

The World Bank Inspection Panel at Thirty-Three: Measuring Accountability Failure

A mechanism that investigates the Bank but not the borrower who causes the harm, whose findings are advisory to a Board on which that borrower sits, and that does not systematically verify whether relief reaches the people harmed – an empirical assessment of access, performance, and a remedy that, even where it is delivered, reaches a documented handful of cases against a portfolio of more than ten thousand projects

Revised Edition — June 2026

MDB Reform Advisory | mdbreform.com

Executive Summary

In thirty-three years of operation, the World Bank's Inspection Panel received 187 complaints across more than 50 countries and formally investigated 42 of them — roughly one investigation for every 250 projects it evaluated. This paper asks the question that the Panel's own anniversary publications do not: what does that utilisation rate mean, and is the mechanism's low caseload evidence of a well-functioning system or of a structurally constrained one?

The evidence assembled here is most consistent with the second interpretation. The Panel was not designed to fail. When it was created in 1993 — the first accountability mechanism at any international financial institution — it was a genuine institutional innovation. But from its earliest years, the Bank's Management, its Board, and in time the Panel itself developed a set of practices that progressively narrowed the mechanism's reach and diminished its impact on affected communities.

The paper documents seven inter-related failure modes, five of which are developed as detailed case studies in the Annex:

- **Volume failure.** 187 complaints in thirty-three years, against a portfolio of thousands of projects and tens of millions of affected people. The attrition runs the length of the chain: 187 complaints, 131 registered, 42 investigated, and a fraction of those resulting in independently verified remedy.
- **The management response escape hatch.** Under the founding resolution, if Management admits limited non-compliance and promises corrective action before the Panel recommends investigation, the Board can decide not to investigate. Richard Bissell, a former Panel Chair, documents this loophole directly: an elaborate Management response addressed to the Board could persuade it to reduce or reject an investigation at the eligibility stage, and his own survey of early cases found that only 10 of 28 registered complaints produced any project-level impact (Section 4.2).
- **Procedural delay as containment.** The Panel codified two delay procedures — Footnote 7 and the Pilot for Early Solutions — that gave Management unsupervised time to negotiate with complainants outside any formal accountability framework. Yale Law School's Lowenstein Clinic documented in 2017 that these delays, rather than producing early solutions, typically prolonged communities' frustration while preserving the lending relationship — the pattern documented in the Annex's only case of formal Pilot use, Badia East, Nigeria.
- **The Amaravati suppression.** In October 2017 the Panel's report recommending investigation of the Amaravati Capital City project in India was published on the Bank's website and removed within days, officially described as an "inadvertent administrative error." The Board then approved two successive deferrals spanning two years. The project was ultimately abandoned when the Indian state government withdrew the loan request — the first of three procedural closures the Annex traces on the same site over the following eight years, the most recent in April 2025.

- **MAP non-enforcement.** When the Panel does find non-compliance, the remedy takes the form of a Management Action Plan — drafted by the same Management whose project produced the non-compliance, self-monitored, and closed when Management reports completion. Independent verification requires a separate Board decision and is the exception, not the rule.
- **Retaliation without recourse.** Human Rights Watch documented in 2015 that in 18 of 34 complaints reviewed, requesters reported threats or reprisals linked to their criticism of Bank projects. The Bank subsequently issued anti-retaliation guidance, but the Panel has no enforcement authority over borrower governments or implementing agencies, as the Annex's Ethiopia PBS case shows directly: the Panel's own field interpreter was imprisoned for nearly three years after the investigation he supported was published.
- **The upstream blind spot.** Beneath the Panel sits the layer where most grievances are meant to be resolved: the project-level grievance redress mechanism every Bank operation is required to establish. The Bank does not measure whether these mechanisms work. A project can report a GRM as "established" while publishing no data on complaints received, resolved, or escalated; the SEEFOR case in the Annex documents this from the inside, including the parallel digital architecture one project team had to build once it recognised its own paper-based GRM was not generating reliable information. The Panel's low caseload is partly a function of a layer beneath it that the Bank does not track.

These are not impressionistic charges. The paper's central contribution is to measure the constraint rather than assert it, drawing on the IEG master database of 10,453 evaluated projects, and three findings stand out.

- First, a Complaint Index — each region's share of complaints divided by its share of lending — places Sub-Saharan Africa at 1.37, the highest of any region: Africa generates more complaints than its lending share predicts, consistent with its weaker performance record, while the genuinely under-filing regions are East Asia and the Middle East.
- Second, within Sub-Saharan Africa the statistical relationship between project performance and complaint frequency is indistinguishable from zero (Pearson $r = -0.230$, $p = 0.101$, across 52 countries): the countries where Bank projects most often fail are not the countries that complain most often — several of the worst-performing portfolios generate no complaints at all, while complaint volume tracks civil-society capacity.
- Third, the accountability funnel — 187 complaints narrowing to 42 investigations and findings of non-compliance in roughly 38 — terminates in a stage the Bank does not measure at all, namely whether affected communities actually obtain verified remedy. The evidence is most consistent with an interpretation in which low utilisation reflects barriers to access rather than an absence of harm, though complaint data measure utilisation rather than harm directly.

One further point follows from the case evidence in the Annex and Section 4.6, and it is closer to the heart of the matter than a simple count of successes and failures. The mechanism is not one that never delivers remedy. Among the eight cases in this paper where the record permits a clear determination of whether remedy reached affected people, genuine relief arrived, substantially or in significant part,

in four of them — Uganda, Cambodia, Coal India/Vishnugad, and Badia East. That is the more generous reading available from the evidence, and it should be stated as such. But even taking it at face value, four cases of documented remedy sit against 10,453 evaluated projects and 42 investigations, in a system that does not systematically track the outcome in any of them, successes included. A remedy system that works occasionally, unsystematically, and without independent verification is not a smaller version of a working remedy system. At the scale this portfolio operates, it is functionally indistinguishable from no system at all — not because individual successes do not matter to the people who experienced them, but because an accountability mechanism is judged by whether harmed people can reliably expect relief, not by whether relief is occasionally possible.

The merger of the Panel into a single Independent Accountability Mechanism, approved by the Board on June 8, 2026, reorganises the mechanism without addressing any of the seven failure modes documented here. The new Vice-President/Director-General structure creates a single governance chain that concentrates, rather than distributes, the point at which a finding can be managed before it reaches the Board. The non-regression commitment is a policy statement, not a structural protection.

The paper closes with seven specific recommendations (Section 8), each tied to a documented break in the chain rather than a general call for stronger oversight. The most consequential three: verify and publicly report remedy delivery rather than only Management Action Plan approval; replace binary "GRM established" reporting with systematic GRM performance reporting — complaints received, resolved, outstanding, and escalated — across all Bank-financed projects; and build a structured data pathway from project-level grievance mechanisms into the new Integrated Accountability Mechanism, so a chronically dysfunctional project GRM becomes visible as a portfolio-level signal rather than only through the rare complaint that survives years later to reach the Panel. None requires new authority from shareholders.

The Inspection Panel was the world's first citizen-facing accountability mechanism for an international financial institution. It deserved better from the institution it was designed to hold accountable — and so did the communities it was designed to serve.

1. Introduction

When the World Bank's Board of Executive Directors established the Inspection Panel in September 1993, it created the first body at any international financial institution through which private citizens could hold a multilateral lender accountable for the consequences of its lending. The premise was straightforward: communities harmed by a Bank-financed project, who believed the harm followed from the Bank's failure to observe its own policies, could file a complaint and trigger an independent investigation. Over the following thirty-three years the Panel received 187 such complaints, registered 131, and formally investigated 42. Set against a lending programme that the Bank's own Independent Evaluation Group has assessed at more than 10,000 projects and over \$900 billion in commitments, those figures pose a question the Panel's annual reports do not address: does a caseload this small reflect a portfolio that rarely harms communities, or an accountability mechanism that communities rarely reach?

The question has acquired new urgency. In June 2026 the Boards of the World Bank Group approved the consolidation of the institution's three independent accountability mechanisms — the Inspection Panel, the Dispute Resolution Service, and the IFC/MIGA Compliance Advisor Ombudsman — into a single Independent Accountability Mechanism. The reform was presented as a strengthening of citizen recourse. Whether it strengthens anything depends on a prior diagnosis: are the Panel's limitations a matter of fragmented architecture, which consolidation can repair, or of structural incentives embedded in the Bank's governance, which consolidation leaves untouched? Answering that requires measuring what the Panel has actually done, and against what baseline.

The existing scholarship on the Inspection Panel is substantial but largely qualitative. Researchers and civil-society organisations — among them the Yale Lowenstein Clinic, Wong and Mayer, Human Rights Watch, and the Bretton Woods Project — have documented the Board-approval requirement, the management-response loophole, the proliferation of procedural delays, and the retaliation faced by complainants. What this literature has not done is connect the Panel's caseload to the Bank's project-performance record in any systematic way. No published study has asked whether the countries where Bank projects most often fail are the countries that generate the most complaints — the relationship that would have to hold if the mechanism were reaching the communities it was designed to serve.

This paper supplies that analysis, and in doing so makes a methodological contribution: it introduces three empirical measures for evaluating an accountability mechanism — reproducible, and applicable to any IAM, not only the Panel. Drawing on the IEG master database of 10,453 evaluated projects and the Panel's complete case register, it applies all three to the Inspection Panel for the first time. The first is a Complaint Index, which expresses each region's share of complaints relative to its share of lending. The second is a country-level correlation between project performance and complaint frequency across Sub-Saharan Africa. The third is an Accountability Funnel, which traces the attrition from projects through complaints, investigations, and findings of non-compliance to the verification of remedy. Each is constructed to be reproducible and extensible to other accountability mechanisms.

The findings are consistent and, from an accountability standpoint, troubling.

Sub-Saharan Africa — a region carrying one of the Bank's weakest project-performance records — files complaints at a rate disproportionate to its lending weight: a Complaint Index of 1.37, the highest of any region. Across fifty-two African countries the statistical relationship between project performance and complaint frequency is indistinguishable from zero: the countries where projects fail most often are not the countries that complain most often. And at the end of the chain, whether affected communities actually obtain remedy is not a question the Bank systematically measures at all. The evidence is most consistent with an interpretation in which the mechanism's low utilisation reflects barriers to access rather than an absence of harm — though, as the limitations discussion makes clear, complaint data measure utilisation rather than harm directly.

The argument proceeds as follows. Section 2 sets out the quantitative record — the Panel's utilisation rates, the Complaint Index, and the performance–complaint correlation — and the conceptual framework that follows it maps where, structurally, the accountability chain breaks. Section 3 examines the design constraints built into the Panel at its creation. Section 4 turns to the scholarship, the incentives facing the Board and borrower governments, and the structural contradiction at the centre of the mechanism. Section 5 develops the Sub-Saharan Africa access-barrier argument in detail. Section 6 assesses the June 2026 merger against this evidence, including how the integration debate itself measured the problem. Section 7 concludes, and Section 8 sets out specific recommendations. Five case studies illustrating these patterns, the fifth examining the project level beneath the Panel's own caseload, are presented in the Annex.

2. The Numbers: What Thirty-Three Years of Data Show

2.1 The Master Funnel

The funnel below sets out the headline counts; the attrition between its stages is what repays attention. Of 187 requests received between 1994 and 2026, the Panel registered 131, closing roughly 30 percent before formal entry. Of those registered, 42 proceeded to investigation — an attrition of nearly 70 percent from registration to inquiry, and of 78 percent measured from the original complaint. Yet once a case was investigated, the Panel found non-compliance in approximately 38 of the 42, close to 90 percent: the mechanism overwhelmingly validates the concerns of the communities whose cases it actually hears.

THE ACCOUNTABILITY FUNNEL	
<i>World Bank Inspection Panel, 1994–2026</i>	
Bank portfolio (evaluated)	Bank portfolio (evaluated)
Communities potentially affected	Tens of millions across 100+ countries
Complaints filed	187 — 1.5% of approved projects
Registered by Panel	131 — 30% closed before formal entry
Investigated	42 — 78% attrition from complaint to investigation
Non-compliance found	~38 of 42 investigated — communities' concerns largely validated when heard
MAP independently verified	Minority — not systematically tracked
Remedy confirmed delivered	Not systematically measured by the Bank

The last two rows are shaded red not because the numbers are wrong, but because they are unknown. The Bank does not systematically track whether MAPs are fully implemented or whether remedy reaches the communities that filed complaints. Accountability mechanisms are designed to deliver remedy. The absence of a measurement system for the final two stages of this chain is itself an accountability finding.

