Policy Research Working Paper 6297

Capacity Constraints and Public Financial Management in Small Pacific Island Countries

Tobias A. Haque
David S. Knight
Dinuk S. Jayasuriya

The World Bank
East Asia and the Pacific Region
Poverty Reduction and Economic Management Unit
December 2012

Policy Research Working Paper 6297

Abstract

Drawing on Public Expenditure and Financial Accountability assessment scores from 118 countries, this paper provides the first comparative analysis of public financial management performance in small Pacific Island Countries (PICs). It applies a Tobit regression model across the full cross-country sample of Public Expenditure and Financial Accountability scores and country variables to identify potential causes for the observed underperformance of Pacific Island countries relative to other countries of similar income. First, the analysis finds small population size to be negatively correlated with Public Expenditure and Financial Accountability scores, with the "population penalty" faced by small Pacific Island countries sufficient to explain observed underperformance. Second, through application of a new capacity index of Public Expenditure and Financial Accountability dimensions, it finds strong evidence in support of the hypothesis that small population size impacts scores through the imposition of capacity constraints: with a limited pool of human capital, small countries face severe and permanent challenges in accessing an adequate range and depth of technical skills to fulfill all functions assessed through the Public Expenditure and Financial Accountability framework. These findings suggest that approaches to strengthening public financial management in small Pacific Island countries should involve: i) careful prioritization of public financial management capacity toward areas that represent binding constraints to development; ii) adoption of public financial management systems that can function within inherent and binding capacity constraints, rather than wholesale adoption of "best practice" imported systems; and iii) consideration of options for accessing external capacity to support public financial management systems on a long-term basis, from regional agencies, the private sector, or donors.

This paper is a product of the Poverty Reduction and Economic Management Unit, East Asia and the Pacific Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at thaque2@worldbank.org.

Capacity Constraints and Public Financial Management in Small Pacific Island Countries

Tobias A. Haque, David S. Knight, and Dinuk S. Jayasuriya[1]

Keywords: **Public financial management, Pacific Island Countries, institutional development, governance, small-island states**

JEL: **H11, H60, H83**

Public Sector Governance **(PSM)**

[1] Tobias Haque and David Knight are economists in the Poverty Reduction and Economic Management unit of the World Bank, Pacific Division. Dinuk Jayasuria is a consultant and research fellow at the Development Policy Center of the Crawford School, Australian National University. Valuable comments and input were provided by Vivek Suri and Virginia Horscroft. Matt Andrews kindly shared the PEFA dimension categorizations used in his various papers. Stephen Hartung assisted in developing the "capacity" index for PEFA dimensions.

1. Introduction

Public Financial Management is a key concern of governments and development practitioners. Institutions governing public finances have a determining impact on the economic and social costs and benefits of revenue collection and expenditure. Some have recently gone so far as to claim that 'successful "societal evolution" hinges on the systems and procedures societies develop to manage public finance and procurement' (Porter, Andrews, Turkewitz, and Wescott 2010). Reflecting its importance for the achievement of development outcomes, increasing attention and resources are being devoted to strengthening the public financial management systems of developing countries worldwide (World Bank IEG 2008). Global trends towards increased investment in Public Financial Management reform are now being played out in the Pacific, where donors – dealing with recent and projected increases in aid flows and strong political imperatives to make greater use of country systems while ensuring value for money and fiduciary control – have substantially increased efforts in public financial management reform. Most of the independent Anglophone Pacific countries currently have PFM reform programs of some sort underway, often supported by international technical assistance.

Following similar regional and global analyses, this paper provides the first detailed quantitative analysis of PFM performance in small Pacific Island countries (PICs) using the Public Expenditure and Financial Accountability (PEFA) framework. The PEFA framework allows standardized assessment of PFM systems against "good practice" norms, with performance against 31 indicators, and 73 dimensions of the public financial management system scored using an A-D scale, and "A" scores representing "international good practice". At the time of writing, the PEFA framework had been applied in 118 countries, including nine PICs. Using cross-sectional data from PEFA assessments, we examine the PFM performance of twelve small PICs (Fiji, Kiribati, Nauru, Niue, Tuvalu, Vanuatu, Samoa, Tonga, Cook Islands, Marshall Islands, Palau, and Solomon Islands) in relation to broader global patterns. We approach this analysis in two steps.

Firstly, we examine the performance in PEFA assessments of small PICs relative to other countries with similar characteristics. We present simple comparisons between PICs and other country groups. We find that the performance of PICs lags the performance of countries at similar levels of income in other regions. We further find that this result is driven by a common pattern of unusually poor scores against a small set of particular PEFA dimensions, including those relating to procurement, internal audit, and strategic budgeting. Controlling for other factors that have been shown to influence performance in PEFA assessments, we identify population size as an important driver of PEFA performance and find that the observed poor performance of small PICs can be largely explained by their small populations.

Secondly, we examine the causal linkages between small population and poorer PEFA performance, both globally and in the Pacific. Using a number of tests, we find strong support for the hypothesis that small countries perform more poorly due to the capacity constraints.

Facing sustained and severe shortages of trained staff, small countries struggle to successfully complete all of the processes assessed through a PEFA assessment. While smallness exerts an overall negative impact on PEFA scores, this negative impact is most strongly felt in areas where highly specialized capacities are required to fulfill PEFA-assessed functions and especially where high-capacity functions need to be undertaken by multiple staff and beyond central agencies at the level of line ministries.

Policy implications of these results remain to be further developed. Initially, this analysis suggests the need for: i) realism when establishing targets for PFM reforms in the region, given the extent to which performance is likely to be constrained by shortage of capacity; ii) careful prioritization of scarce technical capacity towards carefully prioritized PFM reforms that address binding constraints to broader development progress; iii) a rebalancing in the deployment of technical assistance towards line ministries, where capacity constraints appear to be most severe; and iv) consideration of options for outsourcing of various technical and highly-specialized roles on an ongoing basis, given the low likelihood that such capacities can be sustainably sourced from local labor markets in small PIC settings.

The paper is structured as follows. In the second section we provide a review of the existing literature regarding quantitative analysis of PEFA scores and institutional development in small country contexts. In the third section, we provide a description of data sources and methodology. In the fourth section we summarize overall Pacific PFM performance and identify the country characteristics that may be driving poorer performance in small PICs. In the fifth section, we test the hypothesis that capacity constraints are the causal linkage between small population size and lower PFM performance through the application of existing analytical frameworks and the introduction of a capacity index for PEFA indicators. In the final section we present policy conclusions and recommendations for further work.

2. Literature

Increased investment in public financial management reform has been accompanied by the development and widespread application of tools to assess PFM performance and measure reform progress. Public Economic and Financial Accountability (PEFA) Assessments have become ubiquitous. The resulting standardized scorecards provide a quantitative basis for identifying patterns in PFM systems and assessing the effectiveness of reform efforts across different countries and regions. Some of these assessments have been general in nature, and provided basic insights regarding relationships between PFM performance, as measured through PEFA scores, and other country characteristics. De Renzio (2009), for example, using simple bivariate analysis, attempts to identify relationships between overall PEFA performance and performance against the basic “dimensions” measured in the PEFA framework using scores from all 60 PEFA assessments undertaken at the time of writing. He finds simple relationships between PFM performance and income, region, population, resource dependency, and various governance indicators. Of particular interest to those working in small states, he finds that

population is weakly negatively correlated with PEFA scores, but can identify no clear pattern. He notes, however, that "binary associations are not necessarily significant from a statistical point of view and could therefore be potentially misleading". Perhaps more usefully, De Renzio (2009) also presents a basic multivariate regression analysis, including assessment of the impact of population size. Results from the multivariate, in contrast to findings from the bivariate analysis, show that population is positively correlated with PEFA scores. No causal explanation for the observed relationship is suggested, other than that "there could be economies of scale in investing in budget systems in larger countries" (De Renzio 2009: 10).

Andrews (2010c) confines his analysis to Africa, investigating patterns in PFM performance and relationships between PFM performance and country characteristics using PEFA data. Through coding PEFA indicators according to different categories of process, Andrews reaches a series of powerful conclusions regarding the political economy of PFM reform in Africa and the tendency of PFM systems to take on the "form" of international good practice, without necessarily altering underlying and politically-influenced functions (see also Andrews 2009). For African countries, scores for budget preparation processes are comparatively stronger than those for budget execution and oversight processes, reflecting the fact that it is easier for developing, and especially fragile countries, to develop a budget than to ensure its effective and accountable implementation. Actual practices also tend to lag the implementation of new laws and processes, with stronger scores against indicators that can be improved by stroke-of-the-pen measures than those that impact on the way things are done and resources allocated on the ground. Finally, scores are generally higher against indicators for processes that can be implemented by a narrow, concentrated set of actors. Conversely, processes are generally weaker when they involve multiple players, especially outside of central agencies, such as Ministries of Finance. Andrews (2010c) also reaches some broader conclusions regarding country characteristics influencing PFM performance in Africa, showing correlations between improved PEFA performance and sustained economic growth, stability, and higher non-resource domestic revenues. Andrews (2010c) also identifies some possible impacts of colonial heritage.

Porter, Andrews, and Wescott (2011), drawing on Andrews' earlier analytical framework, use PEFA data to reach various conclusions regarding the particular characteristics of PFM systems in fragile and post-conflict settings. They show that the disconnect between the "laws, rules, and procedures adopted for better public finance and procurement and their actual functionality" noted by Andrews (2010) is particularly pronounced in post-conflict environments.

No research to date has focused on patterns in PEFA scores across PICs, despite increased programming support and policy attention to this area. This is perhaps unsurprising, given the relatively recent availability of PEFA data and the small population of these countries. Further, no analysis undertaken to date has specifically addressed the potential impact of human capital shortages on PFM performance.

