Discussion Note
Work of the Technical Expert Groups:
Applying the ISPMS Criteria to Existing Indicators

*This is a draft for facilitating discussions at the meeting of the ISPMS Steering group in Bonn on August 29-30 2013. It draws on the background papers produced by the Technical Expert Groups in the areas of tax administration, procurement, PFM, and public administration and civil service. Several facts and figures in this draft are subject to further verification/confirmation.
Background

Public sector management systems are critical for the implementation of government policies and delivery of public services. It has become commonplace to recognize that "institutions matter" for economic development (Acemoglu et al., 2012; Fukuyama, 2010; North, 1990). Poor institutions have long been put forward as explaining the relative underperformance of economies and cross-country empirics confirm this relationship with studies relating better institutional quality with higher levels of per capita income and greater economic growth (Acemoglu et al., 2001; Dollar et al., 2003; Knack et al., 1995; Mauro, 1995; Rodrik et al., 2004). While the direction of causality may vary at the country level, a foundational level of institutional quality is necessary for sustained economic growth (Acemoglu et al., 2001; Rodrik et al., 2004).

Yet, evidence on which institutions matter for development effectiveness and in which contexts they are feasible is scarce. In addition, it is often difficult for development practitioners to show what progress is being made (if any) on reforms. Several recent reviews of World Bank projects have concluded that project effectiveness could be improved with a better set of metrics. While a number of indicators covering public management and governance systems have been developed in recent years, coverage across systems, countries and time is uneven. It is not clear that the public sector management areas covered are appropriate or whether when added together they would provide a picture of how the machinery of government in a country worked. Many indicators focused on legal and institutional frameworks rather than the functioning, outputs or outcomes of public management systems. In addition, there have been at least as many failed attempts to develop indicators as there have been successes, with considerable resources expended without commiserate results.

To address these issues, the World Bank, with help of several development partners is currently coordinating a multi-stakeholder effort to identify and advocate for a set of cross-national indicators of the strength of public management systems (ISPMS), in the areas of tax administration, public financial management (PFM), procurement, public administration and civil service systems (PACS), and governance systems (the systems that create the accountability environment for the public sector). The effort aims to learn from prior successful attempts to develop indicators, such as PEFA, and to work as much as possible with existing efforts (such as PEFA, TADAT and MAPS) and the existing data collection infrastructure of governments, donors and NGOs. The idea is not to duplicate any ongoing work or to develop an additional data collection initiative. Where it is found that critical indicators are not being developed or collected, the effort will be to find ways to include these within existing initiatives.

Donors and country governments can use the indicators to monitor and evaluate the results of reforms (e.g. by setting and tracking progress against targets) and international commitments to strengthen country systems, telling high-level stories about progress. The indicators can also serve as broad pointers to problem areas where reforms are potentially needed. In addition, the indicators can be used by researchers to build a larger body of evidence as to how governance and public management systems contribute to development outcomes, including tracking global trends in PSM reform and performance across countries and over time. The ISPMS are not intended for carrying out country level diagnostics and are not expected to capture country-specific characteristics of public management systems.

To identify a set of suitable indicators the initiative has proposed a set of four “utility” criteria and one “feasibility” criterion (see Table 1).
Table 1. ISPMS criteria

<table>
<thead>
<tr>
<th>ISPMS criterion</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>action-worthy</td>
<td>Indicator measures an aspect of PSM performance which makes a difference for development outcomes and/or is intrinsically valuable in itself</td>
</tr>
<tr>
<td>actionable</td>
<td>Indicator is specific enough to point to policy actions that can be taken to make improvements (e.g. not a composite indicator)</td>
</tr>
<tr>
<td>behavioral</td>
<td>Indicator captures the functioning or performance of institutions, rather than specific legal, organizational or institutional forms</td>
</tr>
<tr>
<td>replicable</td>
<td>Indicator is generated transparently and if the measurement was undertaken by someone else, they would reach similar conclusions</td>
</tr>
<tr>
<td>feasible</td>
<td>Indicator could be collected on a regular basis for a large number of countries¹</td>
</tr>
</tbody>
</table>

It is proposed that the ISPMS should be relatively “replicable” measures of those functions of public sector systems (“behavioral”) that can be influenced by governments and donors (“actionable”) and make a difference for development outcomes (“action-worthy”). In other words, ISPMS attempt to measure the degree to which public management systems fulfill those functions that are most likely to be “instrumental” in leading to better development outcomes, both in terms of “downstream” service delivery (e.g. paying teachers on time can improve teacher performance and student educational attainment) and “upstream” final outcomes of fiscal and institutional sustainability (e.g. MTEFs can improve fiscal discipline). Some of these functions (e.g. transparency, accountability etc.) are also intrinsically valuable – indicators measuring performance along these functions can also be considered action-worthy.

¹ Existing indicators were determined to be feasible if they were available for at least 20 countries and had been collected at least twice, but not necessarily regularly (e.g. biannually).
To address the issue of feasibility and sustainability, the first step of the initiative was to sift through available indicators that are already being collected regularly by different agencies/initiatives for a reasonable number of countries and filter a subset of indicators that meet the agreed criteria. This discussion note reviews the steps that the Technical Expert Groups (TEGs) took to apply the criteria, including the datasets that they reviewed and the strategy they pursued for defining “action-worthiness” within the subsystem. It presents the list of available indicators in each subsystem that meet the ISPMS criteria. Additional discussion notes describe (i) the challenges and trade-offs encountered when applying the criteria and (ii) potential strategies for addressing the gaps in coverage of existing indicators that were identified by the TEGs.

