Report Number: ICRR0023204

1. Project Data

Project ID	Project Name	
P150116	MG-Public Sector Performance	
Country	**Practice Area(Lead)**	
Madagascar	Governance	
L/C/TF Number(s)	**Closing Date (Original)**	**Total Project Cost (USD)**
IDA-58350	30-Jun-2020	33,753,475.09
Bank Approval Date	**Closing Date (Actual)**	
13-Jun-2016	31-Dec-2021	

	IBRD/IDA (USD)	Grants (USD)
Original Commitment	40,000,000.00	0.00
Revised Commitment	35,006,046.75	0.00
Actual	33,753,475.09	0.00

Prepared by	Reviewed by	ICR Review Coordinator	Group
Nils Junge	Clay Wescott	Jennifer L. Keller	IEGEC (Unit 1)

2. Project Objectives and Components

a. Objectives

As per the Financing Agreement, the objective of the Madagascar Public Sector Performance Project (PAPSP) was “**to improve revenue management and local service delivery nationwide and in the education sector in selected regions.”**

For this ICRR (Implementation Completion Report Review) validation, the Independent Evaluation Group (IEG) parses the formal objective statement into two objectives:

- To improve revenue management; and
- To improve local service delivery nationwide, including in the education sector in selected regions

b. Were the project objectives/key associated outcome targets revised during implementation?
No

c. Will a split evaluation be undertaken?
No

d. Components
The Project Appraisal Document (PAD divided the project into four components, combining elements of technical assistance (TA) and incentives via results-based financing (RBF). Each component's activities and budget are summarized below (ICR pp 9-10):

Improving Revenue Management. (Appraisal: US$16.8 million, Actual: US$13.8 million). Under this component, the project sought to address revenue management weaknesses by providing TA for the: (a) improvement and consolidation of taxpayers' registration and identification; (b) modernization of the Integrated Tax Administration System (ITAS); (c) capacity strengthening of the audit and control functions for risk-based tax and customs administrations, audit planning and monitoring, as well as a revision of audit procedural manuals; (d) creation of an enabling environment for performance-based management at the tax and customs departments; (e) improvement of revenue collection in the mining sector; (f) improvement of accountability and transparency of the revenue agencies; and (g) development of the leadership, change and knowledge management capacity." Through the Results-Based Financing (RBF) instrument, it sought to incentivize: (a) modernization of the ITAS through equipment and software provisioning, related workshops, and communications; and expansion of the tax base through the improvement of offices and implementation of local tax offices; and (b) strengthening of audit function and control of the Customs Directorate.

Improving Service Delivery Capacity of Local Governments. (Appraisal: US$12.0 million, Actual: US$10.8 million). Under this component, TA was provided for: (a) strengthening of the grant transfer and equalization mechanism; (b) development of a strategy and review of the legal, institutional, and technical framework of revenue collection and the piloting of secure revenue collection; and (c) local governance knowledge management strengthening, including selected impact evaluations. The project used the RBF instrument to incentivize local government service delivery via grants to communes.

Improving Governance Mechanisms in the Education Sector. (Appraisal: US$7.8 million, Actual: US$4.5 million). Under the third component, TA was provided for: (a) improved budget planning, monitoring, execution, and procurement efficiency of funds to the local level; (b) implementation of the School Management Committee; and (c) development of Contracted Project Schools. At the same time, the RBF approach was designed to encourage the channeling of funds to the local level to reinforce budget planning and execution mechanisms by defining the payment of school grants (*Caisses écoles*) as eligible PEC (Contracted Project School (Projet d'école contractualisés) financing. (ICR p. 19) The project targeted three of the country's 22 regions (Alaotra Mangoro, Analamanga, Sofia), covering over 4,800 schools. The Performance-based Conditions (PBCs) were seen as a major encouragement of channeling funds to the local level, combined with the policy dialogue between the Bank and the Government in the context of the project.

Improving Controls and Performance Monitoring. (Appraisal: US$3.4 million, Actual: US$3.8 million.) Under this component, the project provided TA, capacity building, equipment, and operating costs to improve expenditure management by: (a) strengthening controls and oversight mechanisms; (b) promoting transparency, accountability, participation, and media access to information; and (c) strengthening project coordination and performance monitoring, including management of fiduciary, safeguards and monitoring activities, as well as the hiring of Independent Verifiers for PBCs and Eligible Expenditure Programs.

e. Comments on Project Cost, Financing, Borrower Contribution, and Dates

Project Costs: The project's original cost at appraisal was US$40 million. The Bank canceled $5 million and revised the project cost to US$35 million and disbursed US$33.75 million. The rationale for the cancellation was the failure to achieve two-time bound PBCs, equivalent to US$5 million (PBC 1.5 Publication of external ITAS Performance Audit Report and PBC 4.4 Completion of transfers of all investment and operating subventions to communes).

Financing: An IDA credit of SDR 28.4 million (equivalent to US$40.0 million at the time of appraisal) financed the project.

Borrower Contribution: The Government of Madagascar did not contribute financially to the project.

Dates: The Bank approved the project on June 13, 2016. The original closing date was June 30, 2020, but the Bank extended the project by a year and a half to close on December 31, 2021. The Bank completed the Mid-term Review on December 10, 2018. The extension was due to implementation delays during the early stages.

3. Relevance of Objectives

Rationale

This review finds that the objectives were highly relevant since the project directly supported the FY2017-21 Country Partnership Framework (CPF) (which came out after the Bank prepared the project but drew on the Systematic Country Diagnostic, prepared in parallel with the project). The ICR cites three (out of eight) CPF objectives to that effect: *Enhanced and effective decentralization (CPF Objective 3); Enhanced transparency and accountability (CPF Objective 4); and Priority Social and Infrastructure Spending (CPF Objective 5).* Because of Madagascar's status as a fragile, conflict-affected and violence-affected (FCV) country (designated in the wake of the political unrest), the ICR emphasizes the critical need to strengthen the government's ability to collect revenues, and its ability to deliver services in an accountable manner, which it argues are paramount for enhancing political stability.