A growing literature, however, examines the experiences of economic growth and institutional development in very small countries, and makes a compelling case that such countries face particular disadvantages, including as the result of capacity shortages. This literature provides a theoretical basis for our proposition that capacity shortages are the binding constraint to PFM performance in very small countries.

The literature on economic development in small states is too extensive to be summarized here. Bertram and Watters (1985), Winters and Martin (2004), and Gibson and Nero (2007) argue that PICs, in particular, face immutable barriers to economic growth and private sector development, due to diseconomies of scale in production of goods and provision of public services, and high costs of distance. The World Bank has recently contributed to this type of analysis with policy pieces emphasizing geographical constraints to economic growth in the PICs in general (World Bank 2011) and Solomon Islands in particular (World Bank 2010).

More relevantly, a small literature (Brown 2010, Baker 1992, Wittenhall 1992) specifically examines institutional development in small states, typically arguing that smallness leads to important weaknesses in public administration relative to larger countries because of the "limited pool of skilled human resources to perform the vital roles of the public service and a lack of depth in specialization which affects implementation and, by extension, absorptive capacity" (Brown 2010: 56). Without adequate skilled individuals to undertake the vital functions of government, the quality of public administration inevitably suffers. Wittenhall (1992:51) notes that, with weaknesses in education and training and very small populations, governments may simply not be able to find any staff with requisite skills. Baker (1992b) further notes that individuals who attain specialist skills within government are often attracted to broader overseas job opportunities in contexts where "only a few brains need be drained before a serious systemic crisis occurs" (Baker 1992b:16). It is not difficult to imagine how such issues could impact on PFM performance in Pacific Island states, where a recent PEFA assessment conducted in Kiribati noted the administrative problems created by the existence of nine long-term unfilled director-level financial management positions in the Audit Office and Ministry of Finance.

More evidence regarding PEFA performance and the potential impact of capacity gaps in very small country settings could inform improved understanding of the causes and potential cures for weaknesses in public financial management performance and other areas of public administration.

3. Data and Methodology

We define "small" PICs as those with populations of less than one million. The small PICs for which PEFA data exist are Cook Islands, Fiji, Kiribati, Marshall Islands, Micronesia, Nauru, Niue, Samoa, Solomon Islands, Tonga, Tuvalu, and Vanuatu but exclude larger Pacific states (Papua New Guinea and Timor Leste).

Country variables are drawn from widely used sources. GDP per capita and population data are taken from the World Bank's World Development Indicators (WDI). Data on aid flows is drawn from the OECD DAC database and migration data is from UNDESA databases. A full set of the definitions and sources of variables used is given in Appendix A.

PEFA scores were obtained for 162 PEFA assessments in 118 countries over the period 2005 to 2011 from the PEFA Secretariat. The PEFA assessment comprises 71 individually assessed dimensions which together form 31 high level indicators (including assessment of donor systems, 68 and 28 respectively excluding them). The assessments are made on an ordinal basis, assigning a score of A to D based on a set of criteria, with A indicating highest performance, with an additional '+' score possible for some indicators To facilitate statistical manipulation, we convert the ordinal scores to a numeric value of 1-4, with 4 being equivalent to 'A' and a '+' score ascribed an additional 0.5. This methodology follows existing practice and PEFA Secretariat guidance (PEFA Secretariat 2009). Appendix B provides summary statistics at the PEFA assessment level.

Following Andrews (2010), we exclude from our analysis dimensions that are beyond the direct control of government. Indicators related to aggregate fiscal outturns are omitted (PI-1, PI-2, and PI-3), as are indicators related to donor practices (PI-D1, PI-D2, and PI-D3). Omission of these variables reflects our interest in the success or failure of government administrations in undertaking PFM functions, and avoids results being biased by macroeconomic factors or the different practices of development agencies operating in different countries. These omissions are especially important in any analysis of PICs, which are atypical in both their exposure to macroeconomic volatility and reliance on donor aid flows.

We carry out analysis at the most disaggregated level, using all 65 separate dimensions included in each assessment, rather than relying on the 25 high level performance indicators that aggregate some of the dimensions that have been used in previous analyses.

Missing data is categorized by PEFA assessors as "NA" (not applicable), "NU" (not used for the assessment), or "NR" (not rated due to insufficient information). Our treatment of missing data is in line with suggested practice (PEFA Secretariat 2009). We treat NA and NU scores as missing and exclude them from the analysis, but code NR as a 0 score, since in most cases NR will correspond to a function that is either not being carried out at all, or is being completed to such a poor standard that it would not achieve a D score. This approach avoids a systematic upwards bias in the scores of countries with poorer PFM systems, such as those we are primarily interested in.

Given the limited range of values, we follow the literature on similar constructed, ordinal data (Elbadawi and Randa, 2003; Bates, 2006) in testing for the presence of censoring. As described by Rigobon and Stoker (2006), the presence of censoring can be tested for using a simple Chow test on the stability of the coefficient estimates between a sub-sample of possibly censored data

(scores of 1 and 4 in this case) and others (scores between 1.5 and 3.5). Carrying out such a test provides evidence that the data may indeed be censored. In the presence of censored data, OLS estimates would be biased and inconsistent, and the Tobit estimator – which simultaneously estimates the probability of data being censored and the marginal effects – would be a more appropriate regression model. We use the Tobit regression model throughout, although noting that results are not sensitive to the use of OLS (key results replicated with OLS reported in Appendix C).

Given we do not aggregate PEFA scores according to any groupings, we are able to exploit the full information available in the sample without placing any restrictions on the data. Although the data have a time element, there are only a relatively limited number of repeat PEFA assessments on the same country over time, so it is not feasible to use a time series approach in any systematic manner. Therefore, we use an unbalanced pooled-panel model throughout following the model:

$$Y_{ijt} = \alpha + \beta X_{it} + \gamma Y_j + \varepsilon_{ijt}$$

where Y_{ijt} is the PEFA score for country i, PEFA dimension j at time t. X_{it} is a vector of country-specific variables, and Y_j is a vector of categorical variables relating to the PEFA dimensions. Therefore, repeat assessments of a country's PFM system over time are used as separate observations, although clustering of errors is accounted for by employing cluster-robust standard errors throughout. Inclusion of repeat observations improves the sample size. Country control variables include real GDP per capita and log population, both of which are measured as the 5-year average over the period {t-4, t}[2]. Net migrations rates and the ratio of overseas development assistance to GDP are also used, as similarly formed averages.

No weightings are assigned to dimensions. This may provide cause for caution in interpreting some results. The importance of some PEFA dimensions may be greater than others if weaknesses against some dimensions undermine the relevance of scores against other dimensions, or if certain dimensions have greater impact on overall service delivery or macroeconomic management outcomes. The use of un-weighted averages, however, is consistent with previous literature (Porter, Andrews, and Wescott 2011). Ascribing weights to specific dimensions would be problematic, given the absence of any consensus in the existing literature regarding the relative importance of various PEFA dimensions or particular PFM functions.

4. PEFA Performance of Small Pacific Island Countries and Drivers

In this section we outline overall patterns in performance of small PICs relative to comparators. We identify country characteristics relevant to the small PICs that are associated with stronger performance in PEFA assessments.

[2] If data is not available for all previous periods, we average over available data.

4.1 Small Pacific Island Country PEFA Performance

Overall, PICs tend not to perform strongly in PEFA assessments. Column 1 of Table 1 indicates that membership of the small PIC group is associated with significantly lower average PEFA scores.

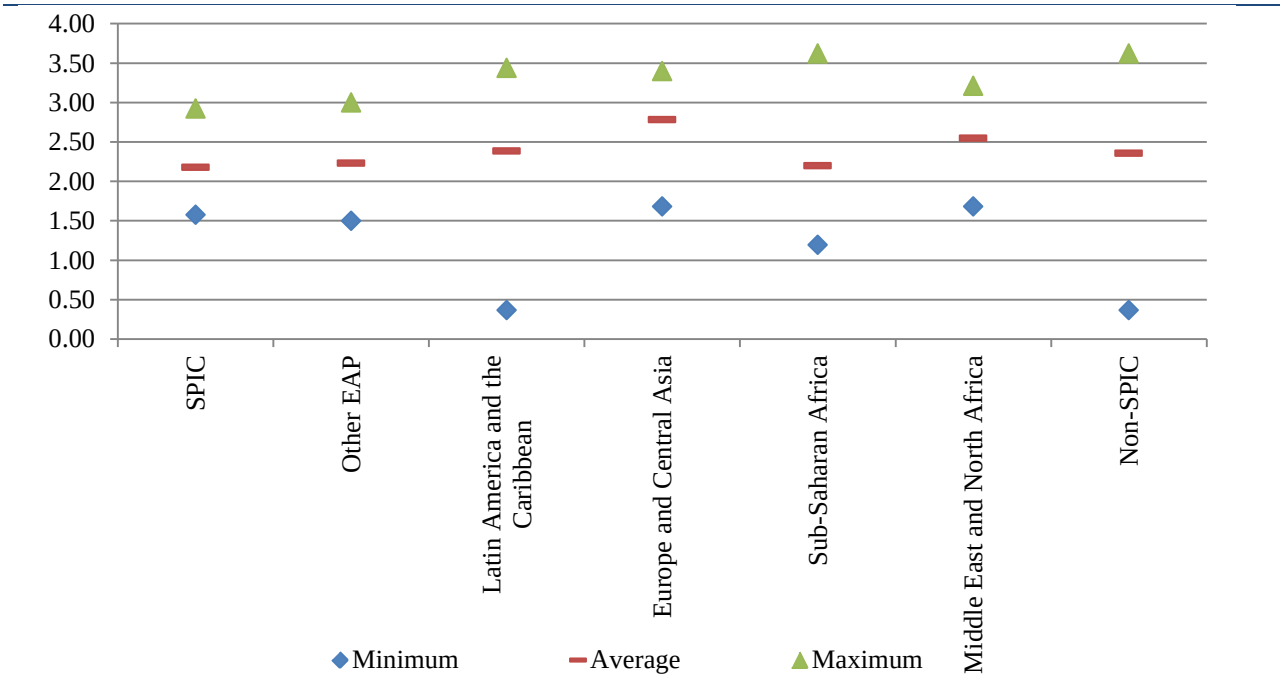
Table 1: Basic Regressions

	(1)	(2)	(3)	(4)
small PICi	-0.602***	-0.813***	-0.033	
	(0.001)	*(0.000)*	*(0.885)*	
LGDPit (log income)		0.327***	0.475***	0.489***
		(0.000)	*(0.000)*	*(0.000)*
LPOPit (log population)			0.200***	0.201***
			(0.000)	*(0.000)*
Time trend	YES	YES	YES	YES
Regional dummies	NO	NO	NO	YES
Number of observations	10,948	10,948	10,948	10,948
Number of clusters	162	162	162	162
Adj. R-Squared	0.003	0.013	0.022	0.025