Overview

The ISPMS Secretariat provided each TEG (public financial management (PFM), tax administration, public administration and civil service (PACS) and procurement) with an initial list of potential indicators drawn from a review of over 20 datasets that met the feasibility criteria (i.e. had been collected at least twice for over 20 countries) and replicability criteria (had a data review/quality assurance process in place).

The Secretariat made a first attempt at applying the ISPMS action-worthiness criteria by mapping the indicators to the CPIA.²

² The Country Policy and Institutional Assessment (CPIA) assesses the quality of a country’s present policy and institutional framework. “Quality” refers to how conducive that framework is to fostering poverty reduction, sustainable growth, and the effective use of development assistance. The CPIA ratings are used in the International Development Association (IDA) allocation process and several other corporate activities. The Bank initiated country assessments in the late 1970s to help guide the allocation of IDA lending resources. The CPIA consists of a set of criteria representing the different policy and institutional
The Secretariat applied the “actionable” criteria by asking the question: “is the indicator specific enough to point to a potential action that could be taken to change its score?” As a result, all quantitative indicators were kept and composite indicators and those based on vague questions were eliminated.

The “behavioral” criteria were fairly strictly applied along the following lines:

- Indicators that measure the contents of government documents (“paper sources”) are de jure.
- Indicators that measure objective and subjective outputs and outcomes and measures of time are behavioral.
- Indicators that measure the provision and use of information are behavioral if they capture the usage of this information by citizens and other actors.

The TEGs reviewed the results of this initial stocktaking exercise, making adjustments as necessary. First, they looked to see if any potential datasets were missing from the initial list. For example, the Tax TEG added OECD Comparative Information Systems to its review. Next, the TEGs closely reviewed the steps taken to assure the quality of the various datasets, using their judgment to determine the adequacy of these actions. The TEGs then investigated options for identifying action-worthiness beyond the CPIA, including looking at other internationally-agreed standards for the subsystem, empirical evidence, theory and intrinsic value. Finally, the TEGs reviewed the application of the actionable and behavioral criteria, making adjustments as necessary.

The work of the TEGs was led by a World Bank staff member with participation from others within the World Bank and representatives from organizations on the Steering Group. The TEG leader worked with a consultant to review the results of the initial stocktaking exercise and conduct research into action-worthiness, drafting a technical input paper for review by TEG members. The work was conducted virtually, with TEG members receiving and submitting comments on the papers electronically. Draft papers have been developed and shared with TEG members in the areas of public financial management, tax and public administration and civil service (PACS). The procurement TEG is currently drafting a paper and the information presented within this note for this subsystem is preliminary.

Public Administration and Civil Service (PACS)

Datasets

Only two datasets that collect indicators on PACS meet the feasibility criteria:

1. **Global Integrity Index (GII)**: an expert survey, covering 109 countries in total so far, and the only one among the expert surveys with repeated data collection. The exact number of data points available varies by country: data for some countries has been collected up to five times, whereas only one round has been collected for others.

2. **Wage Bill and Pay Compression (WB-PC)** dataset. Wage bill indicators are calculated from IMF and WB data, which is drawn from statistical or administrative data reported by national governments, and covers between 55 and 154 countries, up to 8 years, depending on the indicator. Wage compression data is drawn mostly from existing studies that have collected data at the national level, and only covers 35 countries.

As a result, the PACS TEG decided to expand its review to include two additional data sources that have been collected for at least 20 counties but do not yet have time series:

1. **Expert Survey of the Quality of Governance Institute (QoG-ES)**: one round of data, collected between 2008 and 2010, available for 107 countries.
2. **Evans & Rauch (ER)**, also an expert survey collected as part of a research project on the impact of bureaucratic structure on bureaucratic performance and growth, collected in 1999 for 35 countries.

**Applying the criteria**

To better specify “action-worthiness,” the PACS TEG reviewed available literature to identify theory and empirical evidence on the linkages between characteristics of public administration and civil service systems and development and service delivery outcomes. Finding very little to work with, the consultant identified four high-level objectives of PACS that can be considered intrinsically valuable: access on merit, accountability, attractiveness, affordability. The TEG first defined core concepts to be measured within each of these objectives, and then set out to identify existing indicators that that approximate these concepts.

The TEG applied the actionable, behavioral and replicable criteria as described above. The first choice was to identify indicators from the datasets with multiple data points per country. When suitable indicators were not found there, the second choice were databases where at least one round of data collection has been carried out for a significant number of countries.

---


5 Available from: [http://www.qog.pol.gu.se/data/datadownloads/qogexpertsurveydata/](http://www.qog.pol.gu.se/data/datadownloads/qogexpertsurveydata/)

6 Data were collected for 135 countries but only reported for 107 as for some countries the number of respondents was too small. It was collected in three waves, but the second and the third wave were used to expand country coverage, not collect repeat data for the same countries.

