality wedge is addressed with a subsidy big enough to provide incentives for the firm to do as the Bank operation asks and small enough that the Bank is not giving away too much. Of course, the Bank does not know either the size of the wedge or the firm's costs and guesses or negotiates with the firm to determine what it would take to get the firm to act. This is likely to be inefficient. While it is hard to know, it is likely that the subsidies are bigger than needed.[3] In most cases, competition among firms is a better way to determine the size of the required subsidy.

Sometimes firms face risks whose potential gains and losses are hard to estimate. Paying the firm to induce action does not work in this case because the firm has to protect itself against the worst-case scenario and will ask for more than it needs to insure itself against the biggest losses possible. Changing the length of the contract is likely to be more effective at incentivizing action. Firms can bid on the smallest discounted stream of support they need to carry the risk, and the smallest discounted need wins. If the year turns out badly, the contract is extended; the payment does not go up.

The third listed property asks what party the operation affects. With PSD operations, it is the firm. This distinguishes PSD from the other operations the Bank runs. The others mostly contract with a public sector entity. When a PSD operation's objective is job creation, for example, the firm does the hiring and that is why it is the same party paid (subsidized) by the PSD operation. The firm keeps the residual amount of the subsidy, so it knows the smallest support at which it would act and the Bank does not. This is addressed more fully in Section VII.

The admissions test

The definition provides an admission test with two questions designed to conclude whether the proposed operation counts as PSD. One of the three properties—the questions listed above—is not included in those questions. An operation that transacts with no residual claimant is not a PSD operation. The test is unnecessary.

That leaves two questions. Is an incumbent supplier present? Is an externality wedge claimed? A no to the first means the operation does not count as a PSD operation. As stated above, an operation is

3 Chile's Telecommunications Development Fund set a maximum subsidy for each rural payphone project from the regulator's own cost estimate, then awarded the project to whichever company asked for the least. In the first two rounds the winning bids averaged 41 percent of that maximum, and 55 percent of all bids asked for no subsidy at all. By the last round, when only one project of 27 drew more than one bidder, awards averaged 95 percent of the maximum. Björn Wellenius, *Closing the Gap in Access to Rural Communications: Chile 1995–2002*, World Bank Discussion Paper 430 (Washington, DC: World Bank, 2002), 17–18 and table 8.

supplying a public good in a market with no incumbent supplier. A no to the second means that public entry is not licensed, there is no need for it, and the operation does not count either.

A wedge must be claimed for an operation to count as a PSD operation. Claiming a wedge does not mean it has been confirmed against the evidence. Confirmation occurs at the justification hurdle.

What falls outside the definition—the exclusions

Upstream and advisory work is outside the boundary that defines a PSD operation. This work is seen as central to PSD because a substantial share of PSD staff do it. The PSD definition requires a non-sovereign counterparty and a price, and neither upstream nor advisory work meets those requirements. Without a price, this work is a public good carrying the private sector development label.

Wholesale credit also falls partly outside the PSD boundaries because these Bank operations lend through a financial intermediary rather than directly to a firm. The intermediary is the Bank's counterparty but not its residual claimant. The operation has an obligation to show that the Bank's financing has additionality because the intermediary operates in a market where other lenders work.

Bank PSD operations that fall outside the definitional boundary in a different way are competition policy, trade liberalization, and deregulation of entry. These three activities address market power in markets where prices exceed marginal cost but the market still functions. Addressing market power requires pro-competitive reforms that differ from the other two because it seeks to change the conduct of an incumbent supplier and, in some cases, to displace that supplier. Most private sector development operations have the opposite requirement; they are expected to provide something the market does not already provide without displacing existing private suppliers.

The activities are also distinct from other private sector development work because none of the three places public money in a market. These operations implement reforms to the rules governing the market through technical assistance, and the public sector as the counterparty. Because these operations provide no public money a firm and the firms are not the contractual party to the Bank, that is the reason why the question of displacement of the incumbent already in the market does not arise.

Excluding these three activities from what is a private sector development operation would result in a narrower definition than the operational test from which the definition is derived. The operational test applies whenever an operation enters a market in which private suppliers are already present. Since a supplier is present in these types of operations the possible target of displacement, the operation would fail the test and not defined as a PSD operation. That means that the definition requires something more to define these operations as PSD. The operation must also identify an externality, meaning a difference between the private and social return, as the reason for intervention and be defined as a private sector development operation.

The common property: what operations share

Trade facilitation and export development are considered private sector activities. However, they are not specifically listed under Major theme 4, Financial and Private Sector Development in the Bank's historical taxonomy. Among operations that have been coded with those themes, 68 percent have also been coded with Major theme 4. The remaining 32 percent have not. (Table 1).

TABLE 1. The private sector development portfolio in the Bank's taxonomy, FY1975–2025

Bank Sub-theme (Major Theme 4, Financial and Private Sector Development)	Ops	US$bn	Price Set By	Allocation of Support Among Firms
State-owned enterprise restructuring and privatization	643	91.3	negotiation; competition where assets are auctioned	discretion
Regulation and competition policy	935	136.0	no priced transaction	none at firm level
Infrastructure services for private sector development	1,083	171.1	bilateral negotiation	discretion
Micro, Small and Medium Enterprise support	403	36.3	bilateral negotiation, intermediary and borrower	intermediary discretion
International financial standards and systems	213	43.0	no priced transaction	none at firm level
Corporate governance	73	12.8	no priced transaction	none at firm level
Other financial and private sector development	541	73.8	heterogeneous	heterogeneous
Major theme recorded without a named sub-theme (median approval FY1989)	937	123.6	—	—
Distinct operations that have the major theme	3,489	489.4		

Notes: The seven named sub-themes are the set that defines the historical classifier used in Section II: every operation that has any of them carries Major theme 4, and the classifier is Major theme 4 itself. Major theme 4, Financial and Private Sector Development, is the Bank's historical classifier and is not theme code 120000, which Section III applies to the FY24–25 portfolio; the historical taxonomy does not separate financial sector lending from private sector development. Operations labeled with more than one sub-theme appear in each row they carry; the final row reports distinct operations and commitments.

Sources: Author's calculations from World Bank Projects API theme assignments over the FY1975–2025 lending universe described in Section II. Commitment amounts are nominal, recorded as of board approval, with no adjustment for inflation across the fifty-year period; see the methods annex. Abbreviation: API, application programming interface.

An intervention that occurs where an incumbent is present is classified as a private sector intervention. An intervention that occurs where no incumbent is present is classified as public good provision. The distinction matters for industrial parks and special economic zones, value chain and supplier development, and trade facilitation.

II. The performance question

IEG rates operations after closure, assessing achievement of their appraisal objectives and alignment with country plan objectives. This paper examines IEG ratings data to answer four questions: Do private sector operations rate differently from non-private sector operations? Do shifts in PSD orthodoxies affect ratings? What happened to portfolio composition across these changes in orthodoxies? What explains changes to PSD ratings?