The project also aligned with the World Bank's Interim Strategy Note (ISN) priorities, prepared for FY2012-2013, of reversing the post-crisis decline in tax revenues and mobilizing additional funds for key public services. IEG agrees with the project's rationale for the selected components and activities.

The Bank considered improved governance, encompassing greater institutional accountability and decentralization, to be a central pillar in rebuilding the Malagasy state and economy to put it on a course toward more inclusive growth and poverty reduction. The project was aligned with the National Development Plan prepared by the government and approved by Parliament. While the project team did not design it to address the underlying issues that led to the 2009 crisis, it aimed to foster an improved governance environment. All else being equal, it would thus contribute to decreasing the likelihood of political instability, i.e., by improving and strengthening institutional accountability and decentralization, which the ICR notes is "a central pillar of efforts to rebuild the Malagasy state and economy and set it on a course towards more inclusive growth and poverty reduction."

The next step for this review is to assess the strategic nature of the alignment of each sub-objective with the CPF. For the sub-objective, increased revenue management, the problem was that Madagascar's ratio of tax revenues to GDP was one of the lowest in the world. It had declined from 11.2 percent of GDP in 2011 to 9.9 percent in 2015 (PAD p. 2) making revenue management, and especially better tax collection, a priority for the government. TThe theory of change developed by the ICR was that better management and governance at the tax administration would increase revenue collection, through a combination of monitoring, assessment and management tools.

The project design divided the activities that would lead to these outcomes into four areas: updating the single identifier number system, modernizing the ITAS, strengthening risk-based audit functions and controls, and creating a performance-based management system while improving mining sector revenue management. For the Tax Administration Diagnostic Assessment Tool (**TADAT)** outcome, the project assumed that improving revenue management due to better accounting and forecasting, increased performance accountability, and improved risk management oversight systems would improve TADAT indicators. These would result from better planning, the number of suspicious contracts evaluated, and digitalization. The project's support for certifying tax administration officials and performance contracts would feed into these outputs. The project took a holistic approach and assumed that the underlying factors behind poor revenue management were a combination of technical, incentive, and accountability factors (in line with the Bank's Governance and Anticorruption Strategy), and tackled the issue from these multiple angles. This approach was pitched correctly for the country's capacity.

For the sub-objective of improving local service delivery nationwide, including in the education sector in selected regions, both the NDP and the ISN in Madagascar highlighted the urgent need to reverse the decline in public services, worsened by the crisis and declining budgets. Food insecurity, healthcare, and school enrollments were major concerns, particularly in rural areas where poverty rates were almost twice as high as in urban areas. Access to electricity was also among the lowest in the world. The crisis significantly impacted primary education, with enrollment increases dropping sharply and completion rates declining. The education sector accounted for 21 percent of the public budget, but the overall contraction of public spending reduced allocations to the sector. Household contributions rose due to increased reliance on community teachers with limited teaching qualifications. The project targeted three regions with over 4,800 schools (ICR, p. 6).

The theory of change was appropriate for the context: that **school-based management would be strengthened** if: i) more Community-hired (*Fikambanan'ny ray aman-drenin'ny mpianatra*) teachers received subsidies in a timely way, ii) School Management Committees-- Farimbon'Ezaka ho Fahombiazan'ny Fanabeazana eny Ifotony-- (FEFFIs) teachers received subsidies in a timely way, iii) FEFFIs had better management and planning capacity, and iv) a human resource report was published. These mid-term outcomes would happen if procedures were streamlined, more schools had functional FEFFIs, and the teachers' dashboards were published and updated. Based on the activities, the project assumed that technical and informational factors were standing in the way of better school management, hence the support for a general civil servant census, a mapping of teachers, a review of payment processes, and communication campaigns. The project assumed that these activities, although they were discrete, one-time inputs, would have sustainable impacts. Another assumption made by the project is that **increased accountability at community levels** would result in citizens being better informed on public finance management and increased government accountability and transparency. These were to be achieved via the following project outputs: i) public debates by CSOs on the budgets, ii) metrics (local governance barometer and user satisfaction reports), iii) high-risk audits, all of which the project would support via training and capacity building activities. The planned activities indicate that the project assumed a combination of capacity building, greater citizen involvement, and oversight (via audits) would improve local government accountability. These assumptions rested, in part, on there being sufficient interest on the part of community members to play more active roles. Questions arise about whether these outputs would become regular features once the project ended

Rating

High

4. Achievement of Objectives (Efficacy)

OBJECTIVE 1

Objective

To improve revenue management.

Rationale

Relevance of the results indicators

The ToC was credible for the reasons set out in Section 3. The project had two PDO indicators linked to the objective of improving revenue management:

- Increased tax revenue collection (domestic tax revenue/GDP ratio)
- Number of TADAT Performance Outcome Area (POA)-8 indicators 'Efficient Revenue Management' rated B.

The Bank and the government intended to measure the achievement of the PDO "increasing the domestic revenue/GDP ratio" more efficient revenue generation and collection among key revenue agencies.

This support introduced diagnostic tools (the Tax Administration Diagnostic Assessment Tool, the other results indicator linked to this objective), performance contracts, management improvements, and oversight and control mechanisms. The Bank and the Government considered improving revenue management as a necessary condition for government stability in the post-conflict period, and through PAPSP, the Bank aimed to address challenges around informality and leakages, weak capacity, human resources, management issues, and lack of internal controls.

