A Financial Benchmark for Child Protection

Full Benchmark proposal

16 October 2014
Acknowledgements

This paper was prepared by Mokoro Ltd. and the United Nations Children’s Fund (UNICEF). UNICEF would like to thank the authors of the paper, Alta Fölscher and Stephanie Allan. The paper benefited from the inputs of the members of the UNICEF reference group, Miranda Armstrong, Stephen Blight, Laurent Chapuis, Jingqing Chai, Matthew Cummins, Gaspar Fajth, Kendra Gregson, Peter Gross, Karin Heissler, Jean-Claude LeGrand, Smaranda Popa and Lauren Rumble for their inputs that helped shape this paper. Gratitude is also due to Armando Barrientos, from the Brookes World Poverty Institute at the University of Manchester, Bill Bell from Save the Children, Bob Baulch from the RMIT University Vietnam, Philip Goldman from