Finances One records 12,724 World Bank operations approved from FY1975 through FY2025. The IEG project performance dataset contained 12,570 completed evaluations as of 30 June 2026. Linking the two by operation identifier, the author retained 9,490 operations with codable outcome ratings. Of these, 2,780 carry Major theme 4, the Bank's legacy code.[4]

Private sector operations average 4.01 on IEG's six-point scale, against 4.04 for the rest of the portfolio. The estimated gap is 0.02 points without controls. Adding country, instrument type, and evaluation-year fixed effects changes the gap to 0.03 points. Neither estimate is statistically significant.[5]

With standard errors clustered by country, the minimum detectable difference is about one-tenth of a rating point, more than three times either estimated gap. The share rated moderately satisfactory or better is identical in both groups, at 72 percent.

Private sector operations have mean ratings of 4.02 for directed credit approvals, 4.06 for adjustment and privatization, 3.97 for investment climate, 4.00 for jobs and inclusion, and 3.95 for Cascade approvals. A test of equality across the five waves does not reject the null ($F = 0.70$, $p = 0.59$).

The post-2016 taxonomy, theme code 120000, has approvals going back as far as 1990, but the same test just described produces $F = 2.97$ ($p = 0.03$) over the four waves. A higher Cascade-era mean of 4.15 drives the difference, but the range in wave-means is still only 0.16 rating points.[6]

Private sector credit commitments have surged over the same waves, but there was no indication that changes in quality ratings of operations preceded the adoption or abandonment of any of the

4 The six-point scale from highly unsatisfactory to highly satisfactory is mapped to the integers 1 through 6; the published series uses the six-point vocabulary throughout. The historical theme taxonomy does not separate financial sector lending from private sector development, so the series is read as covering the two together. The waves periodize the lending record; Table 2 gives their boundaries. The companion paper counts at least six orthodoxies as theories of the function—privatization, the investment climate, financial inclusion, private capital mobilization, institutional evolution, and industrial policy—and the two sequences do not correspond one to one; the count of five in the text refers to the waves that the ratings series spans.

5 The coefficient on the private sector flag is −0.02 ($p = 0.57$) without controls and −0.03 ($p = 0.38$) with the fixed effects, standard errors clustered by country, N = 9,490. The cluster-robust standard errors are 0.038 and 0.033; at 80 percent power and a 5 percent two-sided test the minimum detectable difference is 2.8 standard errors, roughly 0.11 and 0.09 of a rating point.

6 Operations approved under the current scorecard have not yet reached evaluation, so the sixth wave contributes to the composition series but not to the ratings series.

five orthodoxies. Across the five waves, credit commitments rose from 18 percent to 58 percent. Then back down through 43 and 33 to 20 percent under Cascade. To 17 percent today.[7] The Bank restructured its portfolio around each orthodoxy. And the reported performance of those operations budged less than a tenth of a rating point for five decades. None of the five orthodoxies were preceded by a change in ratings outlook.[8]

The range of ratings given to operations has narrowed considerably since the 1980s. The variance of outcome ratings has halved to 1.0 or less for both private sector and non-private sector operations. Ratings are now more clustered around the middle of the scale: 12 percent of directed credit wave loans and 71 percent of private sector operations approved since 2020 earned the two middle ratings, compared with 57 percent for the rest of the portfolio.[9]

A rating scale with most ratings in two adjacent categories leaves little room to register differences in means, whatever the underlying performance. Comparison of mean ratings from wave to wave is therefore a weak test when ratings are compressed. A system that gives seven out of ten operations one of two middle ratings leaves too little distance between successful and unsuccessful operations.

Ratings differ little across waves, but operations within the same wave need not receive similar ratings. The pre-registered test examines whether disclosure of the terms of proposed private sector finance is associated with a different outcome rating. The sample comprises 300 closed operations drawn at random within each wave. Each appraisal was coded for disclosure of the proposed private sector financing terms, using the fixed protocol. Outcome ratings remained hidden throughout coding, and the author adjudicated all uncertain cases.

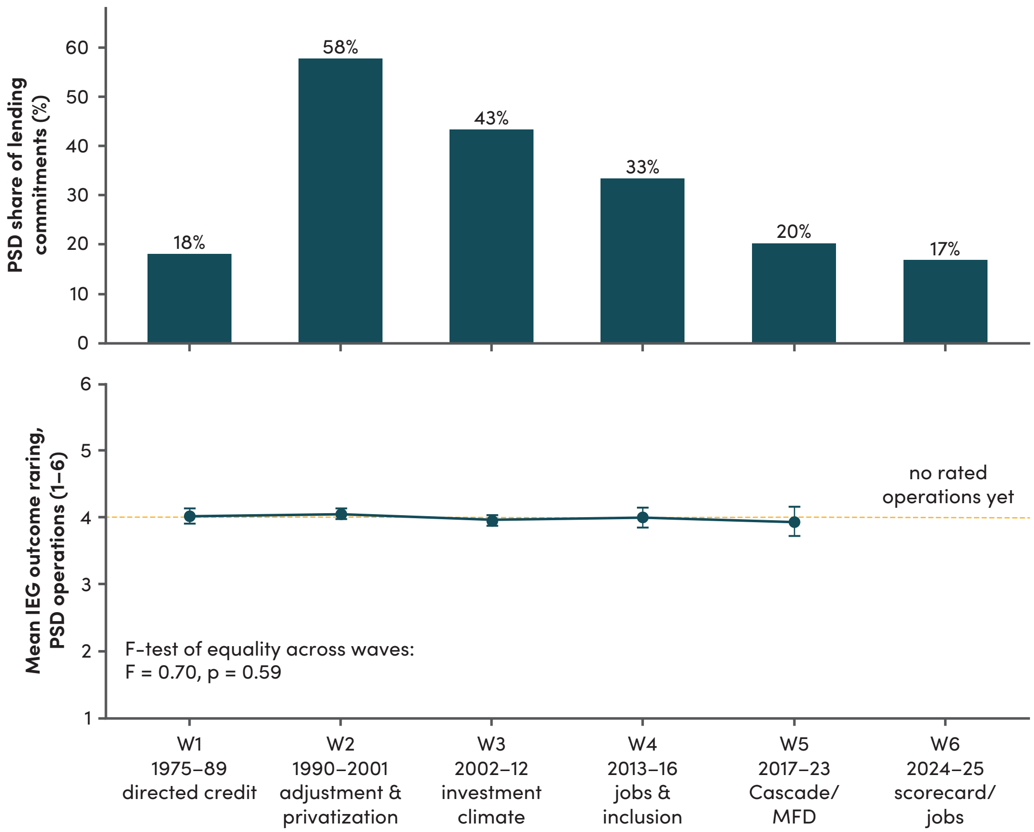
Among the 300 sampled operations, 135 substantially finance private participants. Only 36 of these disclose the subsidy. Their ratings average 3.86, against 3.93 for nondisclosers. The gap of 0.07 points is statistically insignificant ($p = 0.79$), and its sign changes across the three pre-registered cuts. The test has power to detect a difference of roughly two-thirds of a rating point, so the null result does not rule out smaller differences.

7 By operation count, the swing is from 17 percent to 46 percent and back to 16; commitments are used in the text because adjustment-era operations were large.

8 −0.27 ($p = 0.03$) and −0.35 ($p = 0.01$) relative to the first wave, standard errors clustered by country. Under the post-2016 theme taxonomy the estimate for the jobs-and-inclusion wave survives (−0.32, $p = 0.001$) and the Cascade-wave estimate does not (−0.06, $p = 0.54$).