PBCs 1, 2, and 3 (which were the same as IRIs 1, 2, and 3) incentivized these achievements through:

- Modernization of the Integrated Tax Administration System (PBC 1)
- Increased rate of confirmed suspicious customs transactions related to home use in selected Customs Office, in case of physical examination (PBC 2)
- Number of revenue offices subject to external evaluation of their performance contracts/programs (PBC 3)

The PBCs are relevant for improving revenue management by making various elements within the system more accountable, via audits, assessments of customs transactions and evaluations of revenue offices. Indicator PBC 2 could have been formulated more clearly—the rate of confirmed suspicious transactions is outside of administrative control. PBC 3 is straightforward to implement and to measure and encourages compliance. However, context is missing for this indicator, as neither the PAD nor the ICR state how many revenue offices there are. Does the target of 8 offices represent a large or small share? Another problem with the results indicator is that while project activities could increase the numerator of the tax revenue/GDP ratio, they had no effect on the denominator. An indicator more focused on areas supported by the project would have been more suitable.

The relevance of the results indicator is rated **Modest**.

The indicator "Number of Tax Administration Diagnostic Assessment Tool (TADAT) Performance Outcome Area (POA)-8 indicators 'Efficient Revenue Management' rated B" was introduced to improve revenue management. The indicator's logic is that its multiple outcome areas would capture a range of good management attributes supported by the project.

The relevance of the results indicator is rated **Substantial.**.

Outputs

The following key outputs were linked to the objective:

1. Integrated Tax Administration System deployed.
2. Taxpayer registration system improved.
3. 102 Tax administration officials TADAT certified.
4. Performance contracts established and evaluated.

The project also supported the set-up of an interoperability framework between tax, customs, and pension funds registries, improving tax-payer compliance and information accuracy. Compliance was improved through the tax agency's improved ability to detect, through data matching, which firms or individuals are either evading taxes or under-declaring their taxes. However, related to implementation delays, the project did not assess information accuracy before the project close. According to the results framework (RF) summary, the project did not achieve any of the PBC targets. Nonetheless, while the final targets were not met, significant progress was made in these areas that would have been unlikely without the PBC incentives. For example, the government would not have begun the modernization of ITAS (PBC 1).

Outcomes

Taxpayer registrations more than doubled from 200,000 to 515,272, which was, however below the target of 700,000. The project pilot-tested a tax "Short Message Service" (SMS) system, which was highly successful. The project team reported that the technology resulted in a return of $329 on every dollar spent in collecting unpaid taxes (p. 15). (The ICR does not explain whether or not other factors played a role.) In any case, in absolute numbers, the return totaled just US$375,000, a fairly negligible amount, compared to the total tax revenue (approximately 1.1 billion) collected in 2015. Despite the potential benefits of the SMS system, the government stopped using the program following project closure. This was partly attributed to difficulties in planning for and meeting recurrent costs. The project team observed that, in the face of budget constraints, operational costs for technology tend to be the first to be slashed. The program may or may not resume.

Tax revenue growth was linked to an expansion in the tax base, but the growth in tax revenue was interrupted by the impacts of COVID-19. Tax revenue/GDP ratio increased to 10.4% (from 10.0%) to 11.68% in 2018, exceeding the target of 11.4%. However, revenue collection fell back to 9.1 percent of GDP in 2020 as a result of the lockdowns from COVID-19.

The 2019 GDP re-basing was also assessed to have a significant impact on the tax to GDP ratio barely changing. The 2020 restructuring paper noted in the revisions to the Results Framework for this indicator that "The proposed target values were readjusted to align with the GDP rebasing undertaken in 2019 in light of the general population census. Before the GDP rebasing, the tax GDP ratio was estimated at 11.9 percent in 2019. With the rebasing done in 2019, this ratio dropped at 10.4 while the increase of revenue in nominal terms was on a positive trend—total revenue rose from USD 1.35 billion in 2018 to 1.46 billion in 2019. Additionally, the projections were further adjusted to reflect the current macro-economic context due to the COVID-19 outbreak, resulting in an estimated revenue loss of 30% based on a 3-month lockdown hypothesis and a projected recession of -0.8% in 2020.

By 2021, the ratio had risen to 10.4%, thus still failing to meet the target of 11.4% and therefore considered "not achieved" by the project. While the net increase in tax revenue/GDP was modest (following the disruptions related to the COVID-19 pandemic), the ICR argues that in the absence of the project, revenue might have fallen even further and that adaptive measures for revenue management were key to "implementing services that enable digital interaction with taxpayers (for example, a digital tax payment platform) (PAD 48).

Only one TADAT PO-8 indicator, "Efficient Revenue Management," met the target rating of "B" (the target was for two to score at least a "B"), although two improved, based on the government's self assessment. The ICR reported that indicator P8-26 "Contribution to the government tax revenue forecasting process" met the target by changing from D to B, but that the second indicator "P8-28 "Adequacy of tax refund processing" only changed from D to C. The reasons for failing to meet the target was related to "changes in government and related project disruptions as well as the COVID-19 pandemic in the final years of the project" (ICR pp. 14-15).

Increased rate of confirmed suspicious customs transactions (PBC 2) reached 17.2%, close to its target of 20%, from a baseline of 5%. The certification of tax administration officials and performance incentives helped raise the rate. Other factors may have contributed to the change (both promoting the increase or inhibiting a greater increase). For example, changes in administrative/systems procedures or the presence of other programs. The ICR mentions that the government switched to a 100% scanning system in 2020, which experts do not consider good practice. Without fuller information on what other activities and incentives were at play or what other factors may have influenced the change in this indicator, it is difficult to determine the relative importance of the project in this achievement.

Although the targets were not met, the project was effective in moving the indicators along a positive trajectory, developments which, tax revenue as a share of GDP aside, can mostly be attributed to the project. IEG rates the efficacy of the objective as **Modest.**

Rating
Modest

OBJECTIVE 2
Objective
To improve local service delivery nationwide, including in the education sector in selected regions

Rationale
Relevance of the results indicators

Three results indicators were linked to the second objective:

- Share of communes undertaking an IGL self-assessment published on the government's public service delivery performance website.
- Publication of a Human Resource Management Report within the Ministry of Education (MEN).
- Number of female beneficiaries participating in public debates.