7 Available at: [http://econweb.ucsd.edu/~jrauch/research_bureaucracy.html](http://econweb.ucsd.edu/~jrauch/research_bureaucracy.html)
Table 2. PACS indicators available from existing datasets

<table>
<thead>
<tr>
<th>PACS Function/ Objective</th>
<th>Possible indicator (core concept measured)</th>
<th>Existing Indicator from cross-national databases</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1-Access on merit and competition</td>
<td>1. Civil servants are appointed and promoted based on merit rather than on (political or other) connections</td>
<td>GII - 45b. In practice, civil servants are appointed and evaluated according to professional criteria. GII - 45c In practice, civil service management actions (e.g. hiring, firing, promotions) are not based on nepotism, cronyism, or patronage.</td>
</tr>
<tr>
<td>A2-Accountability</td>
<td>2.1. Public sector employees/ civil servants apply the law in an impartial manner</td>
<td>QoG - ES Q2b) When deciding how to implement policies in individual cases, public sector employees treat some groups in society unfairly? QoG - ES Q4 How often would you say that public sector employees today act impartially when deciding how to implement a policy in an individual case? QoG - ES Q8d) Public sector employees strive to follow rules.</td>
</tr>
<tr>
<td></td>
<td>2.2. PS/ CS implement policies decided upon by political leadership, within the bounds of the law</td>
<td>QoG - ES Q8b) Public sector employees strive to implement the policies decided upon by the top political leadership. QoG - ES Q8e) Public sector employees strive to fulfill the ideology of the party/parties in government.</td>
</tr>
<tr>
<td>A3-Attractiveness</td>
<td>3.1. Desirability of public sector employment/ a public sector career</td>
<td>E&amp;R 19. Among graduates of the country’s most elite university(ies), is a public sector career considered: Codes: 1 = the best possible career option. 2 = the best possible option for those whose families are not already owners of substantial private enterprises. 3 = the best option for those who are risk averse. 4 = definitely a second best option relative to a private sector career.</td>
</tr>
<tr>
<td></td>
<td>3.2. Overall remuneration of public sector professionals/ senior management compared to the remuneration of equivalent public sector positions.</td>
<td>QoG-Es Q2) Senior officials have salaries that are comparable with the salaries of private sector managers with roughly similar training and responsibilities.</td>
</tr>
<tr>
<td></td>
<td>3.3. Wage compression ratio</td>
<td>WB-PC: Wage Compression Ratio (the ratio of the highest salary to the lowest on the central government’s main salary scale)</td>
</tr>
<tr>
<td>A4-Affordability</td>
<td>4.1. Government wage bill as % of revenue/ expenditure/ GDP, compared to similar countries</td>
<td>WB-PC: central government wages as % of expenditures WB-PC: central government wages as % of GDP WB-PC: central government wages as % of revenue WB-PC: general government wages as % of expenditures WB-PC: general government wages as % of GDP WB-PC: general government wages as % of revenue</td>
</tr>
</tbody>
</table>
**Tax Administration**

**Datasets**

An initial overview of potential ISPMS indicators identified six possible datasets that meet the feasibility criteria of the ISPMS indicators: PEFA, Global Integrity Indicators, the Institutional Profiles Database, Doing Business Indicators, Transparency International’s Global Corruption Barometer, and the OECD’s Tax Administration in OECD and selected non-OECD countries: Comparative information Series. PEFA data is available for 79 countries, although assessments are carried out on an ad-hoc basis; Global Integrity indicators have been collected at least twice for over 100 countries; the Institutional Profiles Database indicators have been collected three times for over 120 countries; Doing Business indicators are collected annually for over 180 countries; Transparency International data comes from its Global Corruption Barometer, which has been collected for over 100 countries for up to 8 rounds; and the OECD Comparative Information Series data has been collected up to 5 times, most recently for 52 OECD and selected non-OECD countries.

**Applying the criteria**

The IMF is currently leading a multi-stakeholder process to develop and build consensus on a conceptual framework and performance indicators for tax administration. This will be a critical contribution towards defining “action-worthiness” in this field. As this framework is still being developed and cross-national empirical evidence of linkages between tax administration performance and development outcomes does not exist, the TEG input paper pragmatically proposes in the meantime to judge “action-worthiness” on the basis of a results chain developed by reviewing theoretical literature from the fields of economics, political economy and political science. Specifically, indicators were considered “action-worthy” if they focused on the following key areas of tax administration that can plausibly affect development outcomes: taxpayer registration, taxpayer services, return processing, payment processing, audit, appeals, refund processing and arrears collection. In addition, indicators that contributed to “strategic-level” outcomes of reducing the tax compliance gap, improving integrity, and improving tax administration efficiency were considered “action-worthy”.

The TEG applied the actionable, behavioral and replicable criteria as described above. When applying the behavioral criterion, *Global Integrity* indicators dropped out because the sub-questions referred to inputs rather than performance outputs or outcomes. In addition, the Institutional Profiles Database indicators did not meet the actionability criterion.

---

8 Selected non-OECD countries include: Argentina, Bulgaria, China, Cyprus, India, Indonesia, Latvia, Lithuania, Malaysia, Malta, Romania, Russia, S. Arabia, Singapore, and South Africa.
Table 3. Tax indicators available from existing datasets