9 Distribution-sensitive tests reflect the compression and do not contradict the flatness: an ordered logit rejects the equality of the full ratings distribution across waves ($p < 0.001$) while the equality of the means stands, because the two middle categories hold 12 percent of private sector ratings in the first wave and between roughly one half and two thirds in the waves after 2002.

FIGURE 1. The private sector share of lending and the measured performance of private sector operations, FY1975–2025



Notes: The top panel shows the private sector share of lending commitments by vocabulary wave; the bottom panel shows the mean IEG outcome rating of private sector operations on the full six-point scale with 95 percent confidence intervals. Operations approved under the current scorecard have not reached evaluation.

Source: Author's calculations from the World Bank Projects API lending universe and the IEG project performance dataset described in the text.

TABLE 2. Portfolio composition and measured performance by vocabulary wave

Wave	PSD Share of Commitments (%)	PSD Share of Operations (%)	Mean PSD Rating	SD	Rated PSD Operations	Mean Rating, Non-PSD
W1 1975–89 directed credit/DFI-SME	18	17	4.02	1.37	555	3.96
W2 1990–2001 adjustment and privatization	58	46	4.06	1.27	1,041	4.10
W3 2002–12 investment climate	43	35	3.97	1.00	898	4.00
W4 2013–16 jobs and inclusion	33	27	4.00	0.92	167	4.21
W5 2017–23 Cascade/MFD	20	16	3.95	1.21	119	4.24
W6 2024–25 scorecard/jobs	17	17	—	—	—	4.20

Notes: Waves follow the vocabulary periodization described in the text. Ratings are IEG outcome ratings on the six-point scale for operations approved in each wave. Operations approved under the current scorecard have not reached evaluation, and the non-PSD entry for that wave rests on five early evaluations. Abbreviations: DFI, development finance institution; MFD, Maximizing Finance for Development; SD, standard deviation.

Source: Author's calculations from the World Bank Projects API lending universe and the IEG project performance dataset described in the text.

TABLE 3. Populations, frames, and samples used in this paper

Population or Sample	N	Purpose	Result Reported
Panel A. The historical record, FY1975–2025 (Sections I and II; the Bank's pre-2016 major theme)			
Lending operations approved	12,724	Universe for the historical series	—
Carrying the Financial and Private Sector Development major theme	3,489	Section I, the Bank's taxonomy	937 carry no named sub-theme
With a codable IEG outcome rating	9,490	Section II, the ratings series	4.01 private sector, 4.04 rest
of which carrying a private sector theme	2,780		
Closed operations drawn at random within waves	300	Section II, pre-registered disclosure test	36 of 135 disclose; gap 0.07 points, p = 0.79
Panel B. The FY24–25 cross-section (Section III; theme code 120000)			
Bank-approved operations	964	Universe for the cross-section	—
Carrying the private sector development major theme	156	The reported portfolio	—
With a protocol-eligible appraisal document	111	Section III, the coding frame	21 failures, 19 percent
Entering a commercially supplied market	29	Section III, the domain of the test	21 fail, 72 percent
with a majority of commitment in such a market	10	Concentration check	5 fail, 50 percent
with at least 90 percent of commitment in such a market	7	Concentration check	4 fail
Frame drawn on capital the Bank allocates rather than commits	35	Sensitivity to the commitment basis	27 fail
Panel C. Subsamples read a second time (Section III; drawn from Panel B)			
Read by four readers under the same protocol	30	Inter-coder reliability	9 classified identically; a failing majority on every reader's set
Drawn at random from the operations the gate excluded	15	Whether the gate excludes wrongly	None reaches the 50 percent threshold

Notes: The table lists every population, frame, and sample the paper draws on, in the order the sections use them, and reports for each the result stated in the text. Indented rows are subsets of the row above them. Panel A rests on Major theme 4, the Bank's pre-2016 Financial and Private Sector Development classifier, and Panel B on theme code 120000; the two classifiers are not the same set, and Section I describes the difference. Panel C is drawn from the operations of Panel B.

Sources: Author's calculations from the World Bank Projects API lending universe, the IEG project performance dataset, and Bank appraisal documents, each described in the text.

III. Where operations fail the test

A PSD operation must avoid displacing private activity to pass the test. Even if it does not displace private activity, it could still fail if it does not address the binding constraint or market failure.

A lending operation designed to intervene in credit markets must satisfy two conditions to establish additionality. First, the operation must provide credit to firms unable to obtain private financing on terms that permit them to invest. Second, the Bank's lending must not displace credit that private lenders would otherwise provide.

Some firms would not be financed by any lender because expected returns are too low to cover the cost of capital. In particular, a microenterprise that has operated for thirty years and employs only family members is unlikely to qualify for bank lending. That same enterprise exists mostly to smooth household consumption and wait out improvements in the job market. The function of that enterprise is closer to what a transfer payment would do and is the province of a social protection practice and does not belong in the private sector lending portfolio.

The additionality test is meant to identify cases where a firm could be financed, but it faces constraints a PSD operation could address. For example, a young firm could have a sound business model but no collateral or established credit history, leaving it with no financing for its business. To make a loan, commercial lenders often require collateral, an established credit history, or both. If lenders were willing to bear or price the risk differently, this firm would get financing. Under alternative terms, the firm would receive financing, but under prevailing commercial terms, the firm would not. This divergence between the two outcomes defines the market failure.

To incentivize lending to this type of firm requires a public subsidy. Since public money induces private lenders to do something they would otherwise not do without the loan, there is clear additionality with this intervention.

The example was meant to illustrate that there are criteria, four of them, that an appraisal should meet to demonstrate additionality. The first is that the specific market failure causing the firm to be unable to secure financing must be identified. Secondly, the subsidy embedded in the Bank's support must be revealed. This includes both the amount of subsidy provided through risk assumption as well as any subsidy provided through concessional pricing. Thirdly, a dollar or percent value must be assigned to that subsidy. Finally, the appraisal must both establish that the subsidy is needed (i.e. private lenders would not provide a loan without it) and show that the subsidy is minimized (no more subsidy than is necessary to get the bank to lend to this firm).

What the appraisal documents record

World Bank theme code 120000 was applied to 156 operations in FY24–25. Forty-five are removed because the operations lack an appraisal document. This leaves 111 operations for which there exists an appraisal-stage document that meets the coding protocol.

Additional exclusions eliminate most operations before application of the additionality test. Forty-three are not PSD: they fall in other areas of Bank work (conventional public investment, social protection, public-sector capacity building, financial stability). There are 23 budget support operations. Those are removed and subject to a different test. Section IV discusses that different test for budget support operations. Six have no World Bank commitment: financing comes entirely from outside the World Bank Group. In another ten, Bank capital does not enter a market served by commercial firms, so they fall outside the definition of what counts as a PSD operation. Twenty-nine operations remain after all exclusions.

The main result reported in this paper is that 21 of the 29 remaining operations fail the additionality test (Table 5).[10] The number, instead of the rate, is chosen as the value reported because the rate depends on the denominator used. Against the base of all 111 operations with acceptable appraisals, 21 failures yield a failure rate of 19 percent. Against the population of 29 operations that make it to the additionality test, that is a failure rate of 72 percent.

Consideration of whether Bank capital enters a market served by commercial firms narrows the comparison further. Ten of the 29 operations place a majority of their commitments in a commercially supplied market; the remaining nineteen route only a minority there while still meeting the entry criterion. Of the ten majority-commercial operations, five fail to demonstrate additionality, a rate of 50 percent.