Notes: Dependent variable is discrete PEFA score. All regressions use Tobit estimation.
p-values in parentheses are based on robust and clustered standard errors. *** p<0.01.
Constant is included but not reported.

Figure 1 shows the highest, lowest, and mean average score for different countries by country group. The overall PEFA score for PICs is below the global average and also below the average for all other country groups. The range of overall average PEFA scores for small PICs, however, is relatively narrow. The PIC with the highest overall average PEFA score (Vanuatu) performs substantially worse than the strongest performers in the developing world. The poorest performing PIC performs substantially better than the poorest performers globally.

Figure 1: High, Low, and Average Score by Country Group



4.2 Income as a Determinant of PEFA Performance

The most obvious possible explanation of lower PEFA scores for small PIC is lower incomes. A country with fewer resources available to invest in the human capital or pay and provide facilities for public servants is likely to be less capable of implementing the tasks measured by the PEFA framework. Accordingly, a consensus finding from previous work in this area is the substantial impact of income levels as a predictor of PEFA scores (De Renzio 2009; Porter, Andrews, and Wescott 2011). The respective income levels of our sample of PICs and relevant comparators therefore also need to be taken into account when making international comparisons. The majority of small PICs are lower-middle income. The exceptions are Fiji, which is Upper Middle Income and Solomon Islands, which remains a Low Income country. It may therefore be reasonable to expect small PICs to perform worse on PEFA assessments, on average, due to the scarcity of financial resources required to support PFM systems.

In column 2 of Table 1, income is positively correlated to improved PEFA performance. But after controlling for income, small PICs perform even less well than would be expected, with the small PIC dummy becoming more negative and remaining highly significant.

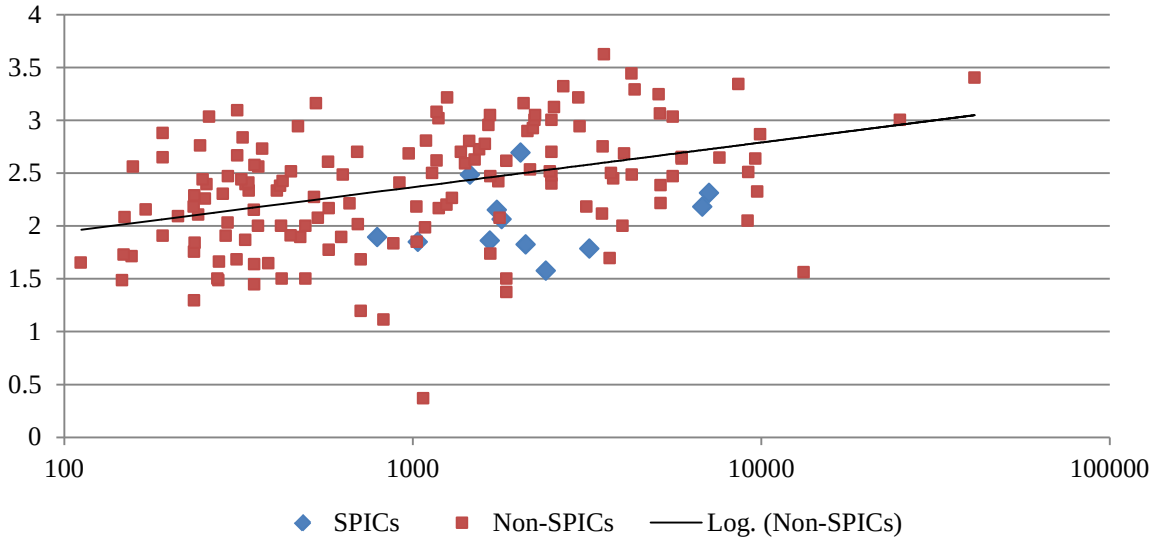
The aggregate picture of poorer-than-expected PICs is somewhat complicated by an examination of specific country performance. As shown in Figure 2, ten out of twelve small PICs have poorer PEFA scores than would be predicted by their level of income. But one country achieves a score almost exactly consistent with the income-predicted level, while another scores significantly

higher than would be predicted by income level. It is worth noting that the methodological approach to the PEFA assessment taken by the single country achieving above the predicted average score has been questioned by PFM practitioners in the region and is commonly considered to have been an unjustifiably favourable assessment.

Across the twelve small PICs in the sample, higher levels of income are clearly no guarantee of higher PEFA scores. As might be expected, the highest scoring countries are among the wealthiest of the group. But the wealthiest country in the region, and the sole upper-middle-income country, has a lower average PEFA score than three of the lower-middle-income countries in the group.

Overall, it is clear that income exerts an important impact on PEFA scores. Lower-than-average PEFA scores of small PICs can be partially attributed to their relatively low levels of income. But differences in income alone do not appear sufficient to explain lower PEFA scores. Further, the PICs that might be expected to have the strongest scores based on their levels of income fail to do so.

Figure 2: PEFA Score and Income Per Capita



Box 1: Similarities in small PICs' PEFA performance

Aggregate PEFA scores provide a very blunt tool for examining PEFA performance. Andrews (2010) groups PFM functions assessed within then PEFA framework according to the functions that they perform. This allows us to examine performance against particular functions (such as procurement, audit, and reporting) rather than relying on either just overall performance, or the very specific individual PEFA indicators which each provide only partial information on any particular function of the system (for example, an individual PEFA indicator within the procurement function is "evidence in the use of open competition for award of contracts that exceed the nationally established monetary threshold for small purchases").

Figure 3 and Figure 4 below show scores against various "functions" for all small PICs. It is clear that procurement processes, internal audit, and donor practices weaken overall PEFA performance for nearly all PICs, with all PICs clustered close together at scores of 2 or less along these three areas. In addition, Strategic Budgeting and Accounts and Reporting (special reporting and annual reporting) are two other areas where all-but-one of the small PICs have scores of 2 or less. These are PFM areas that require substantial technical capacity. As with most countries, small PICs' overall scores are also dragged down by low scores against measures regarding the predictability, reporting, and execution of aid-flows.

Figure 3: PEFA indicators on which small PICs' scores are closer together

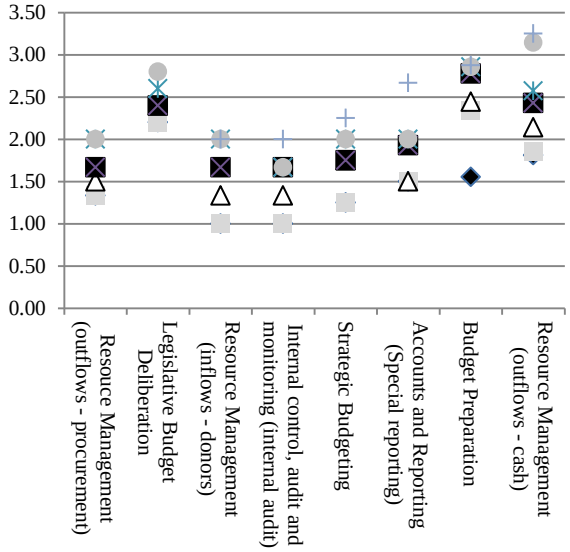
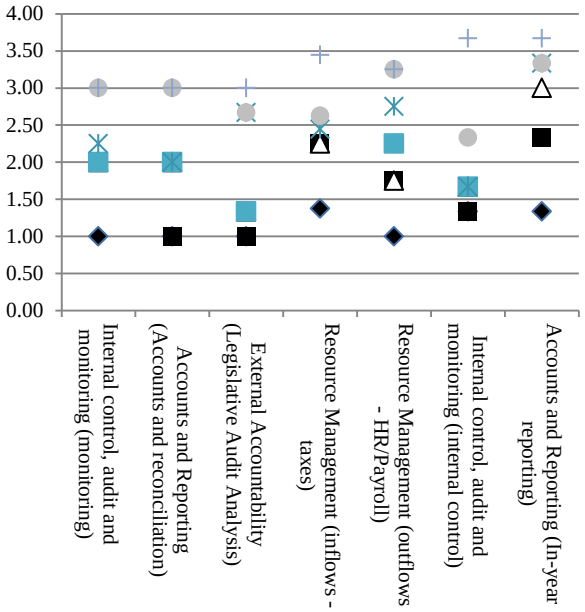


Figure 4: PEFA indicators for which small PICs' scores are wide ranging



While there are some shared areas of poor performance, there are few shared areas of high performance: different small PICs score highly against different functions. With the exception of common fairly strong performance against legislative processes for consideration of the Budget, different countries perform well in different areas.