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Dataset</th>
<th>Indicator title</th>
</tr>
</thead>
</table>
| 1 Strategic       | Doing Business  | Total time spent to pay taxes (hours)  
Proxy for the taxpayer compliance cost |
| 2 Strategic       | Transparency International | Perception of corruption  
The extent to which citizens who have come into contact with tax officials over the past year have paid bribes. |
| 3 Strategic       | OECD CIS        | Tax administration collection costs⁹  
The total tax administration cost¹⁰ as a percent of the total revenue collections |
| 4 Registration    | PEFA (PI-14i)   | Control in the Taxpayer Registration System  
The extent to which taxpayers are registered in a complete database system with comprehensive direct linkages to other relevant government registration systems and financial sector regulations. |
| 5 Taxpayer service| PEFA (PI-13i)   | Clarity and comprehensiveness of tax liabilities  
The extent to which legislation and procedures for all major taxes are comprehensive and clear, with strictly limited discretionary powers of all government entities involved. |
| 6 Taxpayer service| PEFA (PI-13ii)  | Taxpayer access to information on tax liability  
The extent to which taxpayers have easy access to comprehensive, user friendly and up-to-date information on tax liabilities and administrative procedures for all major taxes, supplemented with active taxpayer education campaigns. |
| 7 Payment         | PEFA (PI-3)     | Aggregated revenue outturn compared to budget  
Proxy for the payment gap |
| 8 Payment         | PEFA (PI-15ii)  | Effectiveness of transfer of tax collection to the Treasury  
The frequency with which all tax revenue is transferred into accounts controlled by the Treasury (daily, weekly, monthly, etc.) |
| 9 Payment         | PEFA (PI-15iii) | Frequency of complete account reconciliation between tax assessments, collections, arrears records and receipts by the Treasury (monthly, quarterly, annually, etc.) |
| 10 Arrears        | PEFA (PI-15i)   | Collection ratio for gross tax arrears  
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). |

⁹ While this indicator is easily captured from government budget and revenue information, it may not be fully comparable across countries. The tax administrations’ structures and operational models vary and this affects their cost structure. For example, some administrations receive their tax payments through the banks, other still use their cashiers. Some tax administrations use government or MOF information technology (IT) structure and do not pay for it, while others have to fund themselves their IT area. Some countries have customs and tax administrations integrated in a single agency. Also, larger tax administrations have important economies of scales, which cannot be achieved by tax administrations in smaller countries. It is important to mention also that this indicator seems to be of little use for customs administration, unless VAT collection on imports is included in the coefficient. Even so, for Customs, this variable seems to be typically explained more by the level of an economy’s openness, rather than by true administration efforts. In developing countries, this indicator is frequently used to identify under-funded tax administrations. But, in some circumstance, it is also used to identify over-funded organizations. The way to carry out this analysis is to compare with countries of similar size and level of development.

¹⁰ Tax administration expenditures include three categories: administrative costs, salary costs and IT costs.
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>11</strong></td>
<td>Appeals</td>
<td>PEFA (PI-13iii)</td>
</tr>
<tr>
<td></td>
<td>Existence and functioning of a tax appeals mechanism</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The extent to which the tax appeals system has transparent administrative procedures with appropriate checks and balances, implemented through independent institutional structures and completely set up and effectively operating with satisfactory access and fairness, and its decisions are promptly acted upon.</td>
<td></td>
</tr>
<tr>
<td><strong>12</strong></td>
<td>Return filing</td>
<td>OECD CIS</td>
</tr>
<tr>
<td></td>
<td>Percent of e-filed returns</td>
<td></td>
</tr>
</tbody>
</table>
Procurement

The ISPMS Secretariat has identified a list of existing indicators that potentially meet the ISPMS criteria (table 4). The Procurement TEG leader together with a consultant is currently reviewing this list as well as options beyond the CPIA and MAPS for determining action-worthiness (including the existence of empirical evidence). The results of this review will be presented to all procurement TEG members in the form of an “input paper” for their review and comment.

The initial stock-taking exercise identified seven datasets that met the ISPMS feasibility criterion and had indicators addressing procurement systems: Enterprise Surveys, BEEPS, the Open Budget Survey, Global Integrity Indicators, the Methodology for Assessing Procurement Systems, and PEFA. This exercise applied the utility criteria as described in the “overview” section above.

Table 4. Procurement indicators available from existing datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Q.</th>
<th>Indicator title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enterprise Surveys</td>
<td></td>
<td>Gifts to secure public contracts (% of firms)</td>
</tr>
<tr>
<td>BEEPS</td>
<td>J.6</td>
<td>When establishments like this one do business with the government, what percent of the contract value would be typically paid in informal payments or gifts to secure the contract?</td>
</tr>
<tr>
<td>Open Budget Survey</td>
<td>97</td>
<td>During the past year, have there been credible reported instances in which the procurement of goods and services has not followed an open and competitive process in practice? (response options include: no credible reported instances, some reported instances, significant reported examples of irregularities, and the process was not open and competitive in practice)</td>
</tr>
<tr>
<td>Global Integrity Indicators (GII)</td>
<td>51c, 51j</td>
<td>In practice, the conflicts of interest regulations for public procurement officials are enforced. In practice, companies guilty of major violations of procurement regulations (i.e. bribery) are prohibited from participating in future procurement bids.</td>
</tr>
<tr>
<td>Institutional Profiles Database</td>
<td>A3040</td>
<td>Are the results of national public procurement bids predictable? (scored on a basis of 1-4)</td>
</tr>
<tr>
<td>Institutional Profiles Database</td>
<td>A3041</td>
<td>Are the results of local public procurement bids predictable? (scored on a basis of 1-4)</td>
</tr>
<tr>
<td>Methodology for Assessing Procurement Systems</td>
<td>10a</td>
<td>Appeals/complaints system: decisions are deliberated on the basis of available information, and the final decision can be reviewed and ruled upon by a body (or authority) with enforcement capacity under the law.</td>
</tr>
<tr>
<td>Methodology for Assessing Procurement Systems</td>
<td>10c</td>
<td>Fairness of the complaints system: Procedures governing the decision making process of the review body provide that decisions are a) based on info relevant to the case, b) balanced and unbiased in consideration of the relevant information, c) can be subject to higher level review, d) result in remedies that are relevant to</td>
</tr>
</tbody>
</table>

11 In addition, the overall Global Integrity indicator 51 looks at multiple de jure aspects of the procurement system: a) In law, there are regulations addressing conflicts of interest for public procurement officials; b) In law, there is mandatory professional training for public procurement officials; d) In law, there is a mechanism that monitors the assets, incomes and spending habits of public procurement officials; e) In law, major procurements require competitive bidding; f) In law, strict formal requirements limit the extent of sole sourcing; g) In law, unsuccessful bidders can instigate an official review of procurement decisions; h) In law, unsuccessful bidders can challenge procurement decisions in a court of law; i) In law, companies guilty of major violations of procurement regulations (i.e. bribery) are prohibited from participating in future procurement bids.
correcting the implementation process or procedures.