Twenty-nine operations have total commitments of US$4.92 billion. Failed operations account for US$3.20 billion (65%), while operations that pass account for US$1.72 billion (35%). Of that US$4.92 billion, only US$2.64 billion enter commercially supplied markets. Failed operations contribute US$1.17 billion of that US$2.64 billion.

Value chain interventions, investment climate reforms, entrepreneurship support, and foreign direct investment (FDI) facilitation often misdiagnose the market failure. A value chain intervention, for example, may misidentify which party's behavior causes the failure. In a production network, the constraint generally lies with the buyer. But many value chain interventions attribute the failure to the seller, misdiagnosing the constraint. When a lead firm is its suppliers' major customer, it may drive supplier margins below levels that permit recovery of investments in capabilities or quality. Treating the resulting lack of investment as a supplier-capacity gap directs the subsidy to the supplier even though the buyer with market power created the gap and captures gains from the upgrade. Upgrading suppliers to make them attractive to competing buyers would undercut the monopsony. These four types of operations are nonfinancial private sector development interventions. Wholesale credit lines, the fifth type, are financial sector operations but fall under the PSD umbrella because they support private sector development.

Seven failures are value-chain operations, seven are financial intermediaries, and seven are neither value chains nor financial intermediaries. Four value-chain operations and one revolving credit facility include matching grant windows. Two non-value-chain operations include grant windows as well. Only a handful of failures include the wholesale credit line, the typical private sector instrument discussed above.

The Independent Evaluation Group's 2023 evaluation of IFC additionality in middle-income countries found, across 579 investment operations assessed during FY11–21, that IFC typically delivers financial additionality but struggles with nonfinancial additionality (87 percent of

10 Data construction, retrieval dates, the commitment basis, and the robustness counts are set out in the methods annex.

projected financial additionalities materialized compared to 63 percent of projected nonfinancial additionalities) and identified gaps in its processes for monitoring additionality.[11] The IEG's concept of nonfinancial additionality clears the first two hurdles of this paper's criteria, so the gap they identified is in fact futility, money that lifts no binding constraint, and, as in this paper, asymmetry. One operation fails on additionality of financial resources alone. When looking specifically at the Bank's track record on mobilizing private investment in green energy, the German Institute of Development and Sustainability found that targets set out in country strategy documents for lending and disbursements are often unambitious, not aligned with the country analysis or too vague to incentivize reforms.[12]

Wholesale lines of credit are designed to address information asymmetries in lending, a market failure that causes credit rationing by firm size and sector. They lend through a participating intermediary to provide credit at scale, but do not necessarily change how the intermediary underwrites risk. The operation may therefore leave the market failure unaddressed. Bank funds can also replace commercial funding the intermediary could have raised on its own in the well-developed capital markets characteristic of many middle-income countries. These concerns are not captured by the Bank's rating system because a wholesale credit line is scored narrowly on disbursement, which means wholesale lines of credit can replace private credit without addressing the market failures that justified the intervention.

The test's results are clearest for two operations approved in the same country and fiscal year, both financing firms through a national development bank and both carrying theme code 120000. The Financing Adaptation for Growth operation (P181797) is a US$409 million IBRD guarantee (€360 million at the operation's approval-date exchange rate) and that same figure represents the Bank's exposure to the Industrial Development Bank of Türkiye (TSKB). The Bank provides a rolling first-loss guarantee covering up to 60 percent of the exposure. The operation is designed to mobilize US$681.7 million (€600 million) from commercial lenders.

The operation was justified by noting that international lenders operate in that market but do not lend at the tenor TSKB requires. The appraisal notes that Turkish banks have three-quarters of their liabilities at maturities below three months and less than one-tenth of their loans at maturities above five years. The operation helps fill the gap in long-tenor finance for firm-level adaptation investment.

11 Independent Evaluation Group (IEG), International Finance Corporation Additionality in Middle-Income Countries: An Independent Evaluation (World Bank, 2023). The evaluation analyzed 579 IFC investment services projects evaluated during FY11–21 and found that 87 percent of anticipated financial additionalities were realized, against 63 percent of anticipated nonfinancial additionalities. The asymmetry matches the portfolio finding in this paper, where one operation fails on financial additionality alone.

12 Jürgen K. Zattler and Adrian Schmieg, Does the World Bank Group Efficiently Promote Private Sector Investment? The Case of Energy Transition, IDOS Discussion Paper 22/2025 (Bonn: German Institute of Development and Sustainability, 2025), summarized in Jürgen Zattler, "Is the World Bank Delivering on Its Private Investment Promise? The Case of Green Energy," Center for Global Development, July 2025.

Evidence on the guarantee's additionality comes from the lenders themselves. In late 2024, while the operation was being prepared, the Bank team approached 12 multinational banks. Four gave indicative terms, but only on condition that additional cover be available for the exposure beyond the 60 percent protected by the guarantee. The team found little of this type of cover was available privately. Early in 2025, the Bank requested expressions of interest. Seven of the 11 banks approached then submitted priced quotations. Banks moved from indicative terms to priced offers only after the guarantee covered the risk that had kept additional cover unavailable. The paper treats this as clearing the effectiveness hurdle.[13]

Demand for longer-term lending provides one justification for the operation. The team also identifies an adaptation externality, where neither the firm nor the lender captures or prices all of the benefits from greater resilience, as an additional justification for the operation. The Bank assumes part of the lending risk through the guarantee and charges fees under a published schedule.

Another Türkiye operation in the financial sector, approved months after the first project, clears hurdle 1 but not hurdles 2 and 3. The Second Formal Employment Creation Project (P507276) is a roughly US$500 million IBRD loan to TKYB, the Türkiye State Development and Investment Bank. TKYB will on-lend the funds to businesses in the provinces affected by the 2023 earthquakes. A disruption to credit supply provides a plausible first-tier justification for intervention, so the operation clears hurdle 1. The operation fares worse on clearing hurdle 2. The appraisal states that sub-loans are priced at market rates and that TKYB will bear the credit risk on its own underwriting criteria. To pass the effectiveness hurdle, the appraisal would have had to demonstrate how Bank financing alters TKYB's actions.

Though the appraisal provides information on the pricing of sub-loans, the documentation cannot serve as evidence of a subsidy. The results framework tracks firms reached, jobs created, and the share of firms maintaining or increasing employment. The data sheet labels the operation neither MFD-enabling nor private-capital-enabling. TKYB supplies credit, but the appraisal does not establish what commercial lenders would have supplied without Bank support. The operation therefore also fails the financial additionality hurdle.

In Kosovo, a US$5 million partial credit guarantee forms part of a US$52 million operation. The guarantee supports firms facing collateral requirements above 200 percent of loan value on unguaranteed credit. Guinea's US$100 million fisheries operation also contains a partial credit guarantee window, and its predecessor reports new borrowers making up 49 percent of all borrowers. Both windows disclose their prices and additionality rationales. Neither whole operation is primarily PSD, though each includes a component which clears the test. This is the unit-of-classification problem discussed in Section I.