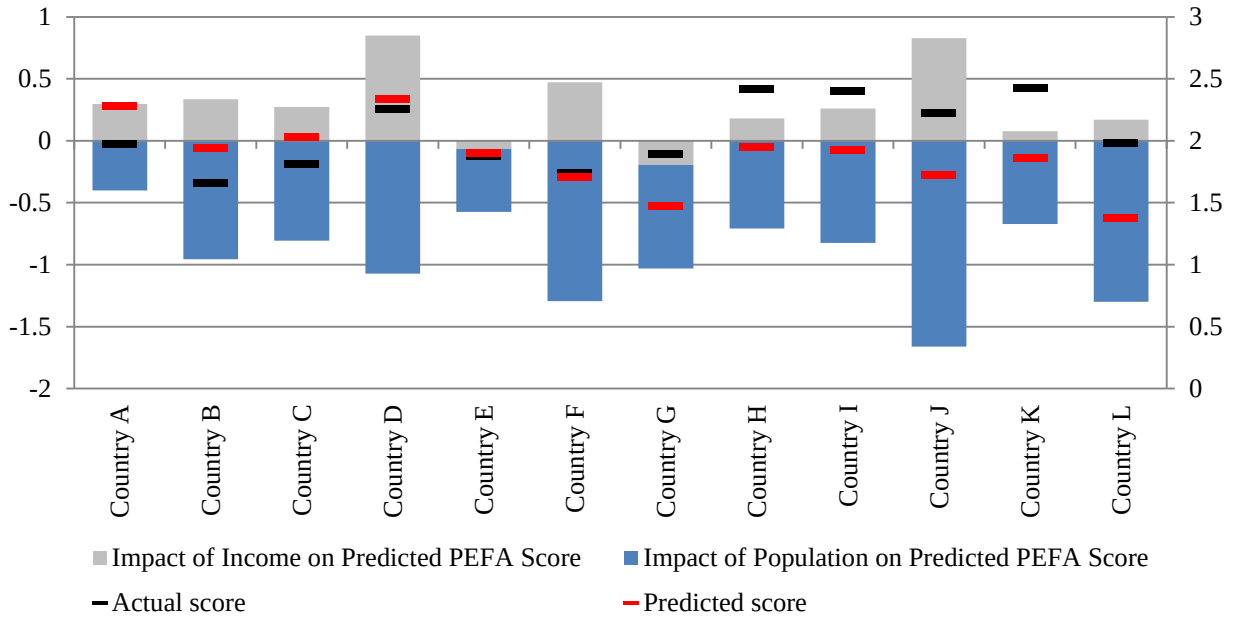
4.3 Population Size as a Determinant of PEFA Performance

Population size has also been identified as a potential determinant of PEFA assessment performance by De Renzio (2009), who – controlling for income and other relevant variables – finds that larger populations are generally associated with stronger PEFA assessment performance. This has clear relevance to our analysis, with the small PICs among the smallest countries in the world. All twelve have populations of less than one million, and ten have populations of less than 500,000.

Consistent with De Renzio (2009) we find a significantly positive relationship observed between population size and PEFA performance. The estimates from Column 3 of Table 1 indicate that a doubling of population is associated with a 0.2 point improvement in average PEFA scores. Notably, when controlling for the impacts of population size and income, the small PIC dummy approaches zero and is insignificant. In other words, the relatively low incomes and small population sizes of small PICs are sufficient to explain their lower average PEFA scores. These results remain unchanged when including regional fixed effects in Column 4[3].

Figure 5 shows the average PEFA scores across all dimensions for small PICs compared to their scores when adjusted to take account of income and population size. The expected impact of population and income on scores is shown, based on observed relationships between these variables and PEFA scores for all countries in the sample.

Figure 5: Small PIC Scores Adjusted for Income and Population Size



[3] Small PIC dummy is omitted due to inclusion for regional dummies, and is not significant if included.

From Figure 5, we can see that most small PICs perform better than or at similar levels to what would be expected, when their income and population size is taken into account. The extent to which these countries are disadvantaged by population size is clear, with a population "penalty" of more than one half of a score (the difference between a "C" and a "C+") experienced by all countries except Fiji.[4]

Results presented above test for a constant relationship between population size and PEFA scores. The assumption of linearity of the marginal effect of population on PEFA scores may not be valid if only very small countries are disadvantaged or if only very large countries are advantaged. This indeed seems to be the case when we rerun our regression on subsamples of countries of different sizes. Results are presented in Table 2. We find that the population effect is significant and largest for countries with populations of less than 500,000, whereas it is insignificant and very close to zero for larger countries.

The positive relationship between population and PEFA performance seems to be confined to small countries. We return to potential interpretations of these results in subsequent sections of this paper.

Table 2: Testing linearity of population effect on PEFA score

	(1) Pop < 500k	(2) 500k > pop > 10m	(3) Pop > 10m	(4) Full sample
LGDPit (log income)	0.487*	0.575***	0.526***	0.391***
	(0.081)	*(0.000)*	*(0.007)*	*(0.000)*
LPOPit (log population)	0.371***	0.026	0.028	
	(0.006)	*(0.784)*	*(0.769)*	
SMALLi				-0.803***
				(0.000)
MEDIUMi				-0.478***
				(0.001)
Time trend	YES	YES	YES	YES
Regional dummies	NO	NO	NO	NO
Number of observations	1,799	4,450	4,282	10,948
Number of clusters	26	69	61	162
Adj. R-Squared	0.010	0.053	0.023	0.022

Notes: Dependent variable is discrete PEFA score. All regressions use Tobit estimation.
p-values in parentheses are based on robust and clustered standard errors. *** p<0.01; * p<0.1.
Constant is included but not reported.

[4] Influence of population and GDP are as compared to the median country in the sample.

4.4 Aid and Migration as Determinants of PEFA Performance

While population size seems to explain much of the relatively poor performance of small PICs, we now consider whether other factors may also play a role.

Associations between aid flows and PEFA performance have been identified in recent literature. De Renzio (2009) finds a simple correlation of around 10%, with higher aid flows being associated with higher PEFA scores. De Renzio (2009) suggests two possible explanations for this result. Firstly, countries with access to high levels of aid, especially in the form of budget support, are also likely to be provided with extensive technical assistance to support PFM system improvement, which might elevate scores. In this case, aid flows would be driving higher PEFA performance. An alternative explanation is that countries with better PFM systems may be entrusted with higher levels of aid by donors. In this case, the observed correlation between higher aid and stronger PEFA scores would reflect higher PEFA scores driving higher levels of aid.

Any relationship between aid and PEFA performance is important to this analysis. With the exception of Fiji, all small PICs receive high levels of aid relative to their size. In Solomon Islands and Kiribati for instance, aid flows are equivalent to 34 and 49 percent of GDP, respectively.

Column 2 of Table 3 indicates that aid is positively linked to improvements in average PEFA scores, consistent with De Renzio's posited relationship.

But such a positive link between aid and outcomes is contrary to the established aid effectiveness literature (for example Rajan and Submaranian 2008) which emphasizes the negative selection bias present in aid flow. Aid flows tends to flow to where it is needed most, i.e. where outcomes are poor, and that will lead to negative link between aid and outcomes on a cross-country basis. While neither possible explanation for the positive link between aid flows and PEFA performance can be ruled out, our results interpreted through the lens of previous work demonstrating negative selection bias might suggest that donors tend to invest in the strengthening of PFM systems when substantial aid programs are being delivered. In so doing, they appear to drive some improvements in PEFA scores.

Large migration flows might impact on PEFA performance through at least two channels. Firstly, extensive outward flows of skilled labor might represent a "brain drain" that impedes effective government functioning and PFM performance. Skilled individuals with access to more lucrative offshore work opportunities leave the public sector, opening up skill gaps that worsen performance on PEFA assessments. Alternatively, labor mobility may lead to improvements in PFM performance if it facilitates the acquisition of skills and experiences and broader knowledge transfers. Those countries with more mobile populations might have access to skills acquired elsewhere that allow them to perform more strongly on PEFA assessments.

Relationships between migration flows and PEFA performance are particularly relevant to any analysis of small PICs. Several small PICs, especially in Polynesia, are notable in the extent to which their populations are mobile, with extensive outwards and returning migration. For both Samoa and Tonga, the migrant population living in other countries but identifying as Samoan and Tongan respectively, exceeds the domestic population. Remittances from mobile workforces provide a large share of GDP for Tonga, Samoa, Tuvalu, and Kiribati. However, we find no evidence of a relationship between migration flows and PEFA performance within our overall sample or for small countries. For larger countries, this result may not be surprising given the very small number of migrants relevant to the population. For smaller countries, one possible explanation of the result is that the two possible effects counteract one another, with countries experiencing "brain drain" often also able to benefit from the knowledge and skill acquisition associated with labor mobility.

Table 3: Additional Possible Determinants of PEFA Performance

	(1)	(2)	(3)
LGDPit (log income)	0.506***	0.633***	0.681***
	(0.000)	*(0.000)*	*(0.000)*
LPOPit (log population)	0.158***	0.234***	0.204***
	(0.000)	*(0.000)*	*(0.000)*
Migrationit	0.012		0.010
	(0.200)		*(0.317)*
AIDit		0.917**	1.090*
		(0.048)	*(0.065)*
Time trend	YES	YES	YES
Regional dummies	YES	YES	YES
Number of observations	9,671	10,259	9,249
Number of clusters	144	151	137
Adj. R-Squared	0.027	0.025	0.027

Notes: Dependent variable is discrete PEFA score. All regressions use Tobit estimation.
p-values in parentheses are based on robust and clustered standard errors. ***p<0.01; **p<0.05; *p<0.1.
Constant is included but not reported.