<table>
<thead>
<tr>
<th>Public Expenditure and Financial Accountability (PEFA)</th>
<th>PI-19</th>
</tr>
</thead>
</table>
| **1. Transparency, comprehensiveness and competition in the legal and regulatory framework**  
The legal and regulatory framework for procurement should: (scored based on # of criteria met)  
(i) be organized hierarchically and precedence is clearly established;  
(ii) be freely and easily accessible to the public through appropriate means;  
(iii) apply to all procurement undertaken using government funds;  
(iv) make open competitive procurement the default method of procurement and define clearly the situations in which other methods can be used and how this is to be justified;  
(v) provide for public access to all of the following procurement information: government procurement plans, bidding opportunities, contract awards, and data on resolution of procurement complaints;  
(vi) provide for an independent administrative procurement review process for handling procurement complaints by participants prior to contract signature. |
| **2. Use of competitive procurement methods**  
When contracts are awarded by methods other than open competition, they are justified in accordance with the legal requirements (in all cases, in at least 80% of cases, in at least 60% of cases, in less than 60% of cases) |
| **3. Public access to complete, reliable and timely procurement information**  
Key procurement information (government procurement plans, bidding opportunities, contract awards, and data on resolution of procurement complaints) is made available to the public through appropriate means. (scored based on types of procurement information made available and extent to which it is available for all procurement operations) |
| **4. Existence of an independent administrative procurement complaints system.**  
(scored based on # of criteria met)  
Complaints are reviewed by a body which:  
(i) is comprised of experienced professionals, familiar with the legal framework for procurement, and includes members drawn from the private sector and civil society as well as government;  
(ii) is not involved in any capacity in procurement transactions or in the process leading to contract award decisions;  
(iii) does not charge fees that prohibit access by concerned parties;  
(iv) follows processes for submission and resolution of complaints that are clearly defined and publicly available;  
(v) exercises the authority to suspend the procurement process;  
(vi) issues decisions within the timeframe specified in the rules/regulations;  
(vii) issues decisions that are binding on all parties (without precluding subsequent access to an external higher authority). |
Public Financial Management

Datasets

There is already a wealth of indicators being used to measure aspects of PFM systems. An initial overview of potential ISPMS indicators identified five possible datasets that meet the feasibility criteria of the ISPMS indicators: PEFA, Global Integrity Indicators, the Institutional Profiles Database, the Open Budget Index (OBI), and the OECD International Budget Practices and Procedures (IBPP) Database. Combined, there are well over 200 potential indicators that could be measured.

Applying the criteria

When the ISPMS utility criteria were applied, the OECD International Budget Practices and Procedures Database was excluded from the list. It is a not a replicable source of ISPMS indicators since the data are collected using country self-assessments. The Global Integrity Indicators and the Institutional Profiles Database were also dropped because they did not meet one of the following ISPMS criteria: behavioral, replicable or actionable.

Many of the OBI indicators were excluded during the initial stocktaking exercise due to the strict application of the “behavioral” criterion. The main objective of the Open Budget Survey is to review budget transparency, particularly with a view to social accountability. Many of the indicators on the Open Budget Survey are therefore related to the publication and quality of information made available to the public, but do not track its use—which is a requirement of the “behavioral” criterion. In practice, assessment of whether the public is using information in the public domain to hold government to account for the use of public resources is not easy to do. Further, this would be greatly influenced by contextual factors outside of the PFM system. Thus, there may be benefit in assessing how accessible governments make information to the public thereby influencing the likelihood that information will be used. The production of citizen’s budgets and holding public consultations on the budget are processes that governments can follow that actively promote public participation in the budgetary debate and therefore merit inclusion. Where budget information is not released on a timely basis to the public it also loses relevance. Consequently the PFM TEG decided to consider these indicators to meet the “behavioral” criteria.

To arrive at a short list of PFM indicators, including both PEFA and OBI that meet the ISPMS criteria, the following additional considerations of “action-worthiness” are applied: (i) the proposed indicators provide a big picture of how the PFM system functions and point to why its performance is inadequate; (ii) these indicators cover key functional areas of the PFM system; (iii) the data for these indicators are already being collected or embedded in available reports; and (iv) they pinpoint performance in areas that crucially matter for effective development and particularly service delivery outcomes. This iterative process relied on PFM functions as defined in the PEFA Framework which is the most widely recognized international standard for measuring the performance of PFM systems. The features of PFM systems that have the most impact on development and especially service delivery outcomes are identified based on the limited empirical evidence and in-depth country assessments examining links between public expenditure management and sector outcomes such as Public Expenditure Tracking Surveys conducted by the World Bank, as well as a theoretical literature. All together these additional criteria assist in the prioritization of action-worthy indicators. A current revision of the PEFA Framework is likely to affect the scope (beyond the central government) and coverage of PEFA indicators (potentially including the use of performance information and
capital investment management among others). Thus, the filtered PEFA indicators will need to reflect changes adopted in the revised PEFA Framework when it becomes available.