13 The two operations are described in the methods annex.

The eight passing operations put more Bank capital into commercially supplied markets than the 21 failures: US$1.46 billion against US$1.17 billion, measured at component level. The average is US$183 million per passing operation and US$56 million per failure. Seven operations commit at least 90 percent of their total financing to a commercially supplied market. Of these, three pass and four fail. The three passes account for US$1.36 billion of the seven operations' US$2.24 billion in total commitments, a little over three-fifths.[14]

Each of the four readers applied the same protocol to a random subsample of thirty operations. The author and Gemini agree at the gate on nineteen cases, with Cohen's kappa of 0.34 (bootstrap 95 percent confidence interval: 0.12 to 0.60). When the decision distinguishes pass, fail, and exclude, they agree on fourteen cases. The corresponding kappa is 0.13, with a 95 percent confidence interval from −0.01 to 0.32.

Gemini excluded several operations the author admitted: fisheries and environment management, transport and tourism infrastructure, and two fiscal development policy loans. The author admitted these operations on the strength of the private sector language in their appraisals. Of the seven operations admitted by both, Gemini failed six. Claude admitted three, all failing at justification. Gemini and Claude agree at the gate on twenty-six of thirty operations (kappa = 0.53). The four disagreements concern admissions by Gemini which Claude excludes. ChatGPT conducted its reading without seeing the other readers' results.

The author and ChatGPT each admit eighteen operations, with twelve in common. Each admits six excluded by the other. Claude's three admissions all appear among Gemini's seven, and Gemini's seven all appear among the author's eighteen. Differences in strictness would explain these nested sets. They do not explain why the author and ChatGPT, admitting the same number, choose different operations.

The six operations admitted by the author and excluded by ChatGPT each direct no more than 14 percent of their commitments to commercially supplied markets. The ChatGPT reading is difficult to explain in this way. If two readers are equally strict in what they admit, then disagreeing about one-third of those admissions means that they are arguing about where the line goes.

The six operations admitted by the author and excluded by ChatGPT each direct no more than 14 percent of their commitments to commercially supplied markets. In every case the author judges those operations as PSD operations based on how the appraisal frames the operation itself. Four of the six operations that ChatGPT admits but the author excludes are budget support or results-based lending operations—all of the Bank's money goes directly to the treasury. The other two are operations with shares just below 50 percent of how much money is channeled to commercially supplied markets. All of the disagreements hinge on what counts as an answer to

14 The construction of the second, third, and fourth readings is described in the methods annex.

two questions: does an operation whose funds go directly to the treasury count and, if so, how much of the total committed funds must go to a market supplied by the commercial sector for the operation to count as substantially private sector financed? Neither question is answered by the appraisal documents. These disagreements help indicate where that boundary breaks down.

Eleven of the eighteen operations admitted and documented by the author fail to clear at least one hurdle. ChatGPT fails all eighteen operations. Fourteen of ChatGPT's eighteen failures do not clear the financial additionality hurdle. ChatGPT's notes identify additional hurdles not cleared by some operations after finding that the operation failed at the first hurdle. Since the coding stops at the first hurdle not cleared, the distribution of failures across the three hurdles depends on which hurdle was failed first. The author's coding of failures determines the distribution shown in Table 5.

The author and ChatGPT reach the same conclusion regarding the number of passes versus fails but differ more at the operation level. There is agreement on the 3-way decision to pass/fail/exclude for thirteen of thirty operations (Cohen's kappa = 0.09). They apply the same obligation to each operation admitted through its gate and agree that a majority fail.

Across the four readings, only nine operations were classified in the same way. The twenty-one discrepancies between the lists are based on differences in how each reader interpreted two questions not addressed in the protocol. The first was whether budget support and results-based lending operations should be put under the PSD umbrella even if proceeds are delivered to the treasury instead of directly into a commercially supplied market. Eight disagreements arose over this question. The other thirteen concerned the share of an operation's commitments entering a commercially supplied market. The protocol did not say how large this share had to be for the operation to count as "substantially" private sector.[15]

The addendum addresses both questions.[16] Budget support is excluded because Bank money must enter a commercially supplied market for the rule to apply. An operation counts as "substantially" PSD when at least 50 percent of its commitments enter such a market.

15 Both clauses were adopted as a dated addendum to the coding protocol on July 20, 2026; the text and its effect on the fourth reading are in the methods annex.

16 After the readings on July 20, 2026.

TABLE 4. Four readings of the 30-operation reliability subsample

Reading	Reader	Operations Admitted (of 30)	Failures Among Admitted Operations
First	Author	18	11
Second	Gemini	7	6
Third	Claude	3	3
Fourth	ChatGPT	18	18
Agreement across the four readings: operations by number of readings admitting them			
All four readings	3		
Three readings	4		
Two readings	5		
One reading	12		
No reading	6		

Notes: All four readings were performed under the original protocol; the addendum of July 20, 2026 was applied afterward. Verdicts are as recorded by each reading before the adjudication described in the text. The author's failures are also reported in the text for the full frame, where they are 21 of the 29 operations that enter a commercially supplied market. Gemini's verdicts are reconstructed from the coding memo of July 20, 2026, which records Claude's verdicts in full and the four operations on which the two readings differ; the original verdict file has not been recovered, and the reconstruction is deposited with the annex. Gate agreement and Cohen's kappa by pair: first–second, 19 of 30 (kappa 0.34, bootstrap 95 percent confidence interval 0.12 to 0.60); first–fourth, 18 of 30 (kappa 0.17, interval −0.20 to 0.51); second–third, 26 of 30 (kappa 0.53); second–fourth, 19 of 30; third–fourth, 15 of 30. Three-way verdict agreement: first–second, 14 of 30 (kappa 0.13, interval −0.01 to 0.32); first–fourth, 13 of 30 (kappa 0.09, interval −0.16 to 0.33).

Source: Author's coding records and the accompanying coding annex.

The failure counts in the thirty-operation subsample are 11 of 18 for the author, 6 of 7 for Gemini, 3 of 3 for Claude, and 18 of 18 for ChatGPT. The author's full-frame count is 21 of 29. Although the readers admit different operations, at least three-fifths of admitted operations fail under every reading. The institutional choice, discussed in Section V, is whether to set the boundary by rule or leave admission to the discretion of the party proposing the operation.

The document-eligibility check removed eleven operations after they had been coded. Their verdicts rested on Project Information Documents, a document type expressly excluded by the protocol.

The exclusions shaved nine percent from the frame into which careful reading had admitted operations, and support Section VII's conclusion about this subsection's sample. Readmitting those eleven operations does not change the conclusion: failed operations constitute 70–74 percent of the larger frame, versus 72 percent here.

The commitment-based frame excludes six operations funded entirely outside the World Bank Group, five of which place capital in a commercially supplied market. In the alternative frame based on capital the Bank allocates rather than commits, 27 of 35 operations fail. At component level, each of thirty subsample operations plus fifteen randomly selected from activities the gate excluded was reread according to a written rule concerning Bank capital's entry into commercially supplied markets. None of the fifteen operations the gate excluded cleared the 50 percent threshold, so the gate does not appear to be holding out operations that belong on the inside. Ten operations that passed the gate inject no capital into a commercially supplied market and exit the frame, all previously recorded as failures. And a prediction registered prior to this exercise, that the threshold would remove proportionally more passing than failing operations, proved incorrect. Failing operations were removed from the measurement at 74 percent, passing operations at 38 percent, evidence that runs contrary to this subsection's argument. This prediction is noted both because it was registered in advance and because it was wrong.