5. Capacity Constraints on PEFA Performance of Small Pacific Island Countries

The preceding analysis suggests that the PEFA performance of small PICs is generally consistent with expectations, given their income level and population size. There are intuitive causal links between higher incomes and improved PFM performance. Richer countries are better able to pay for skilled staff, technical advisors, and up-to-date systems than poorer countries. The governance and political economy factors supporting the effective functioning of institutions for

efficient economy-wide resource allocation are also likely to support institutions that deliver effective public resource use.

The nature of the link between small populations and lower PEFA scores, however, is less intuitive. One might expect small countries to perform more strongly than larger countries in PEFA performance given greater ease in communications and monitoring with a smaller number of parties involved in various aspects of public administration, the smaller total numbers of public entities, and – typically – the lesser role in service delivery played by sub-national governments.

In this section, we examine how small population size might impact negatively on PEFA scores. We test the hypothesis that governments of small countries suffer from endemic capacity constraints that undermine successful execution of various PFM functions assessed within the PEFA framework. There is a minimum fixed scale for the functions of government, with countries having to deliver certain key tasks, including those measured in the PEFA assessment, regardless of the size of their populations. Small countries may, therefore, be disadvantaged in two ways. Firstly, because the absolute number of staff within the public service is smaller in small countries, governments are unable to access sufficient staff to undertake even basic functions. There may simply not be enough people to complete all necessary tasks, especially if these tasks involve a large amount of work being done across a large number of agencies. Secondly, there may be an overwhelming shortage of particular technical skills and capacities required to undertake the specialized functions measured within the PEFA framework, including specialized forecasting, accounting, planning, or IT roles. Within a small public service, opportunities for specialization are very limited. There are not enough staff for tight specialization in roles, and not enough work to justify the appointment and retention of those with very specialized skills. With small populations, the number of people with any given set of specialized skills is inevitably lower, leading to problems of recruitment for specialized positions. Overall, it seems reasonable to expect that small countries face all of the same problems as large countries in carrying out PFM functions, but also face additional capacity constraints that are not experienced so severely in larger country contexts.

This hypothesis is consistent with the authors' experiences working on PFM reform in small Pacific Island contexts and the wider literature on institutional development in small country settings. To provide a vivid example, the Kiribati Ministry of Finance and Economic Development has a total professional staff of around 120 and its responsibilities encompass treasury and payments, statistics, aid management and coordination, internal audit, revenue and customs administration and policy, SOE monitoring and reform, and implementation of several large donor infrastructure projects. The investment unit, responsible for management of trust funds and the SOE portfolio, currently consists of two officers. Staff are constantly overloaded and frequently drawn away from core tasks by workshops, consultations, and training sessions. Those acquiring formal qualifications or marketable skills tend to seek employment in better-remunerated roles within donor agencies or overseas. The consequence is a perpetual shortage of

both people and skills to undertake basic, everyday PFM functions, and a complete absence of individuals with the skills to complete technically specialized tasks – a situation replicated to some extent in many Pacific states. This hypothesis that PFM performance is impeded by capacity constraints to an unusual extent in Pacific states is consistent with the content of PEFA assessment reports from small PICs that often refer to staff shortages, the inability of governments to fill vacant posts, and the shortage or absence of individuals with requisite skills, training, and experience for key tasks.

We test the hypothesis that PEFA performance of small countries, including PICs, is due to capacity-constraints associated with small populations by examining relationships between population size and country scores against various groupings of PEFA dimensions.

5.1 PEFA Dimension Groupings

Previous quantitative analysis of patterns in PEFA performance at a global or regional level has often involved the classification of PEFA dimensions into various groupings. These grouping exercises have been carried out to test whether countries perform better or worse in undertaking PFM functions of different types. Through observing differences in performances against different types of function, previous work has reached various conclusions regarding factors driving overall stronger or poorer performance, typically related to the political-economy of PFM reforms.

Consistent with previous quantitative approaches to analysis of PEFA scores, we created a new index by which PEFA dimensions are grouped according to the capacity requirements for achieving a good score against that dimension. Categories were designed to reflect the amount of technical learning (formal or on-the-job) that individuals with direct responsibility for the task would require to complete successfully the function to the specified standard. Indicators were assigned to various categories based on the joint assessment of three World Bank economists working on Public Financial Management issues in PICs. The resulting classifications were then independently reviewed by a financial management specialist with broad global experience. A large degree of consensus was achieved, with only four of the 66 rated dimensions being reassigned to a different category during the review process. This categorization is intended to allow direct assessment of the impact of capacity on PEFA performance. Each PEFA dimension was assigned to one of two categories, shown in Table 4 below.

Table 4: Categorization of PEFA Dimensions by Capacity

Category	Explanation
No or some specialized capacity requirements	Achieving a C score against these indicators would require some very low or basic capacity, that could be gained with on-the-job training
Highly specialized capacity requirements.	Achieving a C score against these indicators would require specialized skills gained through tertiary education, including, for example, training in accounting and auditing, database design and maintenance, and tax assessment.

Countries experiencing a shortage of technical staff should clearly perform worse against dimensions associated with more specialized technical requirements relative to countries where capacity constraints are less severe.

To further test our hypothesis regarding the impact of capacity constraints in small countries, we combine our new index with Andrews' (2011) concentrated/de-concentrated distinction between PEFA dimensions (see Table 5). This binary distinction allows comparative analysis of scores against dimensions that involve actions by central agencies as opposed to dimension where scores can only be improved through the efforts of a larger and more diffuse group of actors.

If capacity is the binding constraint on the functionality of PFM systems, scores against capacity-intensive de-concentrated dimensions might be expected to lag scores against concentrated dimensions. A small number of appropriately qualified public servants can successfully undertake centralized PFM functions within a Ministry of Finance – where the most capable public servants are typically concentrated. International technical assistance to PFM reform, which is also typically concentrated within Ministries of Finance, would also be expected to more effectively plug capacity gaps in undertaking this kind of function. In contrast, performance against capacity-intensive de-concentrated dimensions which rely on a large number of actors is likely to be undermined to a greater extent by a shortage of such staff. International technical assistance to central agencies is unlikely to be effective in plugging capacity gaps in line ministries or at the sub-national government level.

Table 5: Categorization of PEFA Scores by Concentration

Category	Explanation
Concentrated	Indicators relating to parts of the PFM system that can be managed by a small number of centralized agencies – such as ministries of finance (e.g. preparation of multi-year fiscal forecasts)
De-concentrated	Indicators that involve a wider and more diffuse range of players – such as line ministries or sub-national governments (e.g. frequency and transparency of adjustments to budgets by line ministries)

The results of regression analysis testing the impact of income and population size, and controlling for aid flows, are shown in Table 6 and Table 7 in the following pages.

5.2 Population and Capacity

Results regarding the impacts of population size lend weight to our hypothesis that capacity shortages exert a serious constraint on PEFA performance in small PICs. Columns 1 and 2 of Table 6 show that, across our global sample, smaller population size is associated with lower PEFA scores against both high and low capacity dimensions. But, consistent with our hypothesis, the magnitude of the impact is approximately one third greater than in areas where less technical

capacity is required. We test the difference in the strength of the population effect between high and low capacity dimensions in Column 3, where it is shown to be significant. In Columns 4 and 5, we replace the population variable with a set of dummies for small, medium and large countries (the latter are the omitted dummy). The difference in PEFA performance between high and low capacity indicators is most pronounced for small countries.

Table 6: Analysis of PEFA Dimensions by Capacity Requirement

	(1) Low capacity	(2) High capacity	(3) All	(4) Low capacity	(5) High capacity
LGDPit (log income)	0.612***	0.697***	0.632***	0.404***	0.463***
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.001)*
LPOPit (log population)	0.218***	0.285***	0.223***		
	(0.000)	*(0.000)*	*(0.000)*		
AIDit	0.785*	1.318*	0.911**	-0.413	-0.136
	(0.072)	*(0.063)*	*(0.050)*	*(0.365)*	*(0.797)*
HIGHCAPj			-0.837**		
			(0.020)		
HIGHCAPj*LPOPit			0.044*		
			(0.055)		
SMALLi				-0.777***	-1.240***
				(0.003)	*(0.000)*
MEDIUMi				-0.478***	-0.827***
				(0.001)	*(0.000)*
Time trend	YES	YES	YES	YES	YES
Regional dummies	YES	YES	YES	YES	YES
Number of observations	7,862	2,397	10,259	7,862	2,397
Number of clusters	151	151	151	151	151
Adj. R-Squared	0.023	0.031	0.025	0.021	0.031

Notes: Dependent variable is discrete PEFA score. All regressions use Tobit estimation.
*** p<0.01, ** p<0.05, * p<0.1. p-values in parentheses are based on robust and clustered standard errors. Constant is included but not reported.

These results suggest that smaller countries perform more poorly than other countries in their average PEFA scores partly because of a particularly strong "population penalty" against PEFA dimension that involve the application of technical capacity. While smaller populations seem to constrain performance against all PEFA dimensions, shortages of trained and skilled staff bite against those dimensions requiring the application of higher levels of capacity.

We find this pattern of particularly poor performance against dimensions involving higher levels of capacity is reproduced for small PICs. The lagging performance against high capacity dimensions for PICs on average is more pronounced than that for the average low or lower-

middle income country, and substantially more pronounced than the global average. Two-thirds of small PICs perform more poorly where higher levels of capacity are required, and in some countries the difference is pronounced, equivalent of a half grade lag against high capacity indicators (see Figure 6).