Sources: Panel case data (inspectionpanel.org/panel-cases/data, June 2026); IEG master database (March 2026). Complaint share vs project count: $187 \div \sim 12,500$ Bank projects approved 1994–2026 $\approx 1.5\%$. Evaluated commitment base: \$910.7bn.

The losses are therefore concentrated not at the investigation stage but on either side of it — before, in whether a complaint is filed and registered at all, and after, in whether the remedy a finding implies is ever verified or delivered. The subsections that follow, and the sections beyond them, take those two ends of the funnel in turn.

2.2 Scale Metrics: Investigations Per Portfolio Unit

Raw counts are less informative than rates. The IEG master database (March 2026, 10,453 evaluated projects, \$911 billion in commitments) provides the denominator. Across the full 33-year period, the Panel averaged 1.3 investigations per year — fewer than two per year across a portfolio spanning 100+ countries. That rate has fallen further: 2020–2026 produced an average of 0.4 investigations per year.

Metric	Value	What it means
Investigations per \$100 billion lent	4.6	The Bank lent \$22 billion for every investigation it faced
Investigations per 1,000 projects	4.0	One investigation for every 250 projects approved
Investigations per year (2020–2026)	0.4	Down from a peak of 2.4 per year in 2000–2004

Calculations: 42 investigations ÷ \$911bn evaluated commitments × 100 = 4.61; 42 ÷ 10,453 × 1,000 = 4.02. Active portfolio approximately twice evaluated portfolio; adjusted rates: 2.3 per \$100bn, 2.7 per 1,000 projects. Annual investigation averages from Panel case register.

2.3 The Declining Trajectory

The Panel's investigation rate has not improved as the portfolio grew — it has declined. The following table uses approximate data from the Panel's case register:

Period	Complaints	Investigations	Invest./year	Global S+ rate (IEG)
1994–1999	~18	6	1.0	50.5% (pre-2000)
2000–2004	~45	12	2.4	45.4% (2000s avg)
2005–2009	~40	8	1.6	45.4% (2000s avg)
2010–2014	~42	8	1.6	27.1% (2010s avg)
2015–2019	~25	5	1.0	27.1% (2010s avg)
2020–2026	~17	3	0.4	44.1% (2020s, partial)

Complaint and investigation counts approximate from Panel annual reports and case register. IEG S+ rates from IEG master database (March 2026). The 2010–2019 period shows the sharpest divergence: S+ rate at historic lows, investigation rate declining.

The 2010–2019 period warrants particular attention. The global S+ rate fell to 27.1% — the lowest of any decade — as the decade's evaluated commitments reached \$322 billion, the largest of any decade up to that point, and the portfolio expanded into higher-risk FCV country contexts. Investigations fell to five over five years. Project performance and portfolio scale both reached adverse extremes at the same time as the investigation rate declined.

2.4 Regional Distribution and the Complaint Index

The Panel's regional distribution reveals a structural imbalance between where the Bank lends, where projects fail, and where complaints are filed.

Region	Complaints	Complaint Share	Lending Share	Complaint Index	IEG S+ Rate
Sub-Saharan Africa	55	29.4%	21.5%	1.37	36.8%
South Asia	41	21.9%	16.6%	1.32	45.4%
Europe & Central Asia	33	17.6%	14.8%	1.19	50.5%
Latin America & Caribbean	34	18.2%	22.6%	0.80	48.9%
Middle East & North Africa	9	4.8%	7.4%	0.65	49.8%
East Asia & Pacific	15	8.0%	17.0%	0.47	55.8%

Complaint Index = Complaint Share ÷ Lending Share. Index > 1.0 indicates over-representation in complaints relative to portfolio weight; index < 1.0 under-representation. Complaint shares are the exact regional distribution of all 187 requests, each mapped to the country of the project it concerns; lending shares are regional shares of IEG-database cumulative commitments (about \$911bn). Sources: Panel case register June 2026; IEG master database March 2026.

A note on the denominator. Lending shares are computed from evaluated commitments in the IEG database, the most complete and verifiable regional commitment data available. An earlier version of this paper compared Sub-Saharan Africa's complaint share against its share of IDA commitments (roughly a third) rather than its share of total IBRD-and-IDA commitments (about a fifth); because IBRD lending concentrates in middle-income regions, the two denominators differ substantially, and that mismatch produced an artificially low index for Africa. The figures here use one consistent denominator — each region's share of total evaluated commitments — on both the complaint and lending sides.

The table overturns an intuition this paper initially shared. Sub-Saharan Africa, carrying the Bank's weakest performance record and about a fifth of evaluated commitments, generates 29 percent of all complaints — a Complaint Index of 1.37, the highest of any region. Far from being under-represented,

Africa is the single largest source of complaints, which is consistent with its weaker performance: more failing projects generate more grounds to complain. This matches the Panel’s own published figures, which show Africa as the largest region by complaint count (Kunanayagam et al. 2023). The genuinely under-represented regions are East Asia and Pacific (0.47) and the Middle East and North Africa (0.65). The regional figure therefore does not establish an African access barrier; if there is one, it operates within the region, among the countries and communities that never file at all. Section 2.5 tests for it there.

2.5 The Performance-Complaint Correlation: IEG Database Analysis

The reviewers of a preliminary version of this paper asked the question that the Panel's own annual reports do not address: are complaint rates correlated with project performance? If the mechanism is functioning as intended, countries with weaker performance should generate more complaints. If it is constrained in the ways this paper argues, the relationship should be absent or inverse.

Using the IEG master database (3,144 SSA evaluated projects; 52 countries with at least 8 evaluated projects), and Inspection Panel case records for SSA countries, this paper tests the relationship:

Performance Group	Countries	Avg S+ Rate	Total Projects	Total IPN Cases	Cases per 100 Projects
Bottom half SSA (S+ ≤ 35.8%)	26	29.5%	1,623	31	1.91
Top half SSA (S+ > 35.8%)	26	46.9%	1,464	24	1.64
Difference	—	-17 pp	—	—	+0.27 per 100 proj (essentially flat)

IPN complaint counts by SSA country derived from the exact case register (June 2026), each request mapped to the country of the project it concerns; regional-aggregate entries excluded. Countries with at least 8 evaluated projects included (52 countries, 55 complaints). Split at the SSA median S+ of 35.8%. Illustrative country data: Kenya 150 projects, 40.7% S+, 11 complaints; Uganda 108, 31.5%, 10; Cameroon 82, 45.1%, 6; DR Congo 72, 25.0%, 5; Nigeria 142, 31.7%, 4; Mali 96, 29.2%, 0; Madagascar 112, 33.0%, 0.

Applying a Pearson correlation across the 52 SSA countries with at least 8 evaluated projects yields $r = -0.230$ ($p = 0.101$, $R^2 = 0.053$) between S+ rate and complaints per 100 projects. A Spearman rank correlation returns $\rho = -0.224$ ($p = 0.110$). Neither result is statistically significant at conventional thresholds, and the explained variance is about five percent. The sign is negative — worse-performing countries complain marginally more, not less — but the relationship is weak and insignificant. A country’s project-performance record does not predict how often its affected communities reach the Panel.

Figure 1 plots the relationship directly. Each point is one of the 52 countries, positioned by its S+ rate (horizontal) and its complaint rate per 100 projects (vertical), with point size proportional to complaint count; the dashed line is the OLS regression.

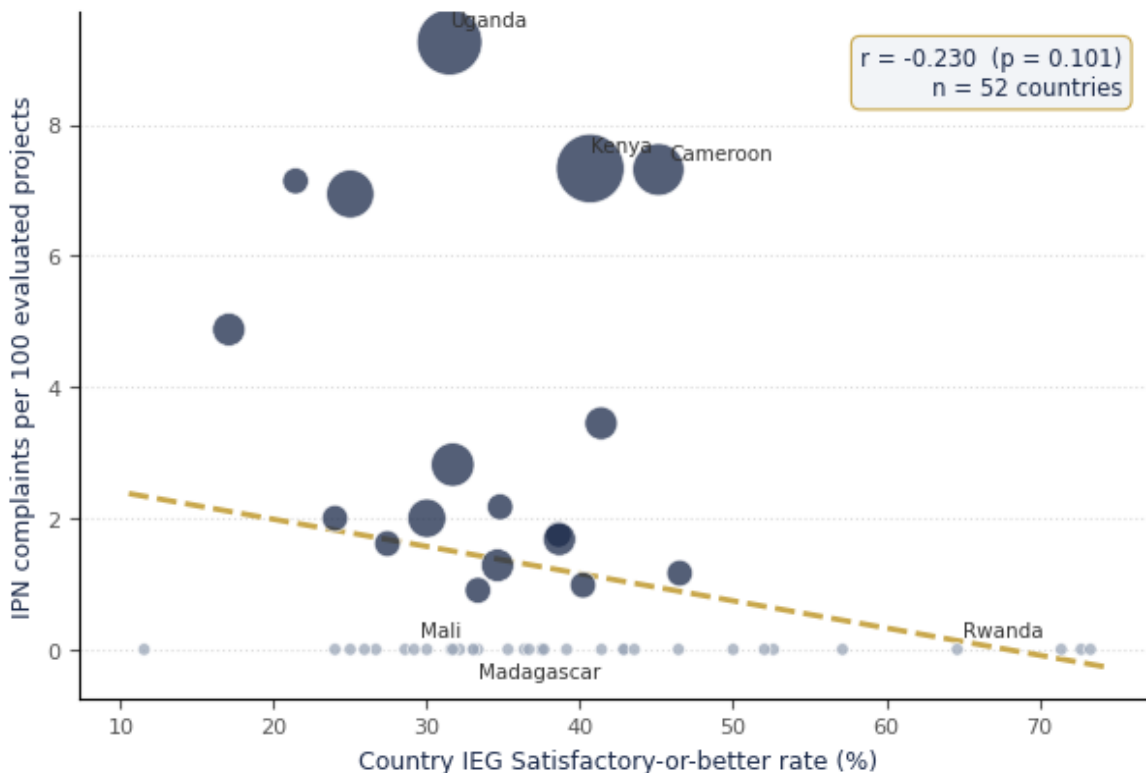


Figure 1. IEG Project Performance versus IPN Complaint Frequency in Sub-Saharan Africa (n = 52 countries with ≥8 evaluated projects). Circle size proportional to IPN case count. Dashed line: OLS regression ($r = -0.230$, $p = 0.101$, $R^2 = 0.053$). Source: IEG master database March 2026; IPN case register June 2026.

The countries with the weakest performance records cluster along the zero-complaint baseline on the left, while the regression line is almost flat — the visual signature of a near-zero correlation.

The worst-performing quintile (average S+ 22.4%) generates only 1.14 complaints per 100 projects. If the accountability mechanism were functioning as designed, that figure would be substantially higher. Countries with the worst performance records (Congo at 11.5% S+, Lesotho at 17.1%, Central African Republic at 26.7%, Mali at 29.2%) generated zero IPN complaints across their entire portfolio history.

The near-zero, statistically insignificant correlation is not a null finding in the conventional sense — it is a substantively important result. It means that complaint propensity in SSA is determined by factors largely unrelated to project performance: civil society capacity, mechanism awareness, retaliation risk, and prior experience with the Panel process. The countries associated with the weakest project performance appear to have the lowest utilisation of the accountability mechanism. This is the access barrier hypothesis stated in empirical terms.

What does predict complaint volume is civil-society capacity: the countries that file most — Kenya, Uganda, Cameroon — rank well above the regional average on the V-Dem Civil Society Participation and Core Civil Society indices and host dense transnational-advocacy networks, while the large, silent portfolios (Mali, Madagascar, Burkina Faso) sit lower and fall into “obstructed” or “repressed” civic space on the CIVICUS Monitor.