Passes varied, more than fails, by how many times the operations were read. Seven passed the first time they were read, in the subsample. Five of those operations were eventually admitted during at least one subsequent reading. None of them passed a subsequent time. Among those later readings, they accumulated 11 failing verdicts, and no additional passes. The author reversed three of those five initial passes.

The three reversals have one thing in common. Each reversal relates to the operation's financing plan, but the reasons for reversals on these plans are different for each. For the West African Economic and Monetary Union credit line, additionality is not corroborated by any participating lender's conduct. The Türkiye development bank loan has a similar fact pattern that has already been discussed. A Bhutan operation's appraisal cites no critical first-tier failure, and a Rwanda development policy loan's concessional element is neither priced nor demonstrated to be needed (the latter determination, in any event, is gated out by the protocol addendum).

The Ethiopia financial sector operation (US$700 million, the second-largest commitment in the subsample) was admitted and failed by three readings and gated out by Claude's. The author decided to side with the machine on this one: the proceeds are used for an equity injection by the government into its state-owned banks and to bolster the financial regulator, and the appraisal explicitly states three times that no operation funds will be used to support a line of credit or other commitment. Bank capital is not entering any commercially supplied market here. It gets gated out, labeled a financial-stability operation rather than a private sector one, and the failing percentage stays at 72 percent.

The rules governing where the Bank's money would go are what differentiated the readings. Once those rules were added to the protocol, the differences across readers narrowed.

TABLE 5. The FY24–25 private sector development portfolio against the test

	Operations	Commitments
Panel A. From the reported portfolio to the test		
Bank-approved operations, FY24–25	964	
Private sector development as a major theme	156	
No disclosed appraisal-stage document	34	
With a disclosed appraisal-stage document	122	
With a protocol-eligible appraisal document	111	
Document not of a type the protocol admits (coded, then removed)	11	
Not substantially PSD	43	
The rule-writing function: budget support and results-based lending	23	US$10.21 billion
No Bank commitment (financing recorded as non–World Bank Group)	6	
No Bank capital entering a commercially supplied market	10	US$1.61 billion
Enters a commercially supplied market	29	US$4.92 billion
Cannot demonstrate additionality	21 (72%)	US$3.20 billion (65%)
Demonstrate additionality	8 (28%)	US$1.72 billion (35%)
Panel B. The 21 operations that cannot demonstrate additionality		
By first failing hurdle		
Justification	16	
Effectiveness	4	
Financial additionality alone	1	
By failure mode		
Subsidy undisclosed or unpriced (12 justification, 2 effectiveness, 1 additionality)	15	
No supported first-tier market failure (all justification)	3	
Social assistance or disaster relief misfiled as PSD (all justification)	1	
Market-rate credit line, additionality unverified (effectiveness)	2	

Notes: Operation counts trace the portfolio from all Bank-approved operations in FY24–25 to the operations that place Bank capital in a commercially supplied market; commitment figures are total Bank commitment amounts and therefore exceed the sums entering a commercially supplied market that the text reports; the share of each commitment that enters such a market was measured at component level, as described in the text. The row reporting operations removed by the eligibility check covers only those that had carried verdicts before the check was applied. Operations governed by the rule-writing test and operations placing no capital in a commercially supplied market are reported as separate rows rather than folded into the excluded count. Classification from appraisal-stage documents by a single reviewer; the accompanying coding annex records each verdict. Parentheses in the mode rows give the joint distribution across the first failing hurdles.

Source: Author's coding of World Bank Projects API data and Bank appraisal documents.

IV. Three institutions, one obligation

Firms in the private sector can be reached by three World Bank Group institutions. IBRD and IDA lend to governments, so their private sector work goes through the public sector in the form of a policy loan, an investment operation, or a state bank line of credit. IFC invests in companies and banks without a public sector counterparty. MIGA insures political risk, taking on risks that keep commercial lenders and investors out, and enhances credit.[17]

Two of the three have additionality policies, but the published policies are more like guidance because they seem to have no teeth. IFC has an explicit, published, and assessed doctrine. The 2018 harmonized framework asks for a narrative description of the additionality, supported by market knowledge, as part of the operation description. Section I sets out why the Bank sits outside a standard it approved; what matters here is that the standard, as written, cannot be checked at the point of approval.

MIGA has provided insurance against political risk since 1988. It began offering insurance against debt default by a host government in 2009 and expanded that coverage in 2013. In 2019, debt default insurance comprised more than one-third of MIGA's total outstanding exposure. At that time MIGA had insured debt default across 34 projects totaling US$11.3 billion in 14 countries. The Independent Evaluation Group evaluated all 34 of these projects in 2022.[18] It found that private insurers and export credit agencies had been selling the same cover since the 1980s and were already selling it when MIGA sought approval in 2009; that MIGA holds about 8 percent of the market for it while private insurers hold close to 80 percent; and that private capacity to insure bankable projects in developing countries is not scarce, even in countries riskier than those MIGA will insure. It was unclear if

17 Multilateral Development Banks, Multilateral Development Banks' Harmonized Framework for Additionality in Private Sector Operations (September 2018). The task force that produced it comprised the African Development Bank, the Asian Development Bank, the European Bank for Reconstruction and Development, the European Investment Bank, the Inter-American Development Bank Group, the International Finance Corporation, the Islamic Development Bank Group, MIGA, and the New Development Bank; the report records that the Asian Infrastructure Investment Bank and the World Bank were consulted and approved the document. The task force confined its scope to non-sovereign and private sector operations and excluded lending to sovereign or public sector counterparties. On evidence, the framework asks that a project description carry a narrative describing the types of additionality present, supported by information demonstrating a judgment based on market knowledge, and it records that accounting for the counterfactual can be difficult.

18 Independent Evaluation Group, The Multilateral Investment Guarantee Agency's Experience with Non-Honoring of Sovereign, Sub-Sovereign, and State-Owned Enterprise Financial Obligation Guarantees, Meso-Evaluation (World Bank, May 19, 2022). The evaluation reviewed all 34 non-honoring projects implemented between 2009 and 2019, of which eight had completed ex post evaluations and seven had been rated; six of the seven were rated satisfactory on development outcome, with economic sustainability the weakest dimension. Market share figures are Berne Union investment insurance data for calendar years 2010 to 2018. MIGA management's comments, published with the evaluation, dispute the private capacity finding on the grounds that the estimate rests on what appears to be a single brokerage report and that the capacity described is available only to the highest-rated credits; MIGA also disputes the mobilization arithmetic and the comparison of country risk ratings across its two product lines. The portfolio figures reported in the text are the evaluation's own: 86 percent of the cover by amount in high- and upper-middle-income countries, each of the 14 recipient countries having issued foreign currency bonds internationally over 2009–19, two-thirds of the amount insured in infrastructure, and issuance growth of 11 percent a year over 2011–19 against 2 percent in the nine preceding years.

the projects would have gone forward absent MIGA, as the product's additionality depends on the case and alternatives are often available from export credit agencies, other development finance institutions, and private insurers.