**Following the PEFA Framework, the selected PFM indicators are broken down into:**

- **System Outturn Indicators** that measure the quality of outturns from the PFM system and
- **Functional Indicators** that measure the effectiveness of specific functional elements of the PFM system. Both types of indicators are important for better understanding of the linkages along the whole results chain between functions of public financial management systems, the outputs those combined functions produce, and then service delivery outcomes (see Figure 2).

*Figure 1: Using Indicators to Understand Linkages between PFM Systems and Development Outcomes along the Results Chain*

The proposed short list of PFM indicators is provided in Table 5 (PEFA Indicators) and Table 6 (OBI Indicators). Where applicable, changes for measuring indicators are suggested, such as converting grades (A, B, C, and D as in PEFA) and binary indicators (as in OBI) into numeric indicators to make them more amendable to cross-national empirical research. Annex A describes the short-listed PEFA indicators in more detail.

**Table 5. A Short List of PEFA Indicators Compliant with ISPMS Based on Prioritization of Action-Worthiness**

<table>
<thead>
<tr>
<th>PFM Objective/Function</th>
<th>Relevant Indicators</th>
<th>Dimension to be measured</th>
<th>Conversion to numerical indicator</th>
<th>Explanation of action-worthiness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Credibility of the budget (System out-turns)</td>
<td>PI-1 Aggregate expenditure out-turn compared to original approved budget</td>
<td>The difference between actual primary expenditure and the originally budgeted primary expenditure</td>
<td>The difference between actual primary expenditure and the originally budgeted primary expenditure</td>
<td>Unaffordable budgets contribute to fiscal stability risks and disrupt service delivery</td>
</tr>
<tr>
<td></td>
<td>PI-2 Composition of expenditure out-turn compared to</td>
<td>Extent to which variance in primary expenditure composition exceeded overall</td>
<td>Variance in primary expenditure composition during the last three years</td>
<td>Changes in composition of budget disrupt service delivery</td>
</tr>
</tbody>
</table>

The relationship between function, output, and service delivery outcomes (see Figure 2)
<table>
<thead>
<tr>
<th>Policy-based budgeting</th>
<th>PI-11</th>
<th>Orderliness and participation in the annual budget process</th>
<th>(i) Existence of and adherence to a fixed budget calendar</th>
<th>Number of weeks that MDAs have for preparation of detailed budgets (after MoF issues a budget circular)</th>
<th>Creates space for budget aligned to priorities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(ii) Timely budget approval by the legislature or similarly mandated body (during the last three years)</td>
<td>Number of days after the beginning of a fiscal year that the budget is approved by the legislature</td>
<td>Delays in budget approval disrupt service delivery and investments</td>
<td></td>
</tr>
<tr>
<td>Predictability and control in budget execution</td>
<td>PI-16</td>
<td>Predictability in the availability of funds for commitment of expenditures</td>
<td>(ii) Reliability and horizon of periodic in-year information to MDAs on ceilings for expenditure commitment</td>
<td>Average number of weeks in advance that in-year information on quarterly fund authorizations are provided to MDAs</td>
<td>Affects MDAs' ability to plan service provision</td>
</tr>
<tr>
<td></td>
<td>PI-17</td>
<td>Recording and management of cash balances, debt and guarantees</td>
<td>(ii) Extent of consolidation of the government’s cash balances</td>
<td>Percentage of government cash balances consolidated on a weekly basis</td>
<td>Centralization of cash can lead to efficiencies in cost of borrowing</td>
</tr>
<tr>
<td>Accounting, recording and reporting</td>
<td>PI-22</td>
<td>Timeliness and regularity of accounts reconciliation</td>
<td>(i) Regularity of bank reconciliations (in weeks) between personnel records and payroll data</td>
<td>Frequency of reconciliation (in weeks) between personnel records and payroll data</td>
<td>Reduced risk of leakage through payroll</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(ii) Timeliness of all central bank accounts reconciliation after end of period (in weeks)</td>
<td>(i) Frequency of all central government bank accounts reconciliation</td>
<td>Frequency of all central government bank accounts reconciliation</td>
<td>Regular bank reconciliations are required to ensure reliability of reports</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(ii) Regularity of reconciliation and clearance of suspense accounts and advances</td>
<td>(i) Frequency of reconciliation of suspense accounts and advances</td>
<td>Frequency of reconciliation of suspense accounts and advances</td>
<td></td>
</tr>
</tbody>
</table>

original approved budget deviation in primary expenditure

PI-4 Stock and monitoring of expenditure payment arrears
(i) Stock of expenditure payment arrears
Stock of payment arrears as a percentage of actual total expenditure and any recent change in the stock
Arrears contribute to cost escalations and also indication of tendency for 'deferred budgeting'