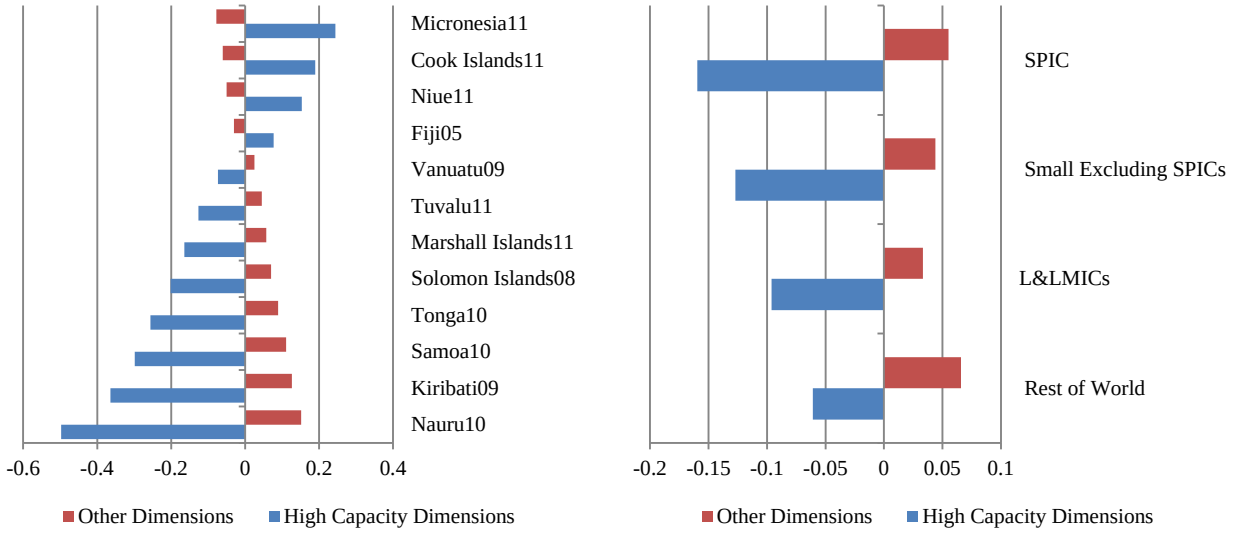
Box 2: Variance in Capacity Constraints Across PICs

Contrary to our expectations regarding the impact of capacity in small-island states, Fiji, Cook Islands, Niue, and FSM (Federated States of Micronesia) perform more strongly on high capacity dimensions than they do on lower-capacity dimensions.

Fiji is the largest of the small PICs, and with a population of close to one million, may be better placed to fill capacity constraints in the public sector. Cook Islands and Niue, on the other hand, are "special cases" to the extent that they are in "free association" with New Zealand. Both Cook Islanders and Niueans are New Zealand citizens and have free access to the New Zealand labor market and education opportunities. The extent to which populations take advantage of the opportunities these relationships offer is reflected by the fact that there are far more Cook Islanders and Niueans living in New Zealand than there are living in Cook Islands and Niue, respectively. The extent of labor mobility may ameliorate some capacity shortages by both allowing higher levels of capacity among local staff through external job and education opportunities and extending the pool of potential employees to much-larger off-shore populations. Close constitutional links to New Zealand also provide benefits in terms of larger per-capita aid flows and intensive technical assistance in various in-line or advisory roles, which may assist in closing any capacity gap.

FSM, through the Compact of Free Association, benefits from similar labor mobility and aid links to the United States. Outward migration has been far more limited and more recent than that experienced by Cook Islands and Niue. But aid flows and other forms of technical assistance from the United States might also buoy PEFA performance against higher-capacity dimensions. Marshall Islands, however, also benefits from Compact arrangements, but does not exhibit the same pattern of higher performance against high-capacity indicators. FSM's slightly larger population, decentralized federal structure, and larger stock of citizens working abroad may account for this difference.

Figure 6: Difference from Average Scores by Capacity Classifications



5.3 Population and Concentration

Our results show that population has a differential impact between concentrated and de-concentrated dimensions where specialized technical capacities are required. Dimensions that are both de-concentrated and require the application of technical capacity are a challenge for all countries, with results against these dimensions lagging for all country groups, as seen in Columns 2 and 4 of Table 7. But this difference is particularly pronounced for countries with populations of less than 500,000. The lag in performance against dimensions that are both de-concentrated and high capacity is significantly larger for small countries than in countries with larger populations.

These results further support our contention that population impacts on PEFA scores through the imposition of capacity constraints. Our results suggest that performance is most significantly constrained by a shortage of staff with technical skills disbursed across line agencies. Small countries may be more able to bring appropriate skills to bear in dealing with technically difficult "concentrated" dimension, either through the concentration of limited available local capacity within Finance Ministries or through accessing international technical assistance. However, they face significant disadvantages relative to larger counties when needing to access sufficient capacity to deal with technical functions carried out at the level of service delivery agencies. The small pool of qualified and skilled public servants typically concentrated within central agencies and with good access to international technical assistance can better handle the technical tasks required of them. But outside of central agencies, a lack of skilled and qualified staff undermines successful implementation of processes assessed through the PEFA framework.

Table 7: Analysis of PEFA Dimensions by Capacity Requirement and Concentration

	(1) Concentrated	(2) Deconcentrated	(3) Concentrated	(4) Deconcentrated
LGDPit (log income)	0.573***	0.687***	0.575***	0.686***
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*
LPOPit (log populatio)	0.220***	0.225***	0.267***	0.274***
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*
AIDit	1.004**	0.799	1.163**	0.969*
	(0.031)	*(0.128)*	*(0.019)*	*(0.079)*
HIGHCAPj	-0.623	-1.134***	0.116	-0.383***
	(0.268)	*(0.002)*	*(0.114)*	*(0.000)*
HIGHCAPj*LPOPit	0.048	0.046*		
	(0.185)	*(0.057)*		
SMALLit			0.298	0.363
			(0.221)	*(0.251)*
HIGHCAPj*SMALLit			-0.074	-0.275**
			(0.730)	*(0.037)*
Time trend	YES	YES	YES	YES
Regional dummies	YES	YES	YES	YES
Number of observations	4,921	5,338	4,921	5,338
Number of clusters	151	151	151	145
Adj. R-Squared	0.020	0.033	0.020	0.033

Notes: Dependent variable is discrete PEFA score. All regressions use Tobit estimation.
*** p<0.01, ** p<0.05, * p<0.1. p-values in parentheses are based on robust and clustered standard errors. Constant is included but not reported.

6. Conclusions and Policy Implications

Small PICs perform poorly in PEFA assessments relative to most relevant comparators. The poor performance of small PICs is driven by low scores against certain common PFM functions, including procurement, internal audit, and strategic budgeting, across all small PICs. The relatively low incomes of small PICs are an inadequate explanation for poor performance, with small PICs performing worse than other countries at similar levels of income. Population size, however, also exerts a significant influence on PEFA performance. Smaller countries face a significant size penalty in PEFA scores, taking account of differences in income. Overall, the weaker average performance of Small PICs can be entirely explained by their population sizes and levels of income.

A range of evidence suggests that population impacts on PEFA performance of small countries through the imposition of capacity constraints. Population size matters most for countries that have populations of less than 500,000. Population ceases to have any significant impact for countries with populations of between 500,000 and ten million. Because there are fixed capacity requirements associated with the basic functioning of a PFM system in any country, population impacts are most pronounced for very small countries which have inadequate skills among their

populations or positions in their public services to meet these requirements. Population impacts come to matter less for countries that have populations adequate to provide a critical mass of required skills and resources to the public sector.

While smaller countries perform worse across all PEFA dimensions, lags in performance are most pronounced against dimensions requiring higher levels of technical capacity. While Pacific Island country public services are short of people at all skill levels, they are especially constrained by an absence of qualified and skilled staff to carry out specialized PFM functions. Population size exerts the most significant constraint on PEFA performance through dimensions that are both high capacity and de-concentrated. Available technical capacity in small countries is often concentrated within central ministries, which also typically benefit from external technical assistance to PFM processes and reforms. Performance against PEFA dimensions lags further for functions undertaken beyond central agencies, where available capacity is further constrained. Capacity constraints associated with smallness seem to outweigh any potential benefits of smallness, such as a smaller number of administrative units, less spending through sub-national governments, or easier communication.

The results presented in this paper are, to some extent, unsurprising. Most advisors with experience in small PICs will be very aware of the constraints imposed by skill-shortages, along with general staff shortages and turnover of key individuals. But they remain unconfirmed within the literature and – too often – are not reflected in policy approaches to PFM reform in small-country settings.

As donors become increasingly concerned with the strength of PFM systems in Small PICs, pressure is building on governments to improve PEFA scores. The primary lesson from our analysis is that the scope and ambition of PFM reform programs needs to be appropriately calibrated to the context. At least some constraints to achieving good PEFA scores are inherent and unavoidable. Small PICs are among the smallest countries in the world. Many are also relatively poor. It is unrealistic to expect the implementation of PFM systems employed in much larger countries at similar levels of income, let alone the achievement of global “good practice” assessed through the PEFA framework. PFM reform plans that attempt to address all areas of “poor performance” identified through reference to PEFA scores or achieve blueprint international good practice may actually exacerbate weaknesses in PFM systems by spreading limited capacity too thinly across functions and agencies.