The contrast between who files and who does not makes the pattern concrete — complaint volume follows civil-society capacity, not project failure:

Country	Evaluated Projects	S+ Rate	IPN Complaints
Kenya	150	40.7%	11
Uganda	108	31.5%	10
Cameroon	82	45.1%	6
Congo (DRC)	72	25.0%	5
Nigeria	142	31.7%	4
Mali	96	29.2%	0
Madagascar	112	33.0%	0
Burkina Faso	103	33.0%	0
Senegal	128	41.4%	0
Central African Republic	45	26.7%	0

Top five rows: the SSA countries generating the most complaints — none of them the region’s worst performers. Bottom five (amber): large portfolios with below- or near-median performance that generated no complaints at all. Across all 52 SSA countries with ≥8 evaluated projects, 34 generated zero complaints, spanning the full performance range — from Mali (29.2%) to Côte d’Ivoire (52.0%) to Rwanda (64.6%).

2.6 Which Projects, and Which Sectors, Draw Complaints

The preceding analysis is conducted at the level of the country. A second test is possible at the level of the individual project. Each of the 187 requests names the operation it concerns; matching those operations to the IEG master database by project identification number resolves 136 requests to 108 distinct evaluated and rated projects. This permits a direct question: do the projects that communities complain about actually perform worse than the portfolio as a whole?

They do. The projects subject to an Inspection Panel complaint carry a Satisfactory-or-better outcome rate of 30.9 percent, against 46.3 percent for the full evaluated portfolio — a gap of more than fifteen percentage points. Put differently, 69 percent of complained-about projects fell short of a Satisfactory outcome, compared with 54 percent of the portfolio. At the level of the individual operation, complaints are an accurate signal: they concentrate on genuinely underperforming projects.

THE TWO LEVELS RECONCILED

At the country level, the number of complaints a country generates is unrelated to how well its projects perform ($r = -0.230$). At the project level, the operations that are complained about perform markedly worse than the portfolio (30.9% vs 46.3% Satisfactory-or-better).

These findings are complementary, not contradictory. When a complaint is filed it correctly identifies a poor project; but whether a complaint is filed at all is governed by access and capacity, not by where harm is most severe. The mechanism works on the cases it sees. The failure is in which cases reach it.

Distribution by Global Practice

Complaints are overwhelmingly an infrastructure-and-land phenomenon. Five Global Practices — Transport, Water, Energy & Extractives, Environment & Natural Resources, and Urban/Land — account for roughly four-fifths of all complaints mapped to evaluated projects. Social Protection, Health, Education, and Governance barely register. The table below sets the Satisfactory-or-better rate of complained-about projects against the rate for each Global Practice across the entire portfolio, so that sector difficulty is held constant.

Global Practice	Complaints	Projects	Compl. S+	Sector S+
Transport	26	16	46%	45%
Water	26	18	27%	36%
Energy & Extractives	23	19	17%	42%
Environment & Natural Resources	15	13	20%	39%
Urban, Resilience and Land	15	13	33%	41%
Agriculture and Food	8	8	25%	36%
Macroeconomics, Trade & Investment	7	5	14%	30%
All other sectors (small samples)	16	16	50%	—
Total / portfolio benchmark	136	108	31%	46%

Compl. S+ = Satisfactory-or-better rate of the complained-about projects in that practice. Sector S+ = Satisfactory-or-better rate of all rated portfolio projects in that practice.

The within-sector comparison sharpens the result. In Energy & Extractives the complained-about projects perform 25 percentage points below the sector average (17% versus 42%); in Environment & Natural Resources the gap is 19 points, in Macroeconomics 16, in Agriculture 11, and in Water 9. The

single exception is Transport, where complained-about projects perform in line with the sector (46% versus 45%) — consistent with the observation that transport complaints are driven by land acquisition and resettlement on otherwise functional road projects, rather than by project failure as such. The small-sample sectors (each fewer than five projects) are reported for completeness but should not be interpreted.

The Sub-Saharan Africa cut. Restricting to Sub-Saharan African operations, the 39 complained-about projects carry a Satisfactory-or-better rate of 17.9 percent, against an SSA portfolio rate of 36.8 percent. Even within the region that performs worst and complains least, the projects that do attract complaints are roughly half as likely to succeed as the regional average — the project-level counterpart to the access-barrier finding of Section 2.5.

Repeat-Complaint Concentration

Complaints are not evenly distributed across projects. A small number of operations generate many requests. The Odra-Vistula Flood Management Project in Poland (P147460) alone accounts for ten separate requests; the Santa Fe road and infrastructure programme in Argentina, the Mongolia Mining Infrastructure project, and the Kazakhstan CAREC road corridor each generated four; Mumbai Urban Transport, the DRC Private Sector Development and Competitiveness Project, and the Greater Beirut/Lebanon water operations each generated three. Repeat complaints signal disputes that the mechanism received but did not resolve in a way that ended the underlying grievance — the project-level signature of the remedy gap documented in the Annex.

A Coverage Gap Worth Noting

Of the 187 requests, 136 resolve to evaluated, rated IBRD/IDA projects (108 distinct operations) and 37 name operations the IEG had not rated at the time of the join — most still active when the complaint was filed, which is itself notable: communities are raising concerns about projects whose outcomes the Bank has not yet formally assessed. A further 14 fall outside the IEG evaluation universe altogether. These are not random: they cluster on IFC/MIGA private-sector operations, guarantee instruments, trust-funded grants, regional operations, and at least one project (Arun III, the Panel's very first case) cancelled before it could be built. The accountability mechanism receives complaints about a wider universe of World Bank Group activity than the compliance-evaluation system ever rates — a small but real instance of fragmented oversight.

3. Design Constraints: What Was Built In from the Start

3.1 The Narmada Origin and What Was Left Out

The Inspection Panel was created in 1993 following the Sardar Sarovar dam controversy in India. The Bank's 1992 Morse Commission found the Bank had not complied with its own policies on resettlement and environmental assessment, and recommended it "step back" from the project. Management produced an Action Plan proposing the projects move forward despite the findings. The Board voted to continue financing.

The Narmada sequence established the institutional template: independent review finds non-compliance; Management produces a response with corrective promises; the Board accepts the corrective narrative; the project continues. When the Panel was designed the following year — under US Congressional pressure during IDA replenishment hearings — it was built with exactly that governance chain preserved. The Board retained authority to approve or reject investigations. Management retained authority to produce remedy plans.

Two proposals were on the table in 1993: a Bank Ombudsman with power to investigate and recommend remedial action, and an independent appeals commission. What was approved was narrower than either: a fact-finding body that could recommend investigation, with the Board retaining the gate-keeping role at every major decision point.

3.2 The Six Structural Constraints

The 2024 External Review Team documented the Panel's structural limitations with unusual directness. The following six constraints define the operational envelope within which the Panel has worked across more than three decades:

Constraint	What It Means	Comparator
Board approval required for investigation	Investigation can only proceed after Board votes to approve Panel recommendation. Average: 4.5 months from complaint to approval.	CAO: no Board approval required
No self-initiation rights	Panel cannot initiate a case. Affected communities must file, in writing, after first approaching Management.	CAO, AfDB IRM, EBRD: self-initiation permitted
No recommendation authority	After finding non-compliance, Panel cannot recommend specific remedial actions. Management designs the remedy.	CAO: recommends and monitors remedy
MAP self-monitoring	Management drafts MAP, Board approves, Management self-reports. Panel verifies only with separate Board approval.	CAO: monitors MAP implementation by default

Constraint	What It Means	Comparator
Policy unreformed for 27 years	From 1993 to the 2020 Toolkit Reforms, the Panel's policy was not substantively revised. CAO was established in 1999 and developed a more advanced framework in parallel.	Most MDB IAMs review policy every 5 years
Eligibility requirements	Requesters must allege harm, link it to Bank policies, show prior contact with Management, and the project must not be >95% disbursed.	CAO: simpler eligibility; lower threshold for access

Sources: 2024 External Review Team; Yale Lowenstein Clinic (2017); Task Force Draft Report (March 2026). CAO comparisons from CAO Policy (2021) and Task Force assessment.

3.3 The Management Response Loophole

Under paragraph 23 of the founding Resolution, the Panel must satisfy itself that Management has failed to demonstrate adequate compliance before recommending investigation. When Management responds to a complaint by admitting limited failures and promising corrective action, the Panel must assess whether those promised actions are adequate. If it finds them so, it can recommend not investigating.

This creates a structural incentive for Management to produce responses that close complaints before formal investigation, regardless of whether they actually address community concerns. The Yale Lowenstein Clinic documented exactly this dynamic in Uzbekistan:

WHEN A PLAUSIBLE LINK TO HARM STILL DID NOT TRIGGER INVESTIGATION

In 2013, the Panel found a “plausible link” between Uzbekistan’s Second Rural Enterprise Support Project and ongoing forced and child labour. Rather than recommending investigation, it deferred its decision for twelve months, based on Management’s promised corrective actions. A year later, with those commitments still only partly met, the Panel deferred again rather than investigate, citing some progress and renewed assurances of more to come. No investigation has been opened to date.

Source: Yale Lowenstein Clinic, Deferring Accountability (2017), citing the Panel’s Eligibility Report and Recommendation, Uzbekistan: Second Rural Enterprise Support Project (December 9, 2013).

3.4 The MAP Chain — Why Remedy Rarely Reaches Communities

The conceptual framework above traces the full chain from project approval to MAP closure; this section isolates the segment that matters most for remedy. Once the Panel finds non-compliance, the Management Action Plan process is where dependence on Management becomes decisive:

THE ACCOUNTABILITY CHAIN: WHERE IT BREAKS

A conceptual framework for the empirical findings in this paper

Stage	What happens	Break point / constraint
Project approved	World Bank Board approves financing; safeguard policies govern design and implementation.	—
Harm occurs	Project implementation produces adverse environmental, social, or economic impacts on communities.	—
Complaint filed	Affected communities file a formal Request for Inspection with the Panel.	ACCESS BARRIER <i>Awareness, CSO capacity, retaliation risk, eligibility bar, instrument scope gaps — most harm never reaches this stage.</i>
Panel registers	Panel determines the complaint is admissible and formally registers it.	PROCEDURAL FILTER <i>30% of complaints closed at registration. Footnote 7 delays, Early Solutions Pilot, management response loophole — additional attrition before formal entry.</i>
Board approves investigation	Board of Executive Directors votes on Panel recommendation to investigate.	POLITICAL GATE <i>Investigation requires Board approval. Board members represent borrowing governments. Political considerations can block or defer investigation.</i>
Panel investigates	Panel conducts field visits, interviews, and policy compliance assessment; publishes findings.	MANDATE SCOPE <i>Panel can narrow scope, excluding most serious harm allegations as outside its mandate. Finding suppressions possible pre-Board.</i>
Management drafts MAP	Management designs the remedy in response to non-compliance findings.	CONFLICT OF INTEREST <i>Author of non-compliance designs the remedy. Panel</i>

		<i>has no recommendation authority over MAP content.</i>
Board approves MAP	Board votes on Management Action Plan.	POLITICAL GATE <i>Board approves Management's own remedy plan. Communities cannot appeal MAP content.</i>
Management implements & self-reports	Management monitors and reports on MAP implementation to Board.	SELF-MONITORING <i>No independent verification as default. Progress reports authored by Management.</i>
Independent verification	Panel independently verifies implementation of Board-approved MAP.	EXCEPTION, NOT DEFAULT <i>Requires separate Board approval. Triggered infrequently.</i>
Remedy to communities	Communities receive documented, verified remedy for harm caused.	NOT SYSTEMATICALLY MEASURED <i>The Bank has no system for tracking whether community-level remedy is delivered.</i>

No shading (Grey): normal stage **Amber:** partial constraint / attrition point **Red:** structural break – political gate or self-monitoring

This framework connects the quantitative findings of Section 2 to the institutional analysis that follows. The access barrier (Stage 3) explains the Complaint Index and the near-zero performance–complaint correlation. The political gates (Stages 5, 7, 8) and the self-monitoring of remedy (Stage 9) explain the MAP non-enforcement pattern. The final stage's absence from Bank measurement systems is the gap visualised in the Accountability Funnel above. Sections 3 through 6 and the Annex document the institutional evidence for each constraint shown here.