(ii) Availability of data for monitoring the stock of expenditure payment arrears
Frequency of preparation of reports on the stock of expenditure payment arrears
| PI-24 Quality and timeliness of in-year budget reports | suspense accounts and advances after end of period | (i) Frequency of in-year budget reports covering both commitments and payments  
(ii) Time elapsed (in weeks) between end of period under review and production of reports | Supports government in making timely adjustments to the budget, thereby enhancing predictability. Also supports realistic expenditure forecasting. |
|---|---|---|---|
| PI-25 Quality and timeliness of annual financial statements | (i) Frequency of in-year budget reports covering both commitments and payments  
(ii) Time elapsed (in weeks) between end of period under review and production of reports | Time elapsed (in months) after the end of the fiscal year and submission of consolidated financial statements for external audit | Strengthening accountability for use of public resources |
| External scrutiny and audit | (i) Scope/nature of audit performed (incl. adherence to auditing standards).  
(ii) Timeliness of submission of audit reports to the legislature | Percentage of central government entities in total expenditure audited annually | Reduced risk of resource leakage |
| PI-26 Scope, nature and follow-up of external audit | (i) Scope/nature of audit performed (incl. adherence to auditing standards).  
(ii) Timeliness of submission of audit reports to the legislature | Time elapsed (in months) between end of the period covered and submission of audit reports to the legislature | Reduced risk of resource leakage |
| PI-27 Legislative scrutiny of the annual budget law | (i) Timeliness of examination of audit reports by the legislature  
(ii) Adequacy of time for the legislature to provide a response to budget proposals | Amount of time (in weeks) the legislature has to review executive budget proposals | Supports alignment of budget allocations with economic and social priorities |
| PI-28 Legislative scrutiny of external audit reports | (i) Timeliness of examination of audit reports by the legislature  
(ii) Adequacy of time for the legislature to provide a response to budget proposals | Time elapsed (in months) between end of fiscal year and the legislature’s completion of scrutiny of audit reports | Reduced risk of resource leakage |

The distinctive offering of the Open Budget Survey, which is largely missing from PEFA assessments, is whether PFM systems facilitate budget transparency and thus invite the participation of citizens in framing budgetary decisions. Table 6 provides OBI indicators included in the short list of PFM indicators. It is worth noting that many of the indicators identified in the Open Budget Survey can complement PEFA indicators by looking specifically at whether intermediate outputs produced and assessed in the PEFA framework are made available to the public.
## Table 6. A Short List of OBI Indicators Compliant with ISPMS Based on Prioritization of Action-Worthiness

<table>
<thead>
<tr>
<th>PFM Objective/Function</th>
<th>Relevant indicators</th>
<th>Proposed Indicator</th>
<th>Explanation of action-worthiness</th>
</tr>
</thead>
</table>
| OBI                    | Number of the 10 key budget reports that are released to the public: | - Pre-budget statement  
- Budget summary  
- Executive's budget proposal  
- Supporting documents to the executive's budget proposal  
- Enacted budget  
- Citizens budget  
- In-year reports  
- Mid-year reports  
- Year-end reports  
- Audit reports | Helps to address gap in PEFA framework on accessibility of budget information |
| OBI                   | Proposed modification: What proportion of expenditure is explained in a "citizens budget" or some non-technical presentation intended for a wide audience that describes the budget and its proposals? | OBI-61 | OBI-82 | OBI-91 | OBI-93 | OBI-104 | OBI-112 |
| OBI                   | How often does the executive release to the public in-year reports on actual expenditure (organized by administrative unit, economic classification and/or function)? | | | | | | |
| OBI                   | For in-year reports on actual expenditure released to the public by the executive, how much time typically elapses between the end of the reporting period and when the report is released? | | | | | | |
| OBI                   | Proposed modification: What is percentage of expenditure covered in mid-year review of the budget released to the public? | | | | | | |
| OBI                   | What level of detail is the focus of the explanation of the differences between the enacted levels and the actual outcome for expenditures presented in the year-end report? | | | | | | |
| OBI                   | Proposed modification: Less than a year after the completion of a fiscal year, what percentage of annual expenditures has been audited and (except for secret programs) the report(s) released to the public? | | | | | | |
Annex A. Discussion of the short-list of PEFA indicators that meet the ISPMS criteria

System Outturn Indicators

*Objective: credibility of the budget*

Where the composition of expenditure is not in line with appropriations, it indicates that services and investments have not been implemented in line with budgeted plans. PEFA indicator PI-2 measures the extent to which reallocations between administrative (functional/program) budget heads in the actual (out-turn) budget have contributed to variance in expenditure composition beyond the variance resulting from changes in the overall level of expenditure compared to the original budget.

There are a number of possible reasons why deviations at the expenditure head might arise. A lack of realism in budget forecasts may be skewed towards certain areas. If for example, a wage increase has been under-estimated, budgetary shifts would be greatest in those ministries whose budgets are dominated by wage expenditures (e.g. education and health). A lack of realism in forecasts may also lead to cash rationing, where selected cuts are made to specific heads in order to meet deficit targets. Where certain agencies introduce new programs or make unbudgeted expenditures during budget execution, deviations may also arise. Deviations can also occur due to under-spending, which is a particular issue in capital spending. Ministries whose budgets are dominated by infrastructure investments would be likely to see proportionately larger deviations in expenditure outturns as a result.

This indicator addresses changes to composition of the budget by budget head, but not by economic type. A common tendency in a number of countries is to hold back payments on public investments in order to achieve deficit targets. There may also be benefit then in monitoring deviations between major economic aggregates (e.g. salaries, pensions, capital investments) and actual outturns.

From the perspective of the recipient of funds, the deviation from budget is of primary importance, irrespective of why that deviation has arisen. To better understand how service delivery is affected by deviations in budgets at the level of expenditure head, it may be of more interest to measure the absolute deviations, rather than deviations from an adjusted budget.

Dimension (i) of PI-4 measures the stock of arrears. Building arrears can contribute to delays in budget execution as a well as cost escalations. Arrears might arise because of insufficient budget allocations at the preparation stage or point to weak expenditure controls that lead to expenditure commitments being made without the corresponding funds available. The stock of arrears is also a useful indicator for assessing whether governments are undertaking ‘deferred budgeting’ as a means to achieve aggregate deficit targets.