To make them more measurable and realistic, and in light of project progress, according to the ICR, in 2020 all three results indicators were revised from the original, which were:

- Percentage of communes with a Local Governance Index (Indice de Gouvernance Locale, IGL) higher than 6.
- Availability of a teachers' information management system.
- Direct project beneficiaries, % of which are female.

The second objective--to improve local service delivery nationwide--was closely linked to the first. A core goal of increasing tax revenue is to enable governments to provide public services. An acute need to address public service delivery capacity was evident as the system had become increasingly degraded, as following the political crisis, already weak social services (including education, health and food security) declined rapidly. The ICR describes infrastructure falling into "an alarming state of despair" (ICR p. 5). Education was selected as a priority area within the public service delivery system, given the sector's "dramatic reversal under the crisis" (ICR p. 6).

PBCs 4 and 5 (the same as IRIs 7 and 9) supported the objective of improving local service delivery, including in the education sector in selected regions, through:

- Improved timeliness of service-delivery fiscal transfers to local governments (PBC 4)
- Share of schools targeted by the component with an operational FEFFI, planning and implementing PECs (PBC 5).

In developing the results indicators—related to local government performance and transparency of school management—and the accompanying PBCs, the Bank project team made the assumption that quicker fiscal transfers would improve social services and that more active FEFFIs would lead to greater accountability and better use of school resources. PBC 4 is a measurable indicator, assuming data is available. Its relevance is based on the assumption that fiscal transfers' timeliness, and therefore reliability, was a key factor inhibiting local government performance, rather than the amount of the transfers they receive. The PAD notes that the "Government transfers only 5 percent of the total revenues to decentralized entities" (PAD, p 4.) which suggests the problem goes much deeper than reliability. It also notes that the capacity of local governments to collect revenue is constrained. The project did not address this issue, however. While the ability of local governments to count on revenue matters, more information is needed on how relevant timeliness is against adequacy. Neither the PAD nor the ICR provides a rationale. The ICR notes that it was difficult to track the indicator, suggesting it may not have been the optimal choice. Considered as a package with the PBC 4, (one focused on local governance performance, the other on fiscal transfers to the local level), the PDO indicator "Percentage of communes with a Local Governance Index (Indice de Gouvernance Locale, IGL) higher than 6" is rated **Substantial.** On the other hand, neither the PAD nor the ICR provide sufficient evidence for why human resource management reporting is a specific barrier to school performance. The reporting was to complement the HR system, as per the ICR, but it is left unclear why reporting, as opposed to the system itself, was a key barrier. Therefore, the indicator Publication of a Human Resource Management Report within the Ministry of Education (MEN) is rated **Modest.** The third indicator, the number of female beneficiaries participating in public debates, was selected to meet the gender tag requirement. The project seems to have made the assumption that women's engagement would contribute to better school management related to their role in the household. The indicator is straightforward to measure if attendance rolls are taken. It is not clear, however, whether women merely attended debates (in order to meet quotas) or actively participated. The indicator is rated **Modest.**

Outputs

Key outputs linked to the second objective were:

1. Communes trained to meet improved selection criteria and performance-based criteria.
2. All communes operating arrears cleared.
3. Conditional grants delivered to all communes within agreed timeframe.
4. Communes received training and technical assistance from the fiscal leadership initiative.
5. Over 400 communes conduct IGL self-assessment and published results online.
6. Timely delivery of Court of Account's annual public report.

7. Share of schools with operational FEFFI increased and implementation of PECs.
8. Public debates by CSOs on budget law and project performance conducted.

The revised *"Indice de Gouvernance Locale"* or IGL (local governance indicator) target was exceeded: 450 communes undertook an IGL self-assessment published on a government public service delivery performance website, exceeding the target of 400. The IGL was intended to help local governments assess their performance and as an incentive for local governments to access central government funding. Given that there are 1,700 communes, the original score would have implied that 1,020 would have reached a score of 6 out of 10. By the end of the project, less than half that number had conducted a self-assessment.

Under the project restructuring, the IGL indicator was changed from hitting a specific IGL score to "undertaking and publishing" an assessment without regard to a score. The project team considered the revised indicator more realistic since it took into account the capacity of communes to adopt the use of the IGL assessment. The indicator “Availability of a teachers’ information management system” was changed to “Publication of a Human Resource Management Report within the Ministry of Education (MEN).” The publication did not happen.

The ICR reports that a dashboard to monitor teachers was created as part of the education sector activities. This involved cleaning an existing MEN database and making it user-friendly and publicly accessible on the Ministry’s website. This would have allowed the teachers to be tracked. However, the dashboard was not updated every quarter, and toward the end of the project, it was replaced by an alphanumeric table with partitioned data. Confidence that the dashboard will be used is low since the Ministry of Education made the decision to abandon it. The goal was considered only partially achieved, and significant concerns over its sustainability exist.

Outcomes. The ICR notes that the timeliness of fiscal transfers to local governments improved insofar as the delays were reduced. The target of 30 days for the indicator "Improved timeliness of service-delivery fiscal transfers to local governments (PBC 4)" was not achieved, with transfers completed only within 60 days by the end of the project. Could the change in transfer timeliness also be due to other factors? This would be possible to tease out through comparative analysis between the participating governments and others, but the ICR does not mention this.