If high scores across the board are not an option, a key challenge for those designing and implementing PFM reform programs is to ensure that scarce capacity is targeted to areas with highest return. If a country has, and is likely to continue to have, only a small pool of qualified accountants, lawyers, IT specialists, or individuals experienced in financial planning, then it is important that these individuals are used in areas where they can have the greatest impact, and sufficiently concentrated in these areas to achieve results. PFM reform efforts should therefore be targeted towards weaknesses in PFM systems that are exerting the greatest constraint on the

achievement of broader development objectives. The PEFA framework can provide a valuable tool in identifying strengths and weaknesses in PFM systems and considering potential implications of these strengths and weaknesses for broader macroeconomic management or service delivery challenges. A priority for future research might include the explicit mapping of linkages between specific PEFA dimensions and progress against particular broader development problems. For now, PFM reform practitioners working in small country settings could usefully prepare PFM reform plans based on an explicit consideration of capacity constraints and clear linkages between the prioritization of PFM reform objectives and broader national development goals. These recommendations are entirely consistent with existing regional guidance on planning of PFM reform, such as PFTAC's PFM Roadmap.

Our findings suggest that this careful planning and prioritization of PFM reforms in small PIC contexts could usefully address two issues that are of particular relevance for small, lower-income states. Firstly, the balance in external assistance between central and line agencies needs to be carefully examined. As others have noted, donors tend to focus technical assistance support to PFM reform within central agencies (Andrews 2011). But the capacity constraints faced by small PICs bite most severely for technically demanding PFM functions that are carried out at the level of line agencies. The areas in which small PICs perform particularly weakly are unlikely to be well-addressed by technical assistance or PFM reform programs centered on Ministries of Finance. Particular emphasis on technical assistance and capacity building in line agencies may be required in these countries, given the apparently severe and binding constraints to PFM performance outside of central agencies.

Secondly, explicit recognition of the likely ongoing nature of capacity constraints may be required. Small PICs lag other countries in performance against dimensions that require specialist capacities. These capacities are not easily built or retained. Capacity gaps reflect fundamental features of small PIC economies – paucity of employment and education opportunities in specialized fields – that are unlikely to change over the medium- or even long-term. Options may need to be considered for accessing international capacity on a continuing basis, if capacity constraints to achieving a priority policy objective in public financial management are unlikely to be surmountable with local resources alone. While some capacities can undoubtedly be developed locally, others are likely to need to be 'imported' on a continuous basis over the long term. Future research could assist in identification of functions that could be most effectively outsourced to regional or global public or private sector agencies, given the typical extent and nature of capacity shortfalls in small PICs.

References

Andrews, M. (2009), 'Isomorphism and the Limits to African Public Financial Management Reform', RWP09-012, Kennedy School, Harvard.

Andrews, M. (2010a), 'Good government means different things in different countries', Governance: an International Journal of Policy, Administration and Institutions, 23(1): 7-35.

Andrews, M. (2010b), 'What would an ideal Public Financial Management system look like?' In Shah, A. (ed.) Budgeting and Budgetary Institutions, World Bank, 2008.

Andrews, M. (2010c), 'How far have Public Financial management Reforms Come in Africa?' Harvard Kennedy School Working Paper, RWP 10-018, May 2010.

Armstrong, H. W. and R. Read (2003), 'The Determinants of Economic Growth in Small States', Round Table, 368:99-124.

Baker, R. (ed.) (1992a), Public Administration in Small and Island States, West Hartford, Kumarian Press, Connecticut.

Baker, R. (1992b), 'Scale and Administrative Performance: The Governance of Small States and Microstates', in R. Baker (ed.), Public Administration in Small and Island States, West Hartford, Connecticut: Kumarian Press, 5-25.

Bates, R. (2006), 'Institutions and Development', Journal of African Economies, 15(1):10-61.

Bertram, G, and R Watters (1985), 'The MIRAB economy in South Pacific microstates', Pacific Viewpoint 26(3), pp 497–512.

Bertram, G. (2003), 'On the Convergence of Small Island Economies with Their Metropolitan Patrons', World Development, 32(2):343-64.

Diamond, J. (2010) 'Towards Good Practice Guidelines in the Sequencing of PFM Reforms'. Unpublished discussion paper.

Duncan, R. (ed) (2011), The Political Economy of Economic Reform in Pacific Island Countries, Asian Development Bank, Manila.

Easterly, W. and A. Kraay (2000), 'Small States. Small Problems? Income, Growth and Volatility in Small States', World Development, 28(11):2013-27.

Elbadawi, I. and J. Randa (2003), 'Assessing the development impact of CDF-like experiences'. World Bank Discussion Paper.

Hedger, E., G. Baudienville and P. Krause (2009), 'Public Financial Management in Fragile States: Review of international knowledge & concepts and three case studies on experience with PFM re-building and reforms', Summary Report, ODI, London, for PRMPS, The World Bank.

Gibson, J. and K.L. Nero (2007), 'Are Pacific Island Economies Growth Failures? Geo-Political Assessments and Perspectives', Pasifika Series Working Papers, Department of Economics, University of Waikato.

Larmour, P. (2005), Foreign Flowers: Institutional Transfer and Good Governance in the Pacific Islands, University of Hawaii Press, Hawaii

Pacific Financial Technical Advisory Center (2010), Public Financial Management Roadmap for Forum Island Countries, PFTAC, Suva.

PEFA Secretariat (2005), Public Financial Management Performance Measurement Framework, PEFA Secretariat, www.pefa.org

PEFA Secretariat (2009), Issues in Comparison and Aggregation of PEFA Assessment Results Over Time and Across Countries, www.pefa.org

Porter, D., M. Andrews, J. Turkewitz, and C. Wescott (2010), 'Managing Public Finance and Procurement in Fragile and Conflicted Settings', WDR Background Paper, World Bank, http://wdr2011.wordlbank.org/procurement

Quist, R. (2009), 'The PEFA Reform Sequence Model Framework', Draft Paper submitted to AIDCO, European Commission, June 2009.

Quist, R. (2010), 'On the PEFA Reform Sequencing Model Framework', Presentation prepared for the EC/IMF Seminar on Sequencing of PFM Reforms, June 9, 2010.

Read, R., (2004), 'The Implications of Increasing Globalization and Regionalism for the Economic Growth of Small Island States', World Development, 32(2):365-79.

Renzio, P. (2009), 'Taking stock: What do PEFA assessments tell us about PFM performance across countries?' ODI Working Paper 302, May 2009.

Rigobon, R. and Thomas M. Stoker (2006), 'Testing for Bias from Censored Regressions', MIT working paper.

Srinivasan, T.N. (1986), 'The Costs and Benefits of Being a Small Remote Island Landlocked or Ministate Economy', World Bank Research Observer, 1(2):205-18.

Streeten, P. (1993), 'The Special Problems of Small Countries', World Development, 21(2):197-202.

Tommasi, D. (2009), 'Strengthening Public Expenditure Management in Developing Countries - Sequencing Issues'. http://capacity4dev.ec.europa.eu/strengthening-public-expenditure-management-developing-countries-sequencing-issues.

Wescott, C.G. (2008), 'World Bank Support for Public Financial Management: Conceptual Roots and Evidence of Impact-- Background Paper to Public Sector Reform: What Works and Why? An IEG Evaluation of World Bank Support'. IEG Working Paper 2008. Washington, D.C.: Independent Evaluation Group, World Bank.

Wettenhall, R. (1992), 'Small States: Some Machinery-of-Government Considerations.' In R. Baker (ed.). Public Administration in Small and Island States. West Hartford, Connecticut: Kumarian Press, 49-68.

Winters, L.A. and Pedro M. G. Martins (2004), 'When Comparative Advantage Is Not Enough: Business Costs in Small Remote Economies', Centre for Economic Performance, London School of Economics, London.

World Bank (2008), Public Sector Support: What works and why? And IEG Evaluation of World bank support, World Bank, Washington DC.

World Bank (2010), Solomon Islands growth prospects: constraints and policy priorities, World Bank Discussion Paper.
http://siteresources.worldbank.org/INTPACIFICISLANDS/Resources/SourcesofGrowthSummary.pdf

World Bank (2011), Pacific Futures, World Bank Discussion Paper.
http://devpolicy.anu.edu.au/pdf/2012/Discussion_Note_Pacific_Futures.pdf

Appendix A: Data description

$SCORE_{ijt}$: Score on PEFA indicator j for country i and time t, converted to numeric value. Source: PEFA Secretariat

$LGDP_{it}$: Log of average of constant USD 2000 GDP per capita over period t – 3 to t. Source: World Development Indicators (Sao Tome & Principe, Nauru, Niue, Cook Islands, Afghanistan from other sources)$LPOP_{it}$: Log of average population over period t – 3 to t. Source: World Development Indicators (Sao Tome & Principe, Nauru, Niue, Cook Islands, Afghanistan from other sources)

AID_{it}: Average ratio of overseas development assistance to GDP over period t – 3 to t. Source: OECD Statistics

$Migration_i$: Average net in-migration rate over period 2000 to 2005. Source: United Nations Population Division "World Population Prospects"

$DECONC_j$: True for PEFA indicators which involved a wider and more diffuse range of players. Source: Andrews (2011)

$HIGHCAP_j$: True for PEFA indicators for which achieving a C would require specialized skill. Source: Tobias and Suri (2011)

$SMALL_i$: True if a country's population is less than 500,000 over most of the sample.

$MEDIUM_i$: True if a country's population is greater than 500,000 and less than 10 million over most of the sample.