The critical characteristic of this chain is that every gate is controlled by the Board, and every content decision is controlled by Management. The Panel produces findings; it does not control what happens with them. The community that filed the complaint participates in Management consultations on the MAP but cannot appeal the MAP's content or verify its implementation independently.

3.5 Comparison with Other MDB Accountability Mechanisms

The Inspection Panel is consistently assessed as operating below the standards of comparable MDB mechanisms on the dimensions most consequential for remedy delivery:

Mechanism	Self-initiate?	Board approves?	Recommends remedy?	Verifies MAP?
World Bank IPN	No (proposed)	Yes — Board vote required	No	Rarely (Board must approve)
IFC/MIGA CAO	Yes	No	Yes	Yes (monitors by default)
ADB CRP/OSPF	CRP: No; OSPF: limited	CRP: Board vote	CRP: No; OSPF: limited	Yes
AfDB IRM	Yes (limited)	No	Yes	Yes
EBRD IPAM	Yes (limited)	No	Yes	Yes
IDB MICI	Limited	No	Yes	Yes
AIIB PPM	No	No	No	No

Sources: Task Force Draft Report (March 2026), Annex on IAM benchmarking; CAO Policy (2021); ADB Accountability Mechanism Policy (2012); AfDB IRM Rules (2015); EBRD Project Complaint Mechanism Rules (2019). AIIB PPM is the newest and least developed mechanism.

The 2024 External Review Team was direct: "the CAO Policy is better aligned with IAM good practices." The Task Force on integration noted this as the rationale for harmonising toward CAO standards in the merged mechanism. The Panel spent 27 years operating below the standard its own institution's private-sector arm had developed and that every comparable MDB mechanism had adopted.

4. Scholarship, Civil Society, and the Structural Contradiction

On the Panel’s thirtieth anniversary the American University Washington College of Law published *Perspectives*, a collection of essays by the people who built, ran, defended, and criticised the mechanism. Read together, the defining feature of that collection is consensus: the Panel’s own members and the outside critics agree on where the mechanism fails. This section draws on those essays to locate this paper’s quantitative findings within the qualitative record, and to state plainly the structural contradiction the data express.

4.1 What the Panel’s Own Record Concedes

The most authoritative defence of the Panel is also, on inspection, the clearest admission of its limits. In the only prior data-driven assessment of the mechanism, eight serving Panel members and staff (Kunanayagam et al., *Perspectives* 11, 2023) report that requests have equalled roughly 1.5 percent of Bank-financed projects and investigations under 0.5 percent of approved projects — the same funnel documented in Section 2. Their list of “positive outcomes” is revealing: it consists of systemic learning, policy clarifications, and lessons-learned advisories. It does not include any systematic measure of whether affected communities obtained relief. The essay concedes the reason directly: the Panel “has no monitoring authority,” its verification function is highly restrictive and requires Board approval, and the same Board mandate that lends the Panel legitimacy also limits how far it can reach.

Dilek Barlas, the Panel’s former Executive Secretary (*Perspectives* 6, 2024), is equally candid on the decisive point: Management’s proposed remedial actions must be agreed with the borrower, and when borrowing governments refuse, the affected people get nothing. She names the cases — Albania Integrated Coastal Zone, where only two of nine families’ claims ultimately won compensation, and Cambodia Land Management, where the government rejected the remedy and the Bank suspended lending for five years. Serge Selwan, a senior Panel adviser writing in 2026 (*Perspectives* 58), adds that the corporate-level Grievance Redress Service now intercepts roughly 1,200 complaints a year before they reach the Panel, and warns that the 2026 integration’s options “structurally dilute accountability” by trading independence for efficiency.

4.2 What the Critical Literature Establishes

The independent scholarship converges on the same structure from the outside. Richard Bissell, a former Panel Chair (*Perspectives* 9, 2023), states the core finding of three decades: remedy has been the issue least addressed by IFI leadership because management and board members find it threatening. He reports that an early survey found only 10 of 28 registered complaints produced any project-level impact, and names the high-stakes cases that produced none — Yacyretá, Itaparica, Lesotho Highlands, Lake Victoria, and the Chad–Cameroon pipeline. He documents the loophole this paper terms the management-response escape hatch: an elaborate Management response addressed to the Board could persuade it to reduce or reject an investigation at the eligibility stage.

Susan Park (*Perspectives* 20, 2024) supplies the formal statement of the problem: the aim of the accountability mechanisms is to render an account of the Banks to their member states, not to the

people the projects harm. Accountability runs upward to shareholders, not downward to affected communities; staff are not sanctioned for working on harmful projects; and the cases “routinely demonstrate the same policies being violated.” Dustin Schäfer (Perspectives 19, 2023) identifies the mechanism: it is the Board, as the representative of shareholders, that prioritises low borrowing costs and thereby “structurally prevents proper due diligence, high quality at entry, and effective accountability.” Board decisions, he observes, are shaped less by the Panel’s findings than by geopolitics, and the Panel is least effective in the largest borrowers’ states. Natalie Bridgeman Fields, who founded Accountability Counsel (Perspectives 14, 2023), confirms the result from two decades of casework: community-level remedy has been rare, powerful borrowers simply withdraw or stonewall to remove the Bank’s leverage, and staff career advancement still depends on moving money out the door.

4.3 The Terminal Question: Did Affected Communities Obtain Relief?

Every metric in this paper — the Complaint Index, the performance–complaint correlation, the 30.9 percent S+ rate of complained-about projects — leads to a single question that the accountability literature, the Panel’s own people, and this author’s field experience answer the same way. The Accountability Funnel ends not at “investigation” but at “remedy,” and remedy is the stage no one measures. Bissell’s 10-of-28, Barlas’s 2-of-9 in Albania, and the Panel’s own omission of any community-relief indicator are not isolated data points; they are the same finding observed from three vantage points. The Panel reliably establishes that harm occurred and that the Bank was out of compliance. What follows for the people harmed is, in the majority of cases, not relief but a Management Action Plan, drafted by the Management whose project failed, agreed with the government that caused the harm, and closed when Management reports completion.

The Ethiopia Protecting Basic Services case, examined in the Annex, is the paradigmatic instance. The Anuak communities of Gambella alleged that Bank-financed block grants were underwriting a villagisation programme that forced them from ancestral land; the facilitator the Panel itself had engaged, Pastor Omot Agwa, was detained without charge and then charged with terrorism, and was not freed for nearly three years (Barlas 2024). The affected communities did not obtain relief. The harm was being done by the borrower government — and that is the rule, not the exception: in the large majority of Panel cases the proximate agent of harm is the implementing government, not the Bank.

4.4 The Internal Contradiction: The Borrower Is the Client

This is where the qualitative literature and the quantitative findings meet. The borrower government occupies three roles at once. It is frequently the proximate cause of the harm a complaint describes. It is the Bank’s client, whose continued borrowing sustains the lending pipeline on which the institution’s budget, influence, and staff incentives depend. And it is a shareholder, represented through its constituency on the Board of Executive Directors that must approve the investigation, approve the Management Action Plan, approve any independent verification, and close the case. The entity that caused the harm is also the customer the Bank cannot afford to lose and a member of the body that adjudicates the complaint against it.

No accountability chain can run cleanly through that configuration. The institution’s structural imperative is to lend — “lending pressure crowds out learning,” in the IEG’s own phrase quoted by Schäfer — and remedy that antagonises the client threatens the pipeline. Park’s observation that accountability is owed upward to shareholders rather than downward to the harmed is not a critique of how the mechanism is run; it is a description of what the mechanism is. The access barrier this paper measures in Sub-Saharan Africa, and the remedy gap the literature documents, are the two observable outputs of the same contradiction: the people best positioned to file are not those most harmed, and the people most harmed are those least able to compel a client government — and a Board that represents it — to deliver relief.

The Panel’s defenders offer one genuine counter, and it deserves to be stated fairly. Barlas argues that the Panel’s effectiveness has always rested on its independence and its direct reporting line to the Board — the Board as the body that shields the Panel from Management interference. That is correct, and it is precisely why the contradiction bites: the same Board that protects the Panel from Management is the body whose members represent the borrowers the Panel investigates. The Board is simultaneously the Panel’s shield and the borrower’s seat. An accountability mechanism whose final authority is the assembled representatives of the governments it scrutinises is independent up to the point where independence would have consequences.

4.5 The Power Asymmetry the Mechanism Was Not Built to Reach

The Ethiopia Protecting Basic Services case makes the asymmetry concrete, and it is sharper than the familiar image of an uneven contest. The complainants were Anuak villagers in Gambella — an indigenous minority among the most marginalised people in one of the Bank’s poorest borrowers. The harm they described, forced relocation under a government villagisation programme, was carried out by the state’s own security forces. The financing vehicle was Protecting Basic Services, a multi-donor flagship that channelled roughly \$17.5 billion into Ethiopian service delivery across its phases — one of the largest operations the Bank has run in Africa, and one whose continuation the Bank and its partners had strong institutional reasons to protect. The complainants were opposed not by one adversary but by an aligned set of them: their own government, the security apparatus implementing the programme, and an institution with a reputational and programmatic stake in the very operation under complaint.

Daniel Bradlow (Perspectives 33, 2025) identifies the design feature that makes the contest structurally unequal rather than merely lopsided. The Panel investigates only the compliance of Bank staff; it does not investigate the conduct of the borrower, and because it does not investigate member states it does not touch their sovereignty — by design. The consequence in Gambella was direct: the entity actually conducting the villagisation was never the subject of the investigation. The mechanism could ask only whether Bank staff had supervised their own safeguards adequately. The agent of harm sat outside the mechanism’s reach from the start — and inside the Board that would rule on the findings.

This is why the usual metaphor understates the case. In the David and Goliath story the weaker party at least faces one opponent on open ground with a real, if remote, chance. The Gambella complainants

had no court — the Bank’s immunity forecloses litigation — no recourse to their own government, and a retaliation risk that materialised: the facilitator the Panel itself engaged was imprisoned for nearly three years. Even advocates of the gentler, dispute-resolution route concede the imbalance. Gina Barbieri (Perspectives 2, 2023), an experienced mediator, argues that the standard mediation model is insufficient for development disputes precisely because the power gap dwarfs that of ordinary conflicts, and that without deliberate scaffolding — capacity building, joint fact-finding, explicit levelling measures — collaborative processes reproduce the asymmetry rather than correct it. The 2026 merger steers complainants toward that route without committing to the scaffolding she identifies as the precondition for a fair one.

A finding in the community’s favour meets one further structural disincentive. Bradlow notes that the cost of a remedial action plan falls either on the borrower or, when the Bank funds it, on the same pool of resources available to all borrowing members — so remedy is, in budgetary terms, a charge against lending. An institution whose imperative is to lend has a built-in reason not to make remedy routine. The asymmetry is therefore not only of power but of incentive: nearly every party with the capacity to deliver relief has a reason to withhold it.

4.6 When the Bank Did Act: The Cambodia Exception

A reader could object that the record is not uniformly one of inaction. On a small number of occasions the Bank did impose a real consequence on a borrower in the wake of a Panel case, and an honest assessment has to ask how often, and under what conditions. The answer sharpens the argument rather than softening it. Across thirty-three years and forty-two investigations, the instances in which a Panel finding was followed by a tangible borrower-facing consequence can be counted on one hand.