The project significantly expanded the number of FEFFIs operating in the three selected regions. However, it faced setbacks during the pandemic as FEFFIs require face-to-face interactions that are incompatible with social distancing measures. The ICR reported that higher levels of FEFFI engagement improved community engagement and school budget management, although evidence was not provided to support these assertions. Results varied significantly between rural areas, where the targets were often underachieved, and urban areas, where they were overachieved: in project areas, 68 percent of FEFFIs were operational, compared to 28 percent in non-project areas. Considering the failure to significantly improve the timeliness of transfers to local governments, although the ICR reports that the transfers were completed, combined with the improved performance of FEFFIs, the efficacy of the objective is rated **Modest.**

Rating
Modest

OVERALL EFFICACY

Rationale

Three of the five outcome indicators and targets of the PDO were revised in the restructured project. The project efficacy is rated modest due to its contributions to revenue management, the flow of funds to local governments, and oversight and control mechanisms. Results were mixed for the five PDO indicators, with two targets exceeded, two partially achieved, and one not achieved. The project faced challenges due to weak pre-existing capacity, frequent government changes, and the COVID-19 pandemic.

US$ 23.5 million out of US$ 32 million (69.3 percent) of project financing took the form of PBC–linked disbursements, which considers the $5 million cancellation of PBC 1.5 Publication of external ITAS performance Audit Report and PBC 4.4. The project allocated the remaining financing for technical assistance. Although not all PBCs successfully incentivized the Government to reach the project targets—for example, the tax and customs administration reform elements (e.g., process improvements and performance contracts) declined over time—the ICR described PBCs as a “powerful incentive to modernize [the Tax Administration General Directorate] and its fragmented systems.” The ICR attributed this to a lack of additional investment budget from the central government, which suggests that there were limits on how much PBCs could, on their own, stimulate sustainable behavior change. Nonetheless, while the project did not meet most of its PBC intermediate results indicator (IRI) targets, their use seems to have been critical in incentivizing the government to move forward on multiple fronts (Modernizing the Integrated Tax Administration System; increasing the rate of confirmed suspicious customs transactions; targeting schools with an operational FEFFI, planning and implementing Contracted Project Schools (PECs).

Overall Efficacy Rating	**Primary Reason**
Modest	Low achievement

5. Efficiency

The ICR conducted a thorough, multi-pronged efficiency analysis. It covered the financial efficiency of revenue mobilization, potential economic benefits, direct and indirect benefits of the project through grants to communes, governance improvements, indirect benefits, and project implementation efficiency.

Money generated resulted in increased tax revenue (as a share of gross domestic product, GDP) and spillover effects from infrastructure investments (benefiting the country’s 1,695 communes). The ICR cites governance improvements and an improved business environment (thanks to reduced border compliance times for exports and imports, by 82% and 85%, respectively).

The ICR replicated Component 1 and Sub-component 2.2's cost-benefit analysis to assess the project’s Value for Money.

At appraisal, the Bank estimated that mobilization activities, at US$17.2 in financing under Component 1 and Activity 2.2, would result in a financial net present value (FNPV) of US$261.05 million (using an 8 percent discount rate). This was based on the Bank’s projections of tax revenue increasing from 10.0% to 11.8% as a share of GDP. The financial analysis at appraisal predicated these changes on specific project activities: improving internal audit and control, improving taxpayer's registration to increase the tax base, rollout of an ITAS which would improve quality of control to reduce revenue leakages, improving transparency in customs management, and improving mining governance to increase formal mining exploitation and then ameliorate the potential of revenues.

The ICR estimated that the revenue increase between 2016-2018 (US$433.9) was already 26 times the project's expenditure on revenue mobilization. Extending this to 2016-2025, the ICR calculations resulted in a higher estimate, of $700.3 million in FNPV. In contrast to the PAD, the ICR did not assume that 100 percent of the returns could be attributed to the project. It took into account other contributing factors, specifically: related technical assistance projects by other development partners; and the project's small scope relative to the country's GDP and government revenue figures, and macroeconomic factors outside project control. The ICR, therefore, assumed that only 80 percent of increased revenues are attributed to the project. The decision to use 80 percent attribution rate was a judgment call made by the Bank, based on a review of available data and consultations with project team members. The project team argues that World Bank was recognized as the leading partner working on increasing revenue generation with the government, and that the PAPSP project was considered the leading project on revenue generation in the country during this period, providing support in terms of digitization of both customs and tax system which mitigated the external shocks. The World Bank, furthermore, "was also the only development partner to have a sustained dialogue on revenue issues" (ICR p. 49). The ICR also includes a sensitivity analysis for different assumptions regarding the Bank's contribution, at 5 percent, 10 percent, 30 percent, and 50 percent. Even using a conservative assumption of 30 percent contribution to the impact, (US254.7) the increased revenue was close to the project target of US$261. FNPV thus exceeded the appraisal estimate, although ICR notes the pre- and post-project estimates are not strictly comparable since some data inputs were unavailable. Financial IRR estimates under the sensitivity analysis ranged from 3.6 percent (at 5 percent attribution) to 73.1 percent (at 80 percent attribution.) The ICR also applies an economic NPV, using a fiscal multiplier of 0.2, which gives correspondingly higher returns.

The improvements in collection were predicated on specific project-supported activities:

In the project's early years, tax revenue collection went up significantly to 11.68 percent of GDP (exceeding the target of 11.4%). However, this upward trajectory was interrupted by the COVID-19 pandemic, and revenue collection fell back to 9.1 percent of GDP in 2020. In 2021 it increased again, to 10.4 percent. If the positive trend continues, as the FNPV assumes, through 2025, the return on investment will be substantial.

The ICR reports that taxpayer registrations increased from 200,000 to 515,272 (a figure below the revised target of 700,000 but above the original target of 320,000), linked largely to the introduction of performance contracts. The ICR argues that the "Adoption of improved processes and use of performance contracts combined with strong ownership and motivation of the tax and customs agencies led to increased taxpayer registrations." However, the ICR does not provide any evidence to this assertion. Moreover, the ICR notes that the system of performance contracts ultimately fell out of use, as buy-in was limited (ICR, p. 16).