PI-4, dimension (ii) can be used to measure whether data on arrears is made available. It is proposed that this be amended to measure the frequency of reporting on arrears.

Functional Indicators

*Objective: policy-based budgeting*

(i) Allocating resources in line with priorities

On the budgeting side, process and information have a key bearing on the likelihood that a budget will represent policy priorities. On the process side, dimensions (i) of PI-11 measure whether the budget process creates the space for sufficient consideration of how the budget might
be linked to policy. More specifically, the existence of and adherence to a fixed budget calendar is reviewed. This indicator should be revised to capture the number of weeks that MDAs have for the preparation of detailed budgets.

(ii) 

**Timely approval of the budget**

Delays in initial disbursements due to late approval of the budget are often identified as key constraints to infrastructure investments and service delivery. PI-11, dimension (iii) measures the timeliness of budget approval by the legislature or similarly mandated body over a three year period. This should be amended to capture the number of days of delays in the last three years as measured by the time elapsed between the beginning of the fiscal year and legislative budget approval.

**Objective:** predictability and control in budget execution

(i) 

**Releasing funds in line with plans**

Dimension (ii) of PI-16 measures the horizon over which information on planned funding is shared with MDAs. It should be revised to capture the average number of weeks in advance that in-year information on quarterly fund authorizations are provided to MDAs. It would be beneficial if information is collected both for recurrent and capital expenditure, as sometimes processes may differ.

The preparation of reliable cash flow forecasts depends upon managers having reliable in-year fiscal reports on which to base their forecasts. PI-17 dimension (ii) looks at the extent of consolidation of the government’s cash balances. Where cash balances are consolidated, governments can make savings in the costs of servicing debt. This should also be incorporated but reworked to capture the percentage of cash balances consolidated on a weekly basis. Indicators on bank account reconciliation and in-year reports will also be relevant in better understanding possible reasons for unpredictable resource disbursements. These are covered in section on reporting.

(ii) 

**Expenditure controls**

Control functions can influence service delivery outcomes through reductions in wasteful spending arising from the misuse of funds (for example, the prevalence of ghost workers or inflated procurement prices). Ideally, a measure of the *de facto* effectiveness of controls in reducing wasteful spending would measure the size of expenditure leakages. To do this would require the type of micro-level survey data undertaken in Public Expenditure Tracking Surveys and so could not be done at a reasonable cost for the IPSMS indicator set. The PEFA does contain indicators though which provide assessments of the effectiveness of PFM functions designed to act as controls for limiting the size of these leakages.

**PI-18 measures the effectiveness of payroll controls in limiting risks of fund misuse in payroll management.** Dimensions (i) which captures the degree of integration and reconciliation between personnel records and payroll data is proposed to measure frequency of reconciliation between personnel records and payroll data.

**Objective:** accounting, recording and reporting

(i) 

**Reporting and Accounting**
Well-functioning PFM systems have mechanisms in place that promote both internal and external accountability for the use of public resources. These mechanisms ensure more efficient and effective use of public funds. It is the responsibility of the executive to make information available that can be used to judge how the government is using funds.

The preparation of reliable cash flow forecasts depends upon managers having reliable in-year fiscal reports on which to base their forecasts. PEFA indicator PI-22 dimensions (i) and (ii) assess the regularity of bank reconciliation and the regularity of reconciliation and clearance of suspense accounts, which is a vital control mechanism in ensuring the accuracy of fiscal reporting. This should then be incorporated but reworked to measure the frequency and timeliness of reconciliation.

PI-24 assesses the quality and timeliness of in-year budget reports. While it is difficult to quantify the quality of in-year reports, dimension (ii) measuring the timeliness of these reports is proposed to be amended to review how frequently in-year budget reports covering both commitments and payments are prepared (monthly or quarterly) and when they are issued (in weeks) after the end of the period under review.

PI-25 looks at the quality and timeliness of annual financial statements that are produced. In similar vein, while it is hard to gauge the quality of financial statements, dimension (ii) capturing the timeliness of these statements is proposed to be modified to capture the time taken between the end of the fiscal year and the production of accounts. The compilation of timely financial statements is a central element of holding the executive to account for the use of public resources. In addition, Timely submissions of financial statements are necessary if audits are to be timely and relevant.

Objective: external scrutiny and audit

(i) External Audit

Accountability for the use of resources also depends on how external oversight agencies use the information that is made available to them. PI-26 reviews the external audit function, which acts as a crucial control for reducing risk of the misuse of funds. Dimension (i) looks at the scope of the audit performed. This indicator is proposed to be adjusted to measure the percentage of total budgeted expenditures audited annually. Dimension (ii) on the timeliness of submission of audit reports to legislature is proposed to be amended to capture the number of months between the end of the fiscal year and the submission of audited accounts to the legislature.

(ii) External Scrutiny

Legislative scrutiny of the ex-ante budget is a key mechanism for ensuring that the prepared budget reflects the preferences of citizens. In order to play this role effectively, the legislature needs sufficient time and information. PI-27 dimension (iii) measures the adequacy of time for the legislature to review the budget. This is proposed to be amended to record the number of weeks between receipt of the executive’s budget proposal and the approval of the annual budget law.

Legislative scrutiny of the audit report is a key mechanism for holding the executive to account for the use of public resources. Dimension (i) is proposed to be amended to capture the actual time taken for legislative scrutiny of the audit report.