Case and Panel finding	Consequence to the borrower	Source of the leverage
Cambodia — Land Management (2011–16)	All new lending suspended for ~5 years until a remedy was reached	Management and Board, under sustained US shareholder and statutory pressure
Tanzania — REGROW / Ruaha (2024)	Disbursements suspended mid-investigation; project cancelled at the government’s request	Management suspension over safeguard non-compliance, amid civil-society, UN and media pressure
Uganda — Transport Sector Development, and related UNRA projects (2015–17)	Project cancelled; civil works on two related agency projects suspended; survivor support delivered through the Emergency Child Protection Response and SCOPE programmes	Management action during the Panel process, ahead of the Panel’s report; sustained community–NGO advocacy and US Board pressure
India — Mumbai Urban Transport (2006)	Partial, roughly four-month suspension of the roads and resettlement components	Management suspension during the Panel investigation; lifted on the state’s compliance

Case and Panel finding	Consequence to the borrower	Source of the leverage
China — Western Poverty Reduction, Qinghai (2000)	No Bank penalty; the Board withheld approval on Management’s terms and the government withdrew to self-finance	Board and shareholders acting on the Panel’s findings; borrower exit
Chad — Chad–Cameroon Pipeline (2006)	Disbursements suspended and the London escrow account frozen	Loan-covenant enforcement over a breach of the revenue-management law — not a Panel remedy

Sources: Panel case register; World Bank Board and Management statements; US Treasury, “Positions on Inspection Mechanism Cases”; Independent Evaluation Group project assessments. The final two rows (shaded) are borderline: in the China case the consequence was the borrower’s own withdrawal after the Board withheld approval, and in the Chad case the suspension enforced a loan covenant on the revenue-management law rather than a Panel finding. Counting generously, genuine borrower-facing consequences following a Panel case number about four across thirty-three years — the most recent, Tanzania, only in 2024 — against forty-two investigations and a portfolio of thousands of projects.

The first thing these cases share is that in none of them did the Panel act against the borrower. It cannot: as Section 4.5 sets out, the Panel investigates the compliance of Bank staff, not the conduct of member states, and because it does not investigate its member states it does not implicate their sovereignty. Every consequence in the table above was imposed through a different lever — a Management decision to suspend or cancel, a Board vote to withhold, the diplomatic and budgetary weight of a major shareholder, or the enforcement of a loan covenant. The detail bears this out: in Uganda, Management cancelled the project months before the Panel completed its investigation report; in China, it was the Board, not the Panel, that withheld approval, after which the government withdrew; in Chad the suspension enforced the revenue-management law rather than any Panel finding; and in Tanzania the cancellation came at the government’s own request. The Panel’s own authority appears nowhere in the chain of consequence. The enforcement cases are not evidence that the accountability mechanism reaches the borrower; they are evidence that, on the rare occasions a borrower faces a consequence, the force comes from outside the mechanism.

Cambodia is the case most often cited as proof that the system can bite, and it repays attention because it shows how much had to align. Four conditions held at once. First, Bank Management concurred with the Panel’s findings rather than contesting them — itself uncommon, and the opposite of the misrepresentation documented in the Ethiopia case. Second, the government openly refused the agreed remedial actions for the families evicted around Boeung Kak Lake in Phnom Penh, removing the diplomatic cover that ordinarily lets the Bank record a case as a work in progress. Third, the Bank’s largest shareholder put sustained, public, and ultimately statutory weight behind enforcement. Fourth, Cambodia was a small IDA borrower with little countervailing leverage; the institutional cost of penalising it was low.

THE SHAREHOLDER'S POSITION, ON THE RECORD

“The United States is deeply concerned about the denial of the Boeung Kak Lake residents’ land claims and their forced eviction ... [and] concurs with the findings of serious violations of Bank policies on project appraisal, supervision and resettlement. ... This cannot be business as usual.”

U.S. Position on the Inspection Panel Investigation, Cambodia Land Management and Administration Project, 9 March 2011 — a document still published on the US Treasury website fifteen years later.

The shareholder went further than a statement. Drawing on Senate Report 113-81, the Consolidated Appropriations Act conditioned the Bank’s re-engagement with Cambodia on defined redress for the Boeung Kak families — legal security of tenure for those who had been excluded from titling, demarcation of the developer’s concession, and livelihood programmes for those forcibly evicted — and required the US Executive Director to report to Congress on the steps taken to postpone re-engagement until those conditions were met. Several observers read the contrast between this treatment and the Bank’s gentler handling of investigations touching its largest borrowers as evidence that enforcement tracks a borrower’s strategic weight more closely than the gravity of the harm — the pattern Schäfer (Section 4.2) describes when he observes that the Panel is least effective in the largest borrowers’ states. The evidence here is consistent with that reading without establishing it; what the Cambodia file shows directly is that the consequence followed not from the Panel’s finding but from a shareholder willing to spend political capital where the cost of doing so was low.

And even then — in the one case where Management agreed, the shareholder pushed, and the Bank actually withheld lending — the remedy did not arrive. The freeze ran for roughly five years and was lifted in May 2016, when the Board approved \$130 million in new lending, with the Boeung Kak families still largely uncompensated; a later government sub-decree granted title to a fraction of them, on terms the authorities described as unrelated to the Panel case (Annex, Case 4). The most recent case repeats the shape: in Tanzania the Bank suspended and then cancelled the REGROW project in 2024 and approved a remedial action plan with a trust fund in 2025, yet the affected communities remain uncompensated and, in 2025, faced the very evictions the safeguards were meant to prevent.

Uganda is the clearest exception to the exception, and the fairest reading of this paper’s evidence requires stating it plainly. After the Panel’s investigation documented severe sexual exploitation and abuse of children connected to the road project, the Bank did not stop at cancelling financing. It funded an Emergency Child Protection Response programme and a follow-on initiative, SCOPE (2017–19), that delivered medical care, psychosocial counselling, legal redress, financial support, and school or vocational reintegration directly to affected children, young women, and their families, alongside payment of delayed resettlement compensation. Independent monitors who tracked the case attribute the response to a combination the Cambodia case also shows in different proportions: sustained community–NGO advocacy, international media attention, a strong Panel finding, and Board-level pressure from the United States that did not let up once the cancellation was announced. Measured against the Accountability Funnel’s own final stage — remedy confirmed delivered — Uganda is the strongest case in this paper’s full register of that stage actually being reached.

Cambodia and Tanzania end where the Accountability Funnel ends: non-compliance established, a consequence imposed on the borrower from outside the mechanism, and relief for the people harmed left undelivered. Uganda did not, and the conditions behind that difference — organised local advocacy, sustained international attention, and shareholder pressure that persisted after the cancellation rather than stopping there — are precisely the conditions absent from the great majority of the Panel’s forty-two investigations. The exception does not break the pattern this paper documents across the full caseload; it identifies what the pattern requires to be broken.

4.7 The Board at Every Gate

The observable record is consistent with this account. The Board has exercised its approval authority to restrict rather than expand the Panel’s scope: it declined to approve the Panel’s investigation recommendation in the Lebanon Greater Beirut Water case (Kunanayagam et al. 2023); the Medupi/Eskom project was approved two days after a community complaint was received; the Amaravati investigation recommendation was twice deferred without precedent; the Ethiopia PBS Management Action Plan was approved over Human Rights Watch’s documented objection that it did not address the Panel’s findings. In each case the decision preserved the operational relationship at the expense of the accountability function.

The borrower-side mirror image is the repeat-complaint pattern. Uganda generates more Panel cases than any other African country not because its projects are the worst — its S+ rate of 31.5 percent across 108 evaluated projects is poor but not exceptional — but because it has the civil-society infrastructure to support filing. Several large, below-median portfolios (Mali at 29.2 percent, Madagascar at 33.0 percent, Central African Republic at 26.7 percent) generated no complaints at all. Complaint volume tracks the capacity to complain, not the incidence of harm — the access barrier of Section 2 and the borrower-client contradiction of this section, observed in the same data.

5. Sub-Saharan Africa: The Access Barrier in Detail

5.1 The Performance Record

The IEG master database allows the Africa performance record to be characterised precisely. Across 3,144 evaluated SSA projects with \$198 billion in commitments, the S+ rate is 36.8% — 9.5 percentage points below the global rate of 46.3% (4,839 S+ of 10,453 total rated). The decade breakdown is sharper:

Decade	SSA Projects	SSA S+%	Global S+%	SSA Commitments (\$bn)
Pre-2000	198	39.4%	50.5%	\$13.8bn
2000–2009	661	33.1%	45.4%	\$39.0bn
2010–2019	913	16.8%	27.1%	\$66.8bn
2020+	602	42.0%	44.1%	\$78.0bn

Source: IEG master database, March 2026. S+ = Satisfactory or Highly Satisfactory. 2010–2019 row highlighted: SSA S+ fell to 16.8% as IDA commitments reached \$66.8bn — the decade of maximum portfolio scale and minimum delivery quality. 2020+ figure is partial and reflects a mix of completed and recently evaluated projects.

The 2010–2019 decade shows the sharpest divergence. The Bank's Africa portfolio expanded dramatically — \$66.8 billion in commitments, more than 70% above the preceding decade's level — while the S+ rate fell to 16.8%. Over the same period, SSA Panel investigations were in single digits. Investigation frequency, portfolio scale, and project performance moved in directions that compounded each other's adverse effects.

5.2 Why Access Within the Region Is Uneven

Sub-Saharan Africa as a region files more complaints than its lending share predicts, not fewer — its projects fail more often. The barrier is within the region: among countries with comparably poor records, some generate many complaints and others none. Five cumulative, mutually reinforcing factors account for that unevenness:

- **Mechanism awareness is near zero in most project contexts.** The Panel acknowledged in the 2024 External Review that information is buried in Project Appraisal Documents. In the typical SSA rural project context — community in a remote area, Bank country office in the capital, project language not the community's language — awareness of the Panel is effectively absent without active CSO mediation. No borrower disclosure requirement exists at project level.
- **Civil society infrastructure is the principal predictor of complaints.** Panel utilisation correlates most strongly with the presence of experienced CSOs that can support

communities in navigating the complaint process. Uganda's ten cases reflect strong CSO capacity and transnational-advocacy links; Mali's zero cases across 96 evaluated projects at 29.2% S+ reflect their absence. Complaint distribution tracks civil-society strength — measurable through the V-Dem Civil Society Participation and Core Civil Society indices and the CIVICUS Monitor — far more closely than it tracks project performance.

- **Retaliation risk is structurally higher in FCV contexts.** HRW documented that 18 of 34 complainants reported threats or reprisals. In African FCV contexts — where governments have security capacity and limited rule-of-law constraints — the rational calculation for potential complainants is that filing a complaint against a government-backed project is more likely to produce personal risk than institutional remedy. The Ethiopia PBS case is the documented extreme of this pattern, not an outlier.
- **Prior experience reduces future complaint propensity.** In communities where a previous complaint produced no observable remedy — or where complainants faced retaliation — the rational response is not to file again. The Badia East mediation outcome that Amnesty International found seriously flawed becomes the observed precedent that shapes expectations for future potential complainants across Nigeria's urban programme.
- **The complaint eligibility bar is higher for public sector projects.** The CAO's lower eligibility standard for IFC/MIGA-related complaints — which the Task Force has recommended extending to the merged mechanism — was not available for the IBRD/IDA projects that dominate the SSA portfolio. This means the communities most dependent on public sector lending faced the strictest access requirements.

5.3 The Zero-Complaint Countries

The following SSA countries generated zero Panel complaints across their entire evaluated portfolio history, despite poor-to-middling performance and, in several cases, very large portfolios. Note the worst performer in the region, Congo (Republic) at 11.5 percent S+, alongside Mali, Madagascar, and Senegal — each with a hundred-odd evaluated projects and not one complaint:

Country	Evaluated Projects	S+ Rate	Commitments	IPN Complaints
Congo (Republic)	26	11.5%	\$0.9bn	0
South Sudan	25	24.0%	\$0.8bn	0
Central African Republic	45	26.7%	\$0.9bn	0
Mali	96	29.2%	\$4.0bn	0
Sudan	41	31.7%	\$2.3bn	0
Burkina Faso	103	33.0%	\$5.3bn	0
Madagascar	112	33.0%	\$5.2bn	0
Guinea	72	33.3%	\$2.2bn	0

Country	Evaluated Projects	S+ Rate	Commitments	IPN Complaints
Niger	88	37.5%	\$4.2bn	0
Senegal	128	41.4%	\$6.6bn	0

Source: IEG master database (March 2026) for performance data; Panel case register (June 2026) for complaint counts. DRC recently received two complaints; highlighted as partial exception. Red rows: countries with S+ below 30% and zero complaints across entire portfolio history.

The DRC is a case that illustrates the pattern in concentrated form: 72 evaluated projects, \$6.8 billion in commitments, 25.0% S+, and until recently zero Panel cases. The DRC portfolio analysis in this series documented a \$150 million project size threshold above which no project achieved S+, and a commitment-weighted S+ rate of approximately 6%. No community in the DRC filed a Panel complaint against that record until very recently. This is not because those communities were well-served.