The ICR recognizes that non-project factors will have contributed to the rise in revenue—development partner technical assistance and macroeconomic factors—and that the project's relatively small scope to the size of the economy must be considered.

The ICR observes that the Bank was one of the only interventions focused on the main sources of income, as well as the only development partner to engage the Government in a sustained dialogue on revenue issues. The task team also pointed out that substantive improvement in four out five indicators supported by PBCs indicated these incentives were effective. All the above speak in favor of significant project attribution.

The assessment of investment grants was based on deduction and not quantified. As the ICR concedes, no sub-project evaluations were conducted, which could have helped assess their efficiency. The ICR analysis of Component 2. Improving Service Delivery Capacity of Local Governments, which was positive, was limited to a qualitative discussion and did not demonstrate any efficiency benefits. It is likely that improved predictability of fiscal transfers to local governments would have reduced transaction costs for service delivery and improved budget planning.

Even using a conservative reading of the project's contribution to revenue generation and without considering economic benefits which were not quantified, the ICR makes a case that the project generated significant returns on investment. Despite the macroeconomic disruptions caused by the pandemic and the delay in project implementation. Nonetheless, given the questions about the effectiveness of the performance bonus, it is not clear that the project contributed substantially to the rise in revenue. IEG, therefore, rates the efficiency **Modest.**

Efficiency Rating

Modest

a. If available, enter the Economic Rate of Return (ERR) and/or Financial Rate of Return (FRR) at appraisal and the re-estimated value at evaluation:

	Rate Available?	Point value (%)	*Coverage/Scope (%)
Appraisal	✓	13.00	44.60 ☐ Not Applicable
ICR Estimate	✓	73.10	51.00 ☐ Not Applicable

* Refers to percent of total project cost for which ERR/FRR was calculated.

6. Outcome

The relevance of the project objectives is rated High, while efficacy is rated Modest. Efficiency is rated Modest.

PDO and intermediate results indicators (IRIs) and targets for the second part of the PDO, "to improve local service delivery nationwide, including in the education sector in selected regions," were revised from outcome to output indicators. This reflected the project team's assessment that the original targets were overly ambitious for the conditions. This was the case with IGL governance assessment scores and the teacher's management information system. In the latter case, the publication of a Human Resource Management Report was deemed adequate for meeting the objective instead of making the system available. For the indicator "percentage of women as direct project beneficiaries," the project substituted a target number for their participation in public debates. IEG notes that a less ambitious approach would have been to lower targets for the outcome indicators, not to convert them to output indicators.

The outcome under revised targets:

Under revised outcome 1, neither of the two PDO indicator targets was fully achieved, and four of the six IRIs were not achieved, including two IRIs with PBCs, and the IEG rates it **Modest**. The rating is based on institutional instability within the tax administration and the inability of the government to sustain newly introduced practices, including process improvements and performance contracts. Development and deployment of the ITAS were not completed by the project end.

Under revised outcome 2. of the targets for three PDO indicators, two were exceeded, and one was partially achieved. The ICR reports that the Project “substantially contributed to improved local government service delivery, enhanced budget planning and execution mechanisms, and strengthened school-based management in the education sector.” It found that investment grants funded diverse community needs, that the timeliness of fiscal transfers to local governments improved (due to the PBC incentive), and that IGL self-assessments led communes to adopt good governance standards, leading to improved service delivery. However, none of this is direct evidence of service delivery improving, only of transfers being made more quickly and local governments conducting self-assessments. The ICR itself points out that the “effectiveness of the IGL to enhance better service delivery in the medium term will largely depend on how the public can use this tool to hold local authorities to account” (ICR p. 18). Neither does the ICR explain why the timeliness of transfer led to these improvements. The project did not seem to have increased spending; it only accelerated its disbursement. Based on the evidence, IEG rates the outcome as **Modest.**

a. Outcome Rating

Moderately Unsatisfactory

7. Risk to Development Outcome

The ICR identifies several risks to development outcomes that will need to be mitigated. Risks to sustaining project outcomes include i) reduced resources; ii) lack ownership for project-supported activities; iii) reversal of progress achieved by ministries and agencies; iv) customs administration achievements fading in the absence of external support, given the lack of commitment to use of performance contracts; v) the risk of domestic and international shocks affecting the fiscal space and limiting project achievement (ICR p. 29).The argument for sustainability rests on the government’s commitment to fiscal decentralization and the widespread acceptance of the IGL assessment tool. Another Bank project, PRODIGY (the Digital Governance and Identification Management System Project) will continue the development of ITAS and dialogue on digitization and reform efforts.

At the same time, the discontinuation of performance contracts by the customs administration and reversals of good practices such as risk analysis tools and mirror statistics is worrisome, reflecting a lack of commitment. Without these features, the operation’s revenue generation targets are at risk. Persistently high corruption, combined with external economic and climate shocks, poses further risks to gains made by the project and to government commitment to good governance.

8. Assessment of Bank Performance

a. Quality-at-Entry

Strategic relevance and approach

The project objectives selected aligned well with the CPF and the Government's national development priorities. The Bank considered development partner support and focused on alleviating the critical constraint of low public revenues. The project design had a strong analytical grounding. However, several critical weaknesses offset these positive elements, as detailed below.

Policy and institutional aspects.

The ICR characterized project design as ambitious, given the "fragile country context and political and institutional instability episodes." Preparation was informed by a Quality Enhancement Review and Reform Program for the Efficiency of the Administration (*Programme de Réforme pour l'Efficacité de l'Administration,* PREA), an anchoring unit of the former Governance and Institutional Development Project II. It had gained legitimacy as the responsible entity for coordinating public sector reforms nationwide. The ICR notes that, while PREA played a coordinating role, it lacked the capacity to prepare the Project Operational Manual and Terms of References, and external technical assistance to support these areas was not envisioned. The PAD anticipated risks related to PREA's capacity and envisioned mitigation measures by reinforcing PREA with core management (project officer dedicated to the project) and additional fiduciary staff. However, the ICR does not mention a dedicated project officer or fiduciary staff.