The review includes an example of an institution disagreeing with its evaluator about market existence, and MIGA disagreeing with its evaluator about a capacity finding. Overall, the portfolio conforms to the finding. Eighty-six percent of the insurance was deployed to high- and upper-middle-income countries. All 14 countries that received insurance sold bonds on international markets during the decade, suggesting they all had access to borrowing. Two-thirds of the total amount insured supported infrastructure; the remaining third financed commercial loans to state-owned banks, mostly for SME on-lending. Section IV considers such loans least acceptable. The review was unable to verify whether cover reduced the government's interest rate, the benefit touted at the product's launch, since MIGA did not share the loan agreements with it. MIGA developed the product to help with volume shortcomings: demand for MIGA covers against initial political risk had stagnated since 1997. From 2011 to 2019, issuance of guarantees grew by 11 percent annually, compared to 2 percent annually in the previous nine years.

This paper excludes MIGA's debt-default insurance business from private sector development because the insured party is an international bank, the borrower is a government or government-owned enterprise, and the party that hires and invests does not appear in the contract. MIGA's main product is an instance of Section I's property, as confirmed by IEG. The original political risk cover included risks from expropriation, transfer restriction, and breach of contract. An investor is the residual claimant under the Agency's original political risk cover if the insured risk is realized.

MIGA is given wide berth to intervene where it deems it necessary. The published operational rules and standards govern MIGA.[19] But these operational rules and standards do not constrain MIGA because they can be changed internally. In fact, the standards record how far Exceptional Coverage has been generalized with the changes made to them, and this is the only layer visible from outside. MIGA's project assessments and validations are commercially confidential and therefore not disclosed. Any claim MIGA makes about the additionality of its guarantees cannot be verified by anyone outside of the institution.

The ambiguity of boundaries across the World Bank Group institutions results in lax discipline in choosing an instrument. For example, if no lender is willing to price a company's risk, a line of credit from a state development bank, an IFC investment in a commercial bank, or a MIGA guarantee can get a commercial lender to price that risk.

19 Convention Establishing the Multilateral Investment Guarantee Agency (1985, as amended), Articles 11–13; MIGA, Operational Policies (effective January 6, 2015), 1.09(d), 1.10(d), 1.24, 1.53–1.55, 3.28, 3.33, 3.38; MIGA Council of Governors Resolution No. 86 (July 30, 2010), amending Articles 11 and 12. The Operational Policies replaced the Operational Regulations in January 2015; paragraph references in the evaluation correspond.

The test disciplines choice of channel by linking the instrument to the location of the market failure. Failures in competition policy or bad regulation are the responsibility of the state to fix. The sovereign must act and the Bank is the channel to support that action. IFC is the channel where a commercial firm or intermediary lender is the binding actor and the failure is poor access to capital. When no commercial actor will price the risk at any premium, MIGA provides a guarantee that makes the investment feasible. In this example, there is more than one of the World Bank Group's institutions that can address the same market failure, but the test selects the channel that requires behavior to change where the failure occurs. The interventions from the IFC, for example, may address the problem where the effects are visible downstream.

Articles of Agreement spell out each institution's ability to contract with a private party and determine price. They do not provide the discipline described above. Based on the Articles alone, IFC and MIGA's investment businesses do not have an economic boundary. Both transact with a residual claimant, both are held to the same additionality standard, and they are separated by product, an investment versus a guarantee.

Institutional discipline: Gate and mechanism

A recent example of what happens when a test goes unenforced is found in the Bank Group's experience with the Cascade Approach.[20] The exercise required staff to determine if private funding was sufficient before committing public money to financing solutions. Staff quickly learned that not answering, or avoiding, the Cascade question did not halt operations, and within a few years the Cascade became a framing exercise for country strategies.

Loan volume now factors indirectly into performance reviews with the implementation of the new Corporate Scorecard. Pressure on lending has intensified. The streamlined performance framework launched in April 2024 reduced more than 150 indicators to 22 used for tracking vertical performance. "Mobilizing Private Capital" is one of the remaining indicators. There is also a methodology under development for tracking "Private Capital Enabled." The metrics track the dollars mobilized by an operation. Staff success has never been more closely linked to the volume of capital committed and disbursed.

To balance the incentives for lending in volume, there should be a part of the Bank assigned to make sure operations are designed with additionality in mind. The Operations Policy and Country

20 The Cascade Approach was introduced in the March 2017 *Forward Look—A Vision for the World Bank Group in 2030—Progress and Challenges* (DC2017-0002) and operationalized as Maximizing Finance for Development in World Bank Group, "Maximizing Finance for Development: Leveraging the Private Sector for Growth and Sustainable Development," Development Committee paper DC2017-0009, September 19, 2017, prepared for the October 14, 2017, Development Committee meeting. The question put to staff, set out in Box 1 of that paper, is whether a sustainable private sector solution exists that limits public debt and contingent liabilities; a negative answer directs the team to policy and regulatory support, or to risk instruments, before public funding is pursued. The same paper committed the institution to build the approach into performance management, the rewards and recognition programs, and the career framework during FY18.

Services (OPCS) may be best-placed and able to do that. OPCS would need a new policy to enforce additionality. That policy would be triggered when an operation proposes to have Bank capital enter a market in which a commercial supplier is already present. The project preparation team would declare this as they begin to prepare an operation.

The declaration can be recorded in the project appraisal data sheet. What would be added under the proposed additionality policy is not another mobilization tag. The existing private capital enabling tag, and the forthcoming MFD-enabling tag, record the private capital an operation expects to bring in. The additionality question is designed to record whether there are commercial suppliers already in the market, before Bank capital enters it.

The Ethiopia financial sector operation was classified as MFD-enabling although no Bank capital entered a commercially supplied market. An operation can satisfy the MFD-enabling condition while the operation does not clear the additionality hurdle.

Staff have discretion to assign and change theme codes, creating a problem because the reporting system does record which operations were reclassified. If a policy were limited to a theme, it could lose or gain operations when their theme changes. Teams would arbitrage the system by switching theme codes, avoiding theme codes that signal that an operation will have to pass a more rigorous review and moving on to theme codes where operations are easier to clear.

To avoid problems with other variables over which operational staff have the discretion to change, the review takes operations with total Bank commitments above US$10 million, as reported in Table 5. Total commitment is used because the figure is recorded outside the operation team. For the share entering a commercially supplied market, the reviewer relies on the staff preparing the operation. Financial-institution supervision, capital markets, payment and settlement systems, and financial stability remain outside the review because these areas already have established additionality practices. Section I sets out the order of the hurdles.

Sixteen operations fail the justification hurdle. Teams would have little difficulty adding the market-based account proposed by the harmonized framework[21] to their appraisals. The harder task falls to the reviewer: establishing whether the diagnosed constraint exists. Verification would require repeating much of the team's country work, often built over months from information obtained on the ground. The reviewer is unlikely to have access to the same information. Rejection authority alone does not resolve this disadvantage, so weak and strong diagnoses risk receiving similar treatment. More complete appraisals would improve the results in Table 5 without necessarily changing operation design.

21 https://www.ifc.org/en/insights-reports/2018/201809-mdbs-additionality-framework

Gate and enforcement

A PSD operation works with firms as counterparties because firms receive the operation's support directly and keep profits generated by the operation. With this kind of operation, the objective is to change the firm's behavior to address the market failure. A subsidy buys off the firm's reluctance to change that behavior. The firm alone knows how much support it requires to change its behavior but has an incentive to make the operation team believe that amount is high—higher, at least, than what it really needs to act.

This is an information asymmetry problem. The terms on which a firm will act are private information that the Bank does not have, and the firm has no reason to reveal them truthfully when asked. Competitive allocation gives the firm a reason to ask for a lower required subsidy. Ask for too much and that risks losing the auction to another eligible firm willing to perform the same task for less. Competition reveals information about the subsidy a firm is prepared to accept.

The Bank uses competitive mechanisms to purchase infrastructure and buildings for its projects, but such mechanisms are an exception for PSD operations. The PSD subsidy window is exempt from that discipline. PSD subsidies do not generally use competition to identify the firms able to meet the operation's objective at the lowest subsidy.

The gate and the mechanism are complements. The mechanism addresses appraisals that say a subsidy is essential when it is not, while the gate addresses appraisals that do not establish a market failure. The portfolio currently permits assessment of whether a subsidy is essential but rejecting an operation that does not establish the market failure requires the gate.

The fragile-state objection

Thin markets, such as those in fragile and conflict-affected states (FCS), limit the use of this test. Typical measures of credit rationing by firm size, concentration ratios, and entry and exit rates—the variables that indicate first-tier market failures—are noisy or unavailable. The gap between what the market would have without the intervention and what it does is large enough that almost any intervention can appear additional.

The operation team does not construct the counterfactual. The counterfactual is done outside the team and should be evaluated by an external reviewer, possibly someone from the Bank's research department. Verification of the counterfactual happens after the fact and is built into the design. With these requirements, it is clear that fragile-state and low-income country operations have a higher bar to overcome, and these economies are where the evidence is likely to be weakest. But this does not make these operations exempt from this requirement. A counterfactual must still be built to be robust enough to withstand the absence of data.

The likely objection to this proposal is that the requirement cannot be satisfied in the countries where private investment is most needed, and that activity in those countries cannot be postponed until the gaps in the information are filled. The objection notes a real problem, but it does not justify an operation, because it pits two potential courses of action against each other rather than identifying a benefit that is not internalized by the firm; the argument for a public subsidy depends on the latter, not the former.

Over FY10-FY21, long-term investments originated for IFC's account in fragile and conflict-affected countries averaged 5.2 percent of the institution's long-term commitment portfolio. For MIGA, new guarantee volume for clients in fragile states represented 9 percent of total issuance. At both institutions, the six largest recipients of foreign investment among the 37 fragile states received 54 percent of IFC's investment and 60 percent of MIGA's guarantee exposure.[22] IDA's Private Sector Window was established to circumvent the limitation implied by the objection pointed to above. PSW did not add to commitment amounts in eligible countries in IDA18, and 55 percent of its total allocation was approved.[23] So, the defense is instead offered on behalf of the operations located elsewhere—that is, in markets with an incumbent supplier where displacement risk is highest. One version of the objection persists, but it is not one of sequence: the commercial pioneer entering an unproven market produces an information externality that later comers free-ride on. Evaluators found little evidence that this benefit occurred.[24]

The private sector portfolio is very likely to contract under this additional review requirement in preparing an operation in fragile and conflict-affected states. Initially, Country Directors[25] (as well as vice presidents) are likely to see lending volume and portfolios fall. But country teams will quickly learn what documentation in an operation appraisal document makes the cut, adapt, and the Bank will be leaner and focused when operations produce results.

Section VI describes the safeguards regime and three features that can be carried over to the proposed reforms. Only the Board can waive or amend an Operational Policy. A rejection may be appealed to the Board, including for PSD operations in FCS. Reviewers must be chosen from outside

22 Independent Evaluation Group, IFC's and MIGA's Support for Private Investment in Fragile and Conflict-Affected Situations (Washington, DC: World Bank, 2021), chapter 2. The shares are computed on the World Bank's harmonized list of fragile situations, which differs from the list IFC applies in its own reporting; the 37 states are those carried on the list in FY20.

23 Independent Evaluation Group, IFC's and MIGA's Support for Private Investment in Fragile and Conflict-Affected Situations. Under IDA18, covering FY18–20, US$1.37 billion of Private Sector Window funds was approved against the US$2.5 billion allocated, and most of the approvals fell in the final quarter of FY20.

24 Independent Evaluation Group, IFC's and MIGA's Support for Private Investment in Fragile and Conflict-Affected Situations. The evaluation records limited demonstration effects across the projects it reviewed, and attributes the result in part to the size of the projects relative to the needs of the countries and to the small number of firms with which the institutions could work. See also Charles Kenny, Moving Beyond Mines and Mobiles: How Can IFC Add Value in Fragile States?, CGD Policy Paper 030 (Washington, DC: Center for Global Development, 2013).

25 Now referred to as Division Directors as a result of the last World Bank Group reorganization.

the regional promotion ladder. It has precedent in the Inspection Panel itself. Inspection Panel staff's advancement is not affected by their verdicts on the operations.

The safeguards regime survives because outside petitioners have standing and can exact a cost when the Board or management ignores the rules. If local communities are unhappy with an operation applied the Bank's safeguard regime, they can petition the Inspection Panel or management directly. When a commercial bank loses out to a lower cost subsidized competitor, there is no recourse as like the one the safeguards regime provides. An additionality obligation has no counterbalancing petitioner when the obligation is not observed.

Shareholders would need to create a comparable mechanism to any issues with the implementation of the additionality obligation, such as installing something similar to the safeguards regime.

Implicit in competitive allocation is the advantage of creating a constituency that can appeal an outcome based on the application of what may be a poorly administered rule. The losing bidders are invested in the rule's integrity, and the bid documents contain the price they were ready to commit to.

The problem is that bidders only exist where the allocation mechanism is used, which is only on a fifth to a third of the portfolio. Support for the gate comes from a much narrower constituency than is needed for its long-term viability. There are no stakeholders over the rest of the portfolio. The Cascade gate failed because of the same disparity between the percentage of the portfolio covered and the percentage of the portfolio where there are claimants that can enforce the gate. Something like the proposed reform would need advocates in a broader constituency within the Bank's governance structure.

V. Conclusion

This may well be the first empirical analysis to examine a full private sector portfolio on an operation-by-operation basis using an additionality standard. Previous studies examining portions of the Bank's private sector portfolio focused on specific components of both design and implementation aspects. Gelb, Tata, Ramachandran, and Rossignol (2015) did this for 20 agglomeration operations undertaken by the Bank without evaluating individual operations based on an explicit additionality standard.[26]

Not all work studied fits a coherent definition of private sector development, but many operations meet the criteria. The Bank should require an account of additionality from every operation entering a commercially supplied market. An operation without that account should not be approved.

26 Alan Gelb, Gaiv Tata, Vijaya Ramachandran, and Ivan Rossignol, "When Agglomeration Theory Meets Development Reality: Preliminary Lessons from Twenty World Bank Private Sector Projects," CGD Policy Paper 054 (Washington, DC: Center for Global Development, 2015).

PSD operations need more impact for each dollar spent. Dwindling aid budgets require it. When an operation cannot determine how much subsidy a firm needs to act, the Bank Group is likely paying more than necessary. Firms have every reason to claim they need more. The Bank should stop asking firms what support they need and require them to reveal it through a competitive process.

The accompanying coding annex contains the protocol, recorded shares, and operation-level verdicts. The protocol can be used by other multilateral development banks to conduct the same analysis. Comparison across the nine institutions is easier because the harmonized framework commits them to the standard used here. The required data—appraisal documents and operation-level data—are disclosed by every multilateral development bank.