6. The June 2026 Merger: What Changes and What Does Not

6.1 What the Merger Addresses

The Board's June 2026 decision addresses several of the structural constraints documented in this paper, though policy details remain to be developed. The Task Force recommended:

- **Eligibility harmonisation toward CAO standards.** If fully implemented, this lowers the access bar for public-sector project complainants. It is the most consequential potential gain from the merger for African communities.
- **Self-initiation rights for the compliance function.** If adopted in final policy, this begins to address the zero-complaint pattern in countries where communities cannot or will not file. It is the reform most likely to change the Complaint Index for SSA.
- **Single entry portal.** Reduces navigational barriers for communities uncertain whether their project is IFC, IBRD, or MIGA.
- **Advisory function across the full portfolio.** Cross-sector learning output from both public and private sector operations.

6.2 What the Merger Does Not Address

None of the four recurring patterns documented in the Annex are structurally resolved:

- **Volume.** Self-initiation is not yet confirmed policy. Even if adopted, it does not address awareness barriers, CSO infrastructure gaps, or retaliation risk in FCV contexts.
- **Management response loophole.** The founding Resolution logic allowing pre-investigation closure through Management corrective promises is embedded in the legal framework and was not proposed for removal.
- **Delay procedures.** Footnote 7 and the Early Solutions Pilot were absorbed into the 2025 AM Resolution architecture without elimination.
- **Finding suppression risk.** The Amaravati episode involved a report removed before Board consideration. The merged mechanism under a VP/DG who reports to the Board concentrates – rather than distributes – the institutional point at which a finding could be managed before it reaches the Board.
- **MAP chain.** Management still drafts the MAP. The Board still approves it. Management still self-monitors. The Board still closes it. The VP/DG still needs Board authorisation to verify implementation. The chain is structurally identical.
- **Retaliation.** No enforcement authority. The merged mechanism cannot protect complainants from government security forces in FCV environments.

The merger also creates transition risks. The 2024 External Review Team noted that the CAO Policy is better aligned with good practice, and that any consolidation must be non-regressive. If policy harmonisation moves the IPN's eligibility bar toward CAO standards, that is a gain. If it simultaneously adjusts the CAO's self-initiation or monitoring standards upward toward the IPN's more restrictive approach, the net effect is regression presented as harmonisation. The policy development phase, not the structural decision, will determine whether the merger produces progress or consolidates current weaknesses.

The most pointed warning comes from inside the mechanism. Serge Selwan, a senior Panel adviser with twenty-five years in the field, argues that the integration task force — staffed by a retired senior manager and a dispute-resolution specialist rather than accountability experts — produced options that “structurally dilute accountability” in the name of efficiency, and repeats what he calls the “original sin” of conflating compliance (owed to the Board) with dispute resolution (owed to operations). His caution is the one this paper would underline: deconstructing accountability is fast and easy; rebuilding it is neither (Selwan, Perspectives 58, 2026).

It is worth measuring the merger against the reform the accountability scholarship actually proposed. Daniel Bradlow (Perspectives 33, 2025) argued for a single mechanism spanning all the MDBs — but with two features the 2026 merger omits. First, its findings would bind: the relevant Board would have to accept them or publish its reasons for rejection, narrowing the discretion documented in Section 4.7. Second, it would be paired with an independent Assistance Fund, financed by state and non-state contributors, able to deliver remedy to harmed communities without drawing on the lending pool — directly addressing the cost-of-remedy disincentive. The World Bank Group’s merger consolidates structure while adopting neither feature: findings remain advisory to the Board, and remedy remains a charge against lending agreed with the borrower. The consolidation that occurred is not the consolidation the reform literature called for.

6.3 The Task Force Solves the Wrong Problem

Three documents now sit alongside each other in the integration debate, and they are not making the same case. The Task Force’s draft report, published for consultation in March 2026, opens by scoring all three WBG accountability mechanisms against seven good-practice criteria — one of them, to its credit, is contribution to remedial action. But that scoring runs to a few qualitative sentences per mechanism, and the report’s remaining forty-odd pages turn to a different question: not whether the mechanism delivers remedy, but how it should be structured, governed, and harmonised. Bradlow and Hunter (Perspectives 33 and 41, 2025) ask a third question again: how that restructuring can proceed without weakening what already exists — preserve independence, preserve direct Board access, commit to non-regression. This paper asks a fourth: does the existing mechanism, however organised and however briefly assessed, actually deliver accountability at the scale its caseload requires? The comparison is sharper as a table than as prose.

2026 Task Force	Bradlow & Hunter	This Paper
Effectiveness scorecard (7 criteria)	Independence	Performance
Remedial action (1 of 7, brief)	Institutional safeguards	Empirical outcomes
Structural governance options	Principles	Results
Policy harmonisation	Legal protections	Complaint data
Web portal and transition plan	Non-regression	The accountability gap
Cost-effectiveness benchmarking	Future risks	Thirty-three years of evidence

The Task Force’s report comes closer than its predecessors to asking the right question, and that should be credited plainly: one of its seven effectiveness criteria is the IAMS’ contribution to remedial action, citing the same scholarship this paper also draws on (Kunanayagam et al., Perspectives 11, 2023). But the treatment is brief — a few sentences per mechanism, drawn from interviews and a small sample of cases — concluding only that numerous IPN cases with findings of non-compliance have led to some management action, without quantifying how numerous against what base, or whether any verified outcome followed. The report does not ask why investigations have averaged little more than one per year, why Sub-Saharan Africa’s worst-performing countries generate the fewest complaints, or why remedy delivery is not tracked at the level this paper measures it. Civil-society observers reached a similar conclusion from outside the process: increasing “remedy to affected people was notably not one of the stated goals” of the Board’s own terms of reference for the merger (Pred, quoted in Devex, April 2026). Bradlow and Hunter’s caution is well-founded on its own terms — their implicit premise is that the existing system works reasonably well and chiefly needs protecting from a clumsy merger. This paper’s evidence does not support that premise either way it might be applied: an under-performing mechanism deserves neither careless reorganisation nor careful preservation. It deserves diagnosis first.

The Task Force’s own language makes the gap easy to see, because its aspirations can be set directly against this paper’s measurements.

Task Force aspiration	What this paper measures
“Enhance awareness, accessibility, transparency, and predictability”	Mali, Madagascar, Burkina Faso, Central African Republic, and the Republic of Congo — five of the weakest-performing SSA portfolios — generated zero complaints across their entire histories; country-level complaint volume is statistically unrelated to project performance ($r = -0.230$, $p = 0.101$).
No regression of policies and mandates	Verified, community-level remedy is not systematically measured by the Bank at any stage of the chain (Section 2.1).
Sustaining and strengthening operational effectiveness	42 investigations from 10,453 evaluated projects across thirty-three years — a rate that has fallen to 0.4 per year in 2020–2026 as the portfolio expanded (Section 2.3).

None of this is offered as a case against organisational reform; consolidating overlapping accountability functions may well be sound institutional design. The point is sequence. The Task Force changed the prescription without first establishing the diagnosis. A reform process that redesigns an institution's architecture before establishing whether the existing architecture delivers outcomes risks treating fragmentation as the disease, when the evidence assembled in this paper points to a different one.

6.4 The Integration Debate Missed the Central Question

The consultation surrounding the 2026 merger focused almost entirely on institutional design. The Task Force discussed structural options for governing the merged mechanism, policy harmonisation across eligibility and remedy-verification rules, and a shared intake portal – touching remedy delivery only as one brief criterion in a seven-part scorecard. Civil-society responses concentrated on preserving institutional independence – the position Bradlow and Hunter defend in Section 6.3 above. Both sides shared an assumption neither side fully examined: that the existing accountability model is fundamentally sound and chiefly needs better organisation or stronger protection. This paper's evidence is consistent with a different reading. Thirty-three years of data show a mechanism that investigates a small fraction of the complaints it could pursue; that is used least where project performance is weakest; that leaves remedy largely unverified; and that, at the project level beneath its own caseload, does not measure whether the grievance systems it depends on are functioning at all (Annex, Case 5). Reorganising the institution's reporting lines does not, on its own, address any of these four deficiencies.

7. Conclusions

What this paper measured. Using the IEG master database (10,453 evaluated projects, \$911 billion in commitments, March 2026) and the Inspection Panel's case register (187 complaints, 131 registered, 42 investigated, 1994–2026), we constructed three original metrics: the Complaint Index, the performance–complaint correlation, and the Accountability Funnel. Together these operationalise a question that has not previously been tested with systematic data: does the pattern of complaint filing in Sub-Saharan Africa correspond to the pattern of project performance failure?

What this paper found. It does not. The Complaint Index for Sub-Saharan Africa is 1.37 — the highest of any region: Africa files more complaints than its portfolio weight predicts, because its projects fail more often. But within the region the Pearson correlation between country-level S+ performance and complaint rates across 52 SSA countries is -0.230 ($p = 0.101$, $R^2 = 0.053$), statistically indistinguishable from zero. The countries with the weakest performance records are not the ones generating the most complaints — several large, poorly-performing portfolios (Mali, Madagascar, Central African Republic) generated none at all, while complaint volume tracks civil-society capacity. The Accountability Funnel shows that verified, community-level remedy — the mechanism's ultimate purpose — is not systematically measured by the Bank. The investigation rate has declined to 0.4 per year in 2020–2026 as the portfolio has expanded.

Why it matters. Accountability mechanisms are not ends in themselves. They are instruments for improving project quality, deterring non-compliance, and ensuring remedy reaches communities harmed by financed operations. A mechanism that is systematically under-utilised in the regions where its portfolio performs worst, and that does not track whether remedy is delivered, cannot perform those functions effectively. The June 2026 merger of the Panel into a single IAM reorganises the institutional architecture. The empirical patterns documented here — structural access barriers, political gate-keeping, MAP self-monitoring — are functions of institutional incentives that the merger does not change. Beneath those incentives lies a contradiction the merger cannot reach: in the majority of cases the proximate cause of harm is the borrower government, which is simultaneously the Bank's client, whose continued borrowing sustains the lending pipeline, and a shareholder represented on the Board that must approve every remedy. An accountability chain whose final authority represents the governments it scrutinises can establish that harm occurred; it cannot reliably compel relief.

What future accountability mechanisms should measure. Three metrics would, if routinely published, transform the quality of accountability assessment at the World Bank and across the MDB system: a Complaint Index disaggregated by region and instrument type, tracking complaint frequency relative to lending exposure; a remedy tracking system recording whether MAP commitments are verified at the community level; and a complaint-to-outcome ratio reporting what share of complaints produce independently confirmed remedial action. None of these requires new data collection systems. They require institutional commitment to measuring what the accountability mechanism is actually delivering — not only that it exists.

Accountability systems should ultimately be judged not by the number of investigations they undertake, but by whether affected communities obtain independently verified remedy. On that measure, the evidence assembled here suggests the World Bank Inspection Panel has work still to do – and so does the institution that created it.

8. Recommendations

The findings above point toward specific, evidence-grounded actions rather than a general call for stronger oversight. Each recommendation below traces to a documented break in the Accountability Chain (Section 4.3) or the Accountability Funnel (Section 2.1), and none requires new authority from shareholders: each is a reporting and monitoring practice available within the Bank's existing mandate.