Fiduciary aspects.

Risk is rated substantial in the PAD and was to be mitigated by implementing an action plan developed during the fiduciary assessment. FM management of PREA was evaluated as adequate. Mainstreaming social accountability mechanisms would also mitigate risk (PAD, pp. 21-23). However, this assessment underestimated the fiduciary risk, and mitigation measures were woefully inadequate, as evidenced by poor fiduciary compliance (see Section 10b).

Risk assessments.

As noted, the selection of indicators compromised the M&E system and, thus, the project's ability to measure results. In the end, only four out of 13 IRIs were achieved. This suggests that expectations were unrealistic, the Bank did not sufficiently consider policy and institutional capacities, or the mitigation measures were inadequate to overcome the identified risks.

IEG agrees the project design was sound with respect to its constituent components. However, the project's ambitions were unrealistic given the political economy dynamics and capacity issues. The decision to choose the RBF approach was based on reasonable concerns over weak governance. Still, the ICR notes that it was the first time this type of instrument had been used in Madagascar, so the government was unfamiliar with its modalities, which led to delays.

Quality-at-Entry Rating
Moderately Unsatisfactory

b. Quality of supervision

Project supervision was fairly weak initially but improved during the final two years of implementation, when the task team, with a task team leader (TTL) now based in the country, took serious steps to address previous shortcomings.

High turnover of TTLs led to irregular task team leadership, with a total of five different TTLs, and a period without any TTL. The project's first phase, before the second and third restructurings, was characterized by inadequate supervision.

The ICR notes that the task team failed to promptly take steps to address unsatisfactory progress on outcomes and implementation. The IP and DO all reached moderately unsatisfactory (MU) levels by the third ISR, and the ISRs sometimes failed to update on all components or indicators. Ratings and narratives were inconsistent, and the reporting quality was weak and not always transparent about the project's issues.

At the same time, the Bank made significant efforts to address implementation weaknesses related to PREA capacity. The project team noted that it developed a roadmap to improve project performance, instituted weekly meetings with the government, and increased supervision intensity. Various measures were proposed to help address PREA capacity weaknesses, but their implementation was stymied, and in the end, implementation was shifted to a Project Implementation Unit.

Given that the Outcome rating is Moderately Unsatisfactory, and the two dimensions of Bank performance are Moderately Satisfactory and Moderately Unsatisfactory, the overall performance is weighted downward.

Quality of Supervision Rating
Moderately Satisfactory

Overall Bank Performance Rating
Moderately Unsatisfactory

9. M&E Design, Implementation, & Utilization

a. M&E Design

The ICR identified several weaknesses with the M&E system design, as discussed above in Section 4. Many indicators did not align with good practice, i.e., specific, measurable, achievable, relevant, and time-bound (SMART). The ICR questions the use of the TADAT indicator, given that results were beyond the control of DGI. In general, outcome indicators are rarely fully under the control of a project. While the ICR criticizes the project for indicators that were output rather than outcome-based (a low-risk approach to setting targets since outputs are easier to influence than outcomes), it points out that the TADAT indicator results were beyond the control of the project. Given these drawbacks, the Bank could have chosen a better indicator than the "Number of TADAT Performance Outcome Areas rated B."

The RF failed to capture the project outcomes and objectives adequately. The first results indicator of "increased tax collection" aligned well with the PDO but was also dependent on GDP, clearly outside any project's control. The basis for selecting direct beneficiaries (final target 3,277) and the link to project objectives were not specified. IGL, a composite good governance index tested by some local governments in 2010, was described in the PAD (p. 88) as a tool for local government reflection and self-assessment.

The project team reported that the Local Development Funds package for investment grants was, in effect, a performance-based grant. Eligibility criteria included the availability of the communes' budget for the upcoming year and the timely availability of communes' administrative accounts. The project provided capacity-building activities on budget planning and budget execution, including: (i) coaching by the Districts and local representatives of the Ministry of Finance on budget planning; (ii) capacity building provided by Local Development Funds on budget planning, project management, procurement and even accounting (budget execution stage) before the grants are transferred; and (iii) in the final two years of the project, close involvement the local representatives of the Court of Account, the "Tribunaux administratifs et financiers" (administrative and financial Tribunals) to review controls on communes financial statements, and provide recommendations and "appui/conseil" for future budget planning. The government expanded and continued the activities following project closure. And a performance for results (P4R) project is in the pipeline to support the government's program in decentralized service delivery

The PAD provided minimal baseline information regarding the IGL, other than that, 3% of communes had scores of 6 or more. There was no information on how many communes had conducted self-assessments, average IGL scores, or distributed IGL scores.

Indicators were not well defined, which the ICR attributes, at least partly, to the absence of a TOC (as indicated above, it was not required at the time of project preparation.) While the reconstructed TOC clarified the project's design decisions, the framework did not guide indicators.

b. M&E Implementation

M&E implementation struggled from the beginning of the project, with M&E rated Moderately Unsatisfactory for most of the project life. There were reporting delays and missing indicators in the ISRs. PREA relied on the Court of Accounts (CA) as the independent verification agency (IVA) for M&E. However, the CA only monitored selected indicators.

Weak M&E capacity and poor indicator design compromised the M&E system. The ICR concludes that assigning M&E responsibilities to PREA, which lacked project management capacity, meant the Project's M&E system was "stalled from the outset." The pandemic further complicated implementation.

Toward the end of the project, some improvements were made as the CA took on greater responsibility, including utilizing a digital library. The RF revisions introduced during the 2020 restructuring addressed some of the issues. The ICR notes that "certain discrepancies between data submitted by the ministries and findings from the CA" persisted until project closing.