Appendix B: Summary Statistics

Variable	Number of observations	Mean	Standard deviation	Minimum	Maximum
PEFA Score	162	2.344	0.543	0.368	3.621
Population	162	25.534	97.990	0.002	1,190.846
GDP per capita	162	2,247.917	4,200.702	94.881	40.992.880

Appendix C: OLS estimates of key results

	(1) Regression 1-3	(2) Regression 1-4	(3) Regression 6-3
small PIC_i	-0.032		
	(0.123)		
LGDPit	0.252***	0.258***	0.340***
	(0.028)	(0.000)	(0.051)
LPOPit	0.103***	0.103***	0.116***
	(0.017)	(0.019)	(0.000)
AIDit			0.512*
			(0.264)
HIGHCAPj			-0.430**
			(0.187)
HIGHCAPj*LPOPit			0.022*
			(0.012)
SMALLi			
MEDIUMi			
Time trend	YES	YES	YES
Regional dummies	NO	YES	YES
Number of observations	10,948	10,948	10,259
Number of clusters	162	162	151
Adj. R-Squared	0.065	0.075	0.074

Notes: Dependent variable is discrete PEFA score. All regressions use Tobit estimation.
*** p<0.01, ** p<0.05, * p<0.1. p-values in parentheses are based on robust and clustered standard errors. Constant is included but not reported.

Appendix D: PEFA Dimensions and Coding

No.	Sub-No.	Indicator	Concentrated?	Capacity?
1		Aggregate expenditure out-turn compared to original approved budget		
2		Composition of expenditure out-turn compared to original approved budget		
3		Aggregate revenue out-turn compared to original approved budget		
4	**1**	(i) Stock of expenditure payment arrears (as a % of actual total expenditure for the corresponding fiscal year) & any recent change in the stock.		
4	**2**	(ii) Availability of data for monitoring the stock of expenditure payment arrears		
5		Classification of the budget	Concentrated	Highly specialized capacity
6		Comprehensiveness of information included in budget documentation	Concentrated	Highly specialized capacity
7	**1**	(i) The level of extra-budgetary expenditure (other than donor funded projects) which is unreported i.e. not included in fiscal reports.	Deconcentrated	Low capacity requirement
7	**2**	(ii) Income/expenditure information on donor-funded projects which is included in fiscal reports.	Deconcentrated	Highly specialized capacity
8	**1**	(i) Transparent and rules based systems in the horizontal allocation among SN governments of unconditional and conditional transfers from central government (both budgeted and actual allocations);	Concentrated	Some technical capacity
8	**2**	(ii) Timeliness of reliable information to SN governments on their allocations from central government for the coming year;	Deconcentrated	Some technical capacity
8	**3**	(iii) Extent to which consolidated fiscal data (at least on revenue and expenditure) is collected and reported for general government according to sectoral categories.	Deconcentrated	Some technical capacity
9	**1**	(i) Extent of central government monitoring of AGAs and PEs.	Deconcentrated	Highly specialized capacity
9	**2**	(ii) Extent of central government monitoring of SN government's fiscal position	Deconcentrated	Highly specialized capacity
10		Public access to key fiscal information	Concentrated	Low capacity requirement
11	**1**	(i) Existence of and adherence to a fixed budget calendar;	Concentrated	Low capacity requirement
11	**2**	(ii) Clarity/comprehensiveness of and political involvement in the guidance on the preparation of budget submissions (budget circular or equivalent);	Concentrated	Low capacity requirement
11	**3**	(iii) Timely budget approval by the legislature or similarly mandated body (within the last three years);	Deconcentrated	Low capacity requirement

No.	Sub-No.	Indicator	Concentrated?	Capacity?
12	**1**	(i) Preparation of multi -year fiscal forecasts and functional allocations	Concentrated	Highly specialized capacity
12	**2**	(ii) Scope and frequency of debt sustainability analysis	Concentrated	Highly specialized capacity
12	**3**	(iii) Existence of sector strategies with multi-year costing of recurrent and investment expenditure;	Deconcentrated	Some technical capacity
12	**4**	(iv) Linkages between investment budgets and forward expenditure estimates.	Deconcentrated	Highly specialized capacity
13	**1**	(i) Clarity and comprehensiveness of tax liabilities	Concentrated	Some technical capacity
13	**2**	(ii) Taxpayer access to information on tax liabilities and administrative procedures.	Concentrated	Low capacity requirement
13	**3**	(iii) Existence and functioning of a tax appeals mechanism.	Concentrated	Low capacity requirement
14	**1**	(i) Controls in the taxpayer registration system.	Concentrated	Low capacity requirement
14	**2**	(ii) Effectiveness of penalties for non-compliance with registration and declaration obligations	Concentrated	Some technical capacity
14	**3**	(iii) Planning and monitoring of tax audit and fraud investigation programs.	Concentrated	Highly specialized capacity
15	**1**	(i) Collection ratio for gross tax arrears, being the percentage of tax arrears at the beginning of a fiscal year, which was collected during that fiscal year (average of the last two fiscal years).	Deconcentrated	Some technical capacity
15	**2**	(ii) Effectiveness of transfer of tax collections to the Treasury by the revenue administration.	Concentrated	Some technical capacity
15	**3**	(iii) Frequency of complete accounts reconciliation between tax assessments, collections, arrears records and receipts by the Treasury.	Concentrated	Some technical capacity
16	**1**	(i) Extent to which cash flows are forecast and monitored.	Concentrated	Highly specialized capacity
16	**2**	(ii) Reliability and horizon of periodic in-year information to MDAs on ceilings for expenditure commitmentc	Deconcentrated	Highly specialized capacity
16	**3**	(iii) Frequency and transparency of adjustments to budget allocations, which are decided above the level of management of MDAs.	Deconcentrated	Some technical capacity
17	**1**	(i) Quality of debt data recording and reporting	Concentrated	Some technical capacity
17	**2**	(ii) Extent of consolidation of the government's cash balances	Concentrated	Some technical capacity
17	**3**	(iii) Systems for contracting loans and issuance of guarantees.	Concentrated	Highly specialized capacity

No.	Sub-No.	Indicator	Concentrated?	Capacity?
18	**1**	(i) Degree of integration and reconciliation between personnel records and payroll data.	Deconcentrated	Some technical capacity
18	**2**	(ii) Timeliness of changes to personnel records and the payroll	Deconcentrated	Some technical capacity
18	**3**	(iii) Internal controls of changes to personnel records and the payroll.	Deconcentrated	Some technical capacity
18	**4**	(iv) Existence of payroll audits to identify control weaknesses and/or ghost workers.	Deconcentrated	Highly specialized capacity
19	**1**	(i) Evidence on the use of open competition for award of contracts	Deconcentrated	Highly specialized capacity
19	**2**	(ii) Extent of justificaction for use of less competitive procurement methods.	Deconcentrated	Highly specialized capacity
19	**3**	(iii) Existence and operation of a procurement complaints mechanism	Deconcentrated	Some technical capacity
20	**1**	(i) Effectiveness of expenditure commitment controls.	Deconcentrated	Some technical capacity
20	**2**	(ii) Comprehensiveness, relevance and understanding of other internal control rules/ procedures	Deconcentrated	Some technical capacity
20	**3**	(iii) Degreeof compliance with rules for processing and recording transactions	Deconcentrated	Some technical capacity
21	**1**	(i) Coverage and quality of the internal audit function.	Deconcentrated	Highly specialized capacity
21	**2**	(ii) Frequency and distribution of reports	Deconcentrated	Highly specialized capacity
21	**3**	(iii) Extent of management response to internal audit findings.	Deconcentrated	Some technical capacity
22	**1**	(i) Regularity of bank reconciliations	Concentrated	Some technical capacity
22	**2**	(ii) Regularity of reconciliation and clearance of suspense accounts and advances.	Deconcentrated	Some technical capacity
23		Availability of information on resources received by service delivery units	Deconcentrated	Some technical capacity
24	**1**	(i) Scope of reports in terms of coverage and compatibility with budget estimates	Concentrated	Some technical capacity
24	**2**	(ii) Timeliness of the issue of reports	Concentrated	Some technical capacity
24	**3**	(iii) Quality of information	Concentrated	Some technical capacity
25	**1**	(i) Completeness of the financial statements	Deconcentrated	Highly specialized capacity
25	**2**	(ii) Timeliness of submission of the financial statements	Concentrated	Highly specialized capacity

No.	Sub-No.	Indicator	Concentrated?	Capacity?
25	**3**	(iii) Accounting standards used	Concentrated	Highly specialized capacity
26	**1**	(i) Scope/nature of audit performed (incl. adherence to auditing standards).	Deconcentrated	Highly specialized capacity
26	**2**	(ii) Timeliness of submission of audit reports to legislature.	Deconcentrated	Highly specialized capacity
26	**3**	(iii) Evidence of follow up on audit recommendations.	Deconcentrated	Some technical capacity
27	**4**	(iv) Rules for in-year amendments to the budget without ex-ante approval by the legislature.	Deconcentrated	Low capacity requirement
27	**1**	(i) Scope of the legislature's scrutiny.	Deconcentrated	Low capacity requirement
27	**2**	(ii) Extent to which the legislature's procedures are well-established and respected.	Deconcentrated	Low capacity requirement
27	**3**	(iii) Adequacy of time for the legislature to provide a response to budget proposals .	Deconcentrated	Low capacity requirement
28	**1**	(i) Timeliness of examination of audit reports by the legislature (for reports received within the last three years).	Deconcentrated	Some technical capacity
28	**2**	(ii) Extent of hearings on key findings undertaken by the legislature.	Deconcentrated	Some technical capacity
28	**3**	(iii) Issuance of recommended actions by the legislature and implementation by the executive.	Deconcentrated	Some technical capacity
D1	**1**	(i) Annual deviation of actual budget support from the forecast provided by the donor agencies at least six weeks prior to the government submitting its budget proposals to the legislature (or equivalent approving body).		
D1	**2**	(ii) In-year timeliness of donor disbursements (compliance with aggregate quarterly estimates)		
D2	**2**	(ii) Frequency and coverage of reporting by donors on actual donor flows for project support.		
D2	**1**	(i) Completeness and timeliness of budget estimates by donors for project support.		
D3		Proportion of aid that is managed by use of national procedures		