- **1. Verify and report remedy delivery, not only MAP approval.** The Accountability Funnel's final two stages — independent MAP verification and confirmed remedy delivery — are the only stages the Bank does not systematically measure. Routine, published verification of whether a Management Action Plan was actually implemented, and whether the harm it addressed was actually relieved, would close the funnel's least accountable link.
- **2. Replace binary GRM compliance reporting with GRM performance reporting.** As the SEEFOR case shows (Annex, Case 5), an ICR can confirm a grievance mechanism was “established” while reporting no information on how it performed. Implementation Completion Reports should be required to state, for every project-level GRM, the number of complaints received, resolved, and outstanding; median resolution time; and the number escalated to the corporate Grievance Redress Service or the accountability mechanism.
- **3. Build a structured escalation pathway from project-level GRMs to the accountability mechanism.** No project-level GRM in the Bank's standard reporting publishes performance data in a form comparable across projects. A standing data feed from project GRMs to the Integrated Accountability Mechanism would let chronically dysfunctional grievance systems surface as a portfolio-level signal, rather than only through the rare individual complaint that survives years later to reach the Panel.
- **4. Target outreach and civil-society capacity-building where the access barrier is largest.** Mali, Madagascar, Burkina Faso, the Central African Republic, and the Republic of Congo combine weak project performance with zero recorded complaints (Section 2.5). The access barrier this paper documents is concentrated in identifiable countries; outreach resources should be too.
- **5. Adopt performance metrics, not architecture metrics, as the test of the merger's success.** The first five years of the Integrated Accountability Mechanism should be judged against utilisation rate, investigation rate, and remedy-verification rate — not against whether the new organisational chart was completed on schedule.
- **6. Address the borrower-as-client contradiction at the Board level.** Where a Board member's own constituency is the subject of a Panel finding, recusal or independent-majority review at the investigation-authorisation and MAP-approval gates (Section 4.3) would reduce the structural conflict documented in Section 4.4, without requiring any change to the Panel's mandate.
- **7. Extend project-level performance benchmarking as a standing practice.** The method this paper applies in Section 2.6 — comparing the Satisfactory-or-better rate of complained-about projects against the sector benchmark — is portable. Applied consistently across the IPN, the CAO, and other MDB accountability mechanisms, it would let institutions

and observers track, year over year, whether complaints are catching the projects that most need catching.

None of these recommendations requires new authority from shareholders. What has been missing is not the power to measure, but the institutional decision to do so.

Methodology

IEG database. The IEG master database (March 2026) covers 10,542 records. After filtering to rated projects (excluding "Not Rated"), 10,453 projects with combined \$911 billion in commitments form the analytical base. S+ is defined as Satisfactory or Highly Satisfactory – the stricter threshold used consistently across the mdbreform.com analytical series. SSA is defined as Western and Central Africa plus Eastern and Southern Africa plus the residual "Africa" category in the database. Regional commitment shares are computed from this database.

Complaint Index. Complaint Share / Lending Share, where Complaint Share = regional complaints ÷ 187 total and Lending Share = regional IEG-database commitments ÷ total IEG-database commitments. Complaint shares are the exact regional distribution of all 187 requests, each mapped to the country of the project it concerns.

Performance-complaint correlation. Country S+ rates from IEG database for SSA countries with ≥8 evaluated projects. IPN case counts derived from the exact 187-case register, each request mapped to its project's country. Median split at 35.8% S+. The correlation analysis is descriptive and does not control for project size, sector, or country characteristics; it is intended as exploratory evidence, not a causal claim.

Panel statistics. Total cases (187), registrations (131), investigations (42), countries (56): Panel data page, inspectionpanel.org/panel-cases/data. Investigation-per-period breakdowns are approximate estimates from Panel annual reports and case register; they are indicative rather than precisely sourced for each period.

Case documentation. Ethiopia/PBS: Inclusive Development International (2012, 2015); HRW (2012, 2015); Oakland Institute (2013–2018); IPN Eligibility Report (February 2013) and Investigation Report, Ethiopia PBS III (February 2015); ICIJ (2015). Badia East: ICIJ (2015); Amnesty International (2014); Yale Lowenstein Clinic (2017). Albania: The Inspection Panel at 15 Years (World Bank, 2009); Yale Lowenstein Clinic, Deferring Accountability (2017). Amaravati: World Bank Corporate Secretariat press release (October 8, 2017); Panel case page; Centre for Financial Accountability (India, 2017); Panel case deferral announcements (December 2017, July 2018); Inspection Panel, Report and Recommendation on AIUDP (P507508, April 2025) and case-registration notice (December 2024); India CAG report on the Land Pooling Scheme; The Federal, "TISS audit of Amaravati land pooling scheme raises credibility concerns" (2026). Comparative IAM data: Task Force on Integration of WBG Independent Accountability Mechanisms, Draft Report for Public Consultation (March 25, 2026); Edwards, S., "World Bank weighs consolidating its three accountability mechanisms," Devex (April 14, 2026).

Scale metrics — denominator choice. The paper uses evaluated cumulative commitments (\$911bn, 10,453 projects) as the denominator for investigations-per-scale figures. This is a conservative choice: the active lending portfolio is larger (approximately \$220bn in annual new

commitments), and using it would produce a higher investigation rate, making the denominator smaller and the rate look worse. The evaluated portfolio is used because it is precisely verifiable from the IEG database by region. The methodology note in Section 2.2 states the evaluated-only basis explicitly. Reviewers should note that "investigations per \$100bn" using the active portfolio would produce approximately 19 per \$100bn — still a low rate but a different one. The paper's chosen figure (4.6) is conservative and uses the most precisely verifiable base.

Limitations. This analysis relies on complaint data from the Inspection Panel's case register rather than independently observed measures of project harm. Complaint frequency is an imperfect proxy for harm incidence because filing a complaint depends on awareness of the mechanism, civil society capacity to support requesters, willingness to file given retaliation risks, and eligibility under the Panel's admissibility criteria. A country with few complaints may have few complaints because its affected communities have low access to the mechanism, not because its projects cause less harm. The Complaint Index and the performance–complaint correlation should accordingly be interpreted as evidence regarding the utilisation of the accountability mechanism, rather than as direct measures of the underlying incidence of project harm. The near-zero correlation ($R^2 = 0.053$) is consistent with the access barrier hypothesis but does not rule out other explanations, including measurement error in complaint counts or differential project risk profiles across countries.

Statistical tests. Pearson r and Spearman rank correlation computed across 52 SSA countries with ≥ 8 evaluated projects, using complaints per 100 projects. Because most countries have 0–2 IPN cases, the distribution is heavily zero-inflated; Spearman $\rho = -0.224$ ($p = 0.110$) accompanies Pearson $r = -0.230$ ($p = 0.101$, $R^2 = 0.053$). Project performance explains approximately 5% of variation in complaint rates. The near-zero, statistically insignificant result is itself the substantive finding: complaint propensity is driven by factors other than project performance, consistent with the access barrier hypothesis.

Causal language. Phrases such as "recurring institutional patterns," "structurally consistent with," and "evidence most consistent with" reflect the analytical standard appropriate to evidence drawn from public documents, case records, and secondary sources. Attributions of intent are not made.

Selected References

Perspectives essays (American University Washington College of Law, Independent Accountability Mechanism Project).

Barbieri, Gina. "The Promise of Collaborative Problem Solving in Enhancing IAM Effectiveness." Perspectives 2 (2023).

Barlas, Dilek. "Key Developments at the World Bank Inspection Panel (2013–2024)." Perspectives 6 (2024).

Bissell, Richard E. "Three Decades of Seeking Elusive Remedies." Perspectives 9 (2023).

Bradlow, Daniel D. “A Proposal for One ‘Super’ Independent Accountability Mechanism.” *Perspectives* 33 (2025).

Fields, Natalie Bridgeman. “‘Use and Improve’ is My Accountability Mantra, Despite 30 Years of Eye-opening Disappointments.” *Perspectives* 14 (2023).

Kunanayagam, Ramanie, et al. “Glass Half-Full or Glass Half-Empty? Thirty Years of Accountability at the Inspection Panel — The Impact of its Work and What the Data Tells Us.” *Perspectives* 11 (2023).

Park, Susan. “Are the MDBs Accountable? Reflecting on the Independent Accountability Mechanisms of the Multilateral Development Banks.” *Perspectives* 20 (2024).

Schäfer, Dustin. “Between Disruption and Legitimation of Development: A Critical Perspective on the Inspection Panel.” *Perspectives* 19 (2023).

Selwan, Serge. “Accountability at the World Bank Group.” *Perspectives* 58 (2026).

Other sources.

World Bank Inspection Panel, case register and data page (inspectionpanel.org), as of June 2026.
World Bank Independent Evaluation Group, ICRR master database (March 2026). Human Rights Watch, reporting on Ethiopia Protecting Basic Services (2012, 2015). Yale Lowenstein Human Rights Clinic (2017). World Bank Group, Accountability Mechanism Task Force Draft Report (March 2026).

Annex A. Case Studies

The five cases below illustrate the recurring patterns discussed in the body of this paper. The first four trace what happens once a complaint reaches the Panel; the fifth examines the project level beneath it. None is offered as evidence of a coordinated strategy; together they are the cumulative expression of structural incentives that, across different cases, levels, and periods, produced outcomes favouring the operational relationship with borrowers over the accountability function. The first four are drawn from public sources — Panel reports, civil-society and journalistic investigations, and the published accounts of former Panel members; the fifth draws on the project’s own Implementation Completion Report and the author’s direct operational experience with it. Together they put faces and places to the statistics: the projects behind the 30.9 percent Satisfactory-or-better rate, the communities behind the Complaint Index, and an uneven, largely unverified record of whether remedy ever reached them — substantial in one case, partial in two, refused in one, and absent across three repeated attempts in a fourth.

Case 1 — Delay Before Registration: Badia East, Nigeria

Of the Panel’s procedural innovations, the Pilot for Early Solutions reaches furthest. It can delay both the investigation recommendation and registration itself: if Management “successfully addresses” the concerns raised, the Panel may decline to register the complaint at all — leaving no official case record, no Panel findings, and no public accountability document behind.

Badia East is the only formal use of the Pilot. In February 2013, roughly 9,000 residents of the Badia East informal settlement in Lagos were forcibly evicted during a World Bank-financed urban and water project; their homes were demolished with little notice and no adequate resettlement. Three residents filed a complaint. Rather than proceed to registration, the Panel steered them into mediation with the Lagos State Government. Internal emails later obtained by the International Consortium of Investigative Journalists show the Panel Chair working, in February 2014, to close the case before the mediation agreement could unravel. By July 2014 two of the three original complainants had rejected the negotiated outcome and asked for an investigation. The Panel declined.

The justification was that a lengthy investigation would not improve outcomes for communities who had already lost their homes. The logic is worth pausing on: the Panel assessed the adequacy of the process it had itself designed, against the outcome that process had produced, found it adequate, and closed the case. The complainants had no avenue of appeal. Amnesty International formally asked the Board to investigate; the Board did not act. Lagos State did eventually pay financial assistance to most residents — the Panel put the figure above 80 percent — but Amnesty International documented that the payments fell well below the value of what was destroyed and excluded any restoration of housing or livelihoods, the standard the Bank’s own resettlement policy requires. The residents of Badia East received a settlement neither they nor Amnesty regarded as adequate, no Panel findings, and no independently verified determination that Bank policy had been met.

Case 2 — Scope Limitation and Retaliation: Ethiopia Protecting Basic Services

No case better captures the power asymmetry at the centre of this paper. Across its phases, Protecting Basic Services was one of the largest operations the World Bank has run in Africa — a multi-donor flagship that channelled roughly \$17.5 billion in block-grant budget support to Ethiopian regional governments for health, education, water, and agriculture. Institutionally, it was exactly the kind of operation the Bank had the strongest reasons to defend.

In September 2012, Anuak indigenous people from the Gambella region — by then refugees in South Sudan and Kenya, represented by Inclusive Development International — filed a complaint alleging that PBS block grants were underwriting the federal government’s Villagization Programme: the forced relocation of tens of thousands of Gambella households off ancestral land that was, in parallel, being leased to foreign agricultural investors. Human Rights Watch documented relocations carried out by federal security forces and accompanied by beatings, arbitrary arrest, rape, and killings. The budget support and the villagisation moved through the same regional-government systems; the complainants’ case was that the Bank’s money and the coercion were operationally inseparable.

ETHIOPIA PBS: A TIMELINE

September 2012 — Anuak refugees from Gambella, supported by Inclusive Development International, file a complaint linking PBS block grants to forced villagisation.

July 2013 — After a five-month standoff in which the Ethiopian government publicly threatens not to cooperate, the Board authorises an investigation.

February 2014 — The Panel conducts a field visit in Gambella, in an environment controlled by the same federal security forces documented as instruments of villagisation-related abuse.

February 2015 — The Panel finds non-compliance on indigenous peoples, risk assessment, and the operational interface between PBS and villagisation — but excludes the forced-relocation and human-rights allegations as outside its mandate.

March 2015 — Three weeks after the report is published, Pastor Omot Agwa — the Panel’s own interpreter on the 2014 field visit — is arrested at Addis Ababa airport.

Outcome — The Board approves a Management Action Plan that preserves the block-grant architecture. PBS keeps disbursing and is relaunched under a new name (ESPES). The displaced Anuak Requesters’ appeal for individual livelihood rehabilitation is not granted; the MAP’s completed actions address PBS’s general operations in Gambella rather than the Requesters themselves.

Two features of the case are worth isolating. The first is scope limitation. The Panel found non-compliance — but explicitly carved the forced relocation and the human-rights abuses out of its inquiry as beyond its mandate, examining instead whether Bank staff had assessed risks and applied the indigenous-peoples policy. Inclusive Development International characterised this as the Panel “unilaterally” narrowing its review to the more palatable aspects of the complaint. The effect, as Section 4.5 argues, was that the entity actually conducting the villagisation — the government — was never the subject of the investigation; only the Bank’s supervision of itself was.

WHAT THE MANAGEMENT ACTION PLAN ACHIEVED

Management’s response emphasised the Panel’s finding that the worst-alleged harms — forced relocation, intimidation — were not direct consequences of PBS, while giving less weight to the policy violations the Panel did find. The Board approved a Management Action Plan that preserved the same untracked block-grant architecture the Panel had questioned. PBS continued to disburse, and the Bank later relaunched an almost identical operation under a new name. Per the Panel’s own Eligibility Report, Bank representatives had earlier told the Requesters directly that “no remedies can be offered by the Bank to address the Requesters’ issues other than continuing dialogue with the Government.” The displaced Anuak Requesters’ own appeal to Bank President Jim Yong Kim for livelihood rehabilitation and resettlement assistance was not granted.

The retaliatory arrest. The second feature is what happened to the people who helped. Three weeks after the Panel’s February 2015 report was published, Pastor Omot Agwa — the interpreter the Panel itself had engaged for the 2014 Gambella field visit — was arrested at Addis Ababa airport en route to a food-security workshop in Nairobi. He was held for six months without charge, placed in solitary confinement, and then charged under Ethiopia’s Anti-Terrorism Proclamation; he spent roughly three years in prison before the charges were dropped. The Panel’s former Executive Secretary records that it was this case that led the Panel to become the first international accountability mechanism to adopt anti-retaliation guidelines (Barlas, Perspectives 6, 2024). Direct appeals were made to Bank President Jim Yong Kim; the Bank stayed publicly silent. This is the case the institution does not cite when it describes the Panel’s contribution to accountability — and it is the case that, for the author, settles the question this paper poses: the affected communities did not obtain relief.

Case 3 — Finding Suppression: Amaravati, India

In May 2017, farmers from Andhra Pradesh filed a complaint about the proposed Amaravati Capital City Development Project, alleging harm from forced land pooling, inadequate consultation, threats to food security, and intimidation by state officials. The Panel registered the complaint, visited Amaravati in September 2017, and produced a report recommending investigation. The report was posted on the Bank’s website in early October 2017 — and removed within days.

THE OFFICIAL EXPLANATION

“The Report and Recommendation on a Request for Inspection — India: Amaravati Sustainable Capital City Development Project was inadvertently published in the World Bank’s website prior to the Board’s consideration. The Corporate Secretariat acknowledges this unfortunate administrative error and has taken immediate action to correct it.”

World Bank Corporate Secretariat, 8 October 2017.

No comparable removal of a Panel investigation recommendation prior to Board consideration has been found elsewhere in the Panel’s history. The Board then approved two consecutive deferrals — December 2017 (six months) and July 2018 (nine months) — each on the basis that Management had proposed additional actions, and neither citing any precedent for double deferral. When the state government withdrew the loan request in 2019, the Board approved the Panel’s recommendation not to investigate on the ground that the project would no longer be financed — not on any finding that the farmers’ concerns had been addressed.

Throughout the deferral period, farmers in Amaravati reported continuing intimidation and harassment by state authorities, documented in the Panel’s own field-visit record. The Bank made no public statement. The accountability process produced two deferrals and a closure without investigation.

The story did not end there. A related operation, the Amaravati Sustainable Infrastructure and Institutional Development Project (ASIIDP), was under preparation on the same site through 2018; the Panel recommended investigating it too, in March 2019, again on livelihood-restoration grounds. The Indian government withdrew its financing request the following month, and the Panel updated its recommendation to not investigate — a second closure triggered by the borrower’s exit rather than any finding on the merits.

In December 2024, the same land, the same capital-city plan, and substantially the same community produced a third Bank engagement. The Amaravati Integrated Urban Development Program (AIUDP) — an \$800 million IBRD credit, matched by an equal commitment from the Asian Development Bank — was approved using Program-for-Results financing, a results-based instrument with a narrower causal-link test than the investment financing used in 2017. Four requesters, later joined by a farmers’ federation, filed a new Request in October 2024, again raising the original land-pooling scheme, displacement, and livelihood loss. The Panel and the ADB’s Compliance Review Panel conducted a joint eligibility mission in March 2025, and both bodies acknowledged the harm directly: the Panel’s own report recorded that it was

THE THIRD CLOSURE

“...made aware of harm suffered by people and communities affected by the development” of the capital city, but could not establish that the new program — as opposed to the legacy land pooling — was its likely cause.

Inspection Panel, Report and Recommendation on AIUDP (P507508), April 2025.

Because PforR financing required the Requesters to show the new program itself, not the legacy land pooling, was the likely cause of harm, the Panel found it could not establish that link at this early stage of implementation and recommended against investigation. The Board approved that recommendation on April 29, 2025.

Three Bank engagements with the same community over eight years have now closed without an investigation ever reaching a finding on the merits: an administrative error in 2017, a borrower withdrawal in 2019, and a financing-instrument technicality in 2025. Each closure used a different procedural door; the destination was the same. India’s own Comptroller and Auditor General has since found that the land-pooling scheme’s objectives remain unmet and flagged over Rs 2,244 crore in wasteful expenditure, and independent reporting in 2026 raised credibility concerns about a state-funded social audit of the same land-assembly process. The harm the original 2017 Requesters described has outlived three separate accountability proceedings.

Case 4 – When Remedy Is Self-Monitored, Refused, or Unverified

The first three cases concern getting a complaint heard. The fourth concerns what happens after a finding of non-compliance, and it shows two recurring patterns in how that finding is handled: Management self-monitoring of its own corrective plans, which can still deliver substantial remedy after years of delay and without independent verification, and outright refusal by the borrower government to deliver the remedy the Panel’s findings imply.

Self-monitoring: Coal India and Vishnugad. After each of its investigations the Board approves a Management Action Plan, and the Panel tracks progress through periodic Management self-reports. It cannot independently verify implementation without a separate Board decision – the Panel must recommend verification, the Board must approve it, and Group Internal Audit must conduct it. That sequence is the exception, not the default; no systematic public assessment of MAP completion rates exists. In the Coal India case, Management took nearly six months to respond to a critical Panel report against the Resolution’s six-week timeline. The India Vishnugad hydropower case generated progress reports across nearly a decade (2016–2022); most resettlement and compensation actions were completed – THDC paid annuities to more than 2,500 affected families and delivered water and resettlement infrastructure – but six families in the hamlet of Hatsari remained in an unresolved dispute over demands beyond the agreed package, and a third Request for Inspection assessing ongoing harm was still pending nearly a decade after the original investigation. The Panel’s own 2022 review credited the end-term evaluation’s finding that livelihoods had been restored – but only after a third formal complaint and a decade of Management self-reporting were needed to reach that determination, with no independent verification along the way.

Refusal: Albania and Cambodia. The starker failure occurs when the borrower simply refuses the remedy. The Panel’s former Executive Secretary documents two cases plainly (Barlas, Perspectives 6, 2024). In the Albania Integrated Coastal Zone project, families in the village of Jale had their homes demolished; the Panel found a direct link to the Bank-financed operation and established that the Board had approved the project believing in a demolition moratorium the government had never actually agreed to – a fact Management had not disclosed to the Board. The Bank’s own retrospective later judged Management’s written response to the Panel “particularly unhelpful and non-informative, and at times in total conflict with factual information” the Bank had possessed all along (The Inspection Panel at 15 Years, World Bank, 2009, cited in Yale Lowenstein Clinic, 2017). Management proposed compensation; the government refused, agreeing only to fund the families’

legal costs. Of nine cases the families pursued in the Albanian courts, two succeeded. In the Cambodia Land Management project, more than 4,000 families around Boeung Kak Lake in Phnom Penh were evicted; the Panel found serious non-compliance, the government rejected the proposed remedial actions, and the Bank ultimately suspended all lending to Cambodia for five years. A later government sub-decree granted title to roughly 700 of the affected families – an outcome the authorities claimed was independent of the Panel case.

These cases share the structure set out in Section 4.4. The remedy a Panel finding implies must be agreed with the borrower; the borrower is often the party that caused the harm; and the Bank, dependent on the lending relationship, has limited willingness to compel a client. The result is the pattern the data in Section 2 describe from above and these cases describe from the ground: the mechanism reliably establishes that harm occurred, and rarely ensures that it is repaired.

Case 5 – The Mechanism Upstream: SEEFOR, Nigeria

The first four cases trace what happens once a complaint reaches the Panel. The fifth examines the layer beneath it: the project-level grievance redress mechanism every Bank operation is required to establish, and whose performance the Bank’s own reporting rarely measures. The Inspection Panel’s caseload looks small in part because the large majority of grievances are meant to be resolved long before they would ever reach Washington. What the project-level evidence shows is that the Bank does not know, in any systematic way, whether that resolution is actually happening.

The State Employment and Expenditure for Results (SEEFOR) project financed more than 4,300 community micro-projects across four Nigerian states for roughly 208,000 direct beneficiaries; the author served as a World Bank Task Team Leader on the project. Its Implementation Completion Report records that a grievance redress mechanism and Community Oversight Committees were established in all four states – the standard safeguards language that accompanies most Bank-financed operations. What the ICR does not report is any measure of how that mechanism performed: the number of complaints received, resolved, outstanding, or escalated; the time taken to resolve them; or the share that reached the Bank’s corporate Grievance Redress Service. The closest the ICR comes to a performance assessment is a single observation, that

SEEFOR IMPLEMENTATION COMPLETION REPORT

“...few community-related concerns [were] reflected in the GRM registers.”

a finding about what the registers recorded, not necessarily about what communities experienced.

The project team’s own response is instructive. Recognising that the paper-based registers were generating limited and unreliable information, SEEFOR built a parallel architecture to find out what was actually happening on the ground: GPS-verified digitisation of all 4,300-plus project locations, a call centre that directly contacted 20,000 beneficiaries, independent third-party monitoring, and an

AI-assisted citizen-feedback tool. That investment was a sound management response to a real problem, not a concealment of one. But it also demonstrates, from the inside, the pattern this paper documents from the outside: the Bank’s standard reporting framework treats the existence of a grievance mechanism as evidence of compliance, without requiring the mechanism to report whether it is functioning. A box checked “GRM established” tells an evaluator nothing about whether an affected person could actually be heard.

This is the layer the Inspection Panel — and the Integrated Accountability Mechanism that succeeds it — cannot see. No project-level GRM in the Bank’s standard reporting publishes complaint volume, resolution rate, or escalation count in a form comparable across projects, which means there is no systematic way to identify, before a formal Panel complaint is filed, which project-level mechanisms are failing the communities they were built to serve. This is the gap Recommendation 2 (Section 8) addresses directly.

THE MDBREFORM.COM ANALYTICAL SERIES

This paper is part of the mdbreform.com accountability and performance series. Key papers: Accountability: The Accountability Merger That Avoids the Accountability Question (June 2026); The World Bank Inspection Panel at Thirty-Three: Measuring Accountability Failure (this paper, revised June 2026).

Country Cases: Nigeria Water | Angola DPF | Ghana FCI | DRC Portfolio | DRC Inga | South Africa Eskom | Rwanda | Somalia | Kenya Education.

Zero Club Series (Parts 1–17+): Sector and country performance records, IEG master database, FY2015–2026.

All analysis at mdbreform.com. Evidence-based assessment for treasury officials, executive directors, and development finance professionals.