At the same time, the revisions the project made to several PDO indicators are a sign that the project team recognized their inadequacy as originally formulated and took action to improve their measurability and achievability, even if it led to less ambitious targets.

c. M&E Utilization

Given the numerous issues with M&E design and implementation, it is unsurprising that M&E was not used to inform project management and decision. No instance of the system being used is mentioned in the ICR.

M&E Quality Rating

Modest

10. Other Issues

a. Safeguards

The project triggered two safeguards, was rated as “Category B” and required a partial assessment. Environmental Assessment (OP/BP 4.01) and Involuntary Resettlement (OP 4.12) were the safeguards triggered. The risks of land acquisition and resettlement were identified but considered unlikely. A Resettlement Policy Framework (RPF) was prepared and disclosed, but no resettlement activities are mentioned in the ICR. Public consultation and information disclosure were held as the RPF was prepared, and feedback was incorporated into the project design. However, the PAD does not include an annex with a mitigation plan, and the ICR does not refer to the RPF or relevant mitigation measures, as no land acquisition or resettlement took place. However, the ICR notes that “required Environmental and Social Management Framework and a Resettlement Action Plan were prepared and disclosed before project appraisal.” The ICR reports that safeguards were rated “Satisfactory” overall throughout the life of the project, and any minor issues that emerged were quickly resolved. Robust and consistent environmental and social management toolkits were made available, and grants disbursed by the Local Development Funds were monitored for environmental and social standards compliance. No negative social or environmental impacts associated with the project were cited in the ICR.

b. Fiduciary Compliance

The project rated poorly on fiduciary compliance, from moderately unsatisfactory to moderately satisfactory (as per ISRs). The ICR identified four contributing factors: i) poor understanding among the implementing agencies of the new procurement system (STEP), which was introduced during the project life cycle, and a new procurement framework; ii) ineligible expenditures; iii) a vacant internal auditor function throughout the project; and iv) delays in updating the manual of procedures. It is unclear from the ICR why the last two issues were not resolved, given that they should have been under project control. Procurement delays were in part due to the low capacity of PREA, the project implementer. Remote working arrangements because of COVID-19 also caused delays in procurement.

c. Unintended impacts (Positive or Negative)

No unintended impacts are mentioned in the ICR

d. Other

The ICR cites the project's poverty reduction impacts, inferring that investments in critical infrastructure in the commune helped spur local development, but it does not provide evidence

11. Ratings

Ratings	ICR	IEG	Reason for Disagreements/Comment
Outcome	Moderately Unsatisfactory	Moderately Unsatisfactory	
Bank Performance	Moderately Satisfactory	Moderately Unsatisfactory	Unrealistic expectations related to known capacity constraints
Quality of M&E	Modest	Modest	
Quality of ICR	---	Substantial	

12. Lessons

Building on the lessons outlined in the ICR, IEG the following lessons were learned from the project:

- PBCs may be less effective if financial incentives do not align with or are not invested in the agency or institution responsible for carrying out reforms. The ICR attributed the declining motivation of the tax and customs administration to continue with reforms to process improvements and performance contracts to it not receiving an additional investment budget. While this may or may not be related to incentives, the agency in question was asked to change or improve its practices without more resources.

- If indicators are not clearly defined, they will be difficult (or impossible) to measure. The ICR points out various weaknesses related to the selection of project indicators, which hampered their measurement. For example, the original RF included an intermediate results indicator, “Project beneficiaries, a percentage of which are women,” which posed a challenge since the project did not specify who is considered a project beneficiary. As the ICR noted, the “scope of the project yields multiple direct and indirect beneficiaries dispersed nationwide” (p. 26). The Bank recognized the issue and replaced the indicator with one clearly defined and more easily measurable (the number of women participating in public debates).

- Adequate attention needs to be paid to the M&E system to ensure it functions as intended and contributes to accountability and lessons. The PIU appears to have largely neglected its M&E function, relying on the IVA and Court of Accounts for reporting. As a result, M&E reporting was sub-optimal. The Bank also could have done more to monitor the M&E activities during supervision. Stronger M&E implementation could have uncovered issues with indicators earlier. The ICR notes that the weak M&E element contributed to the need for project restructuring.

13. Assessment Recommended?

No

14. Comments on Quality of ICR

The ICR is of high quality. It provides a thorough analysis of the project and its activities, a well-designed theory of change, and clear summary tables. It also includes detailed quantitative economic and financial analysis, with substantive economic analysis of the revenue mobilization activities at the local and national levels and an assessment of the project's additional direct and indirect benefits. At the same time, more information on the mechanisms that led to higher tax revenue would have been useful (given how poorly Component 1. activities performed and weak implementation). However, the ICR took a creative approach to reconstruct the financial analysis, given that some data inputs behind the PAD’s financial analysis were not accessible.

The conclusions and ratings were well-founded for the most part. Although many of the IRIs were not met, it makes the case that significant progress was made in a challenging institutional environment. However, the ICR could have explained how the PBC incentives contributed to the outcomes better, given that it rated three out of the five IRIs (to which PBCs were linked) as “not achieved.”

A clearer explanation for what is rated “achieved” vs. “partially achieved” would have helped. For example, while the project moved very close to the 20% target for IRI 2, “Increased rate of confirmed suspicious customs transactions” (from 5% to 17.2%) the IRI was rated “not achieved.”

Some issues remain unclear and could have been explained in more detail, for example, a) the specific mechanisms that led to increases in tax revenue and taxpayer registration; b) the rationale for changing some of the indicators and softening their targets, such as the IGL scores; and c) the narrative on how well the incentives mechanisms worked.

No negative social or environmental impacts associated with the project were cited in the ICR.

a. Quality of ICR Rating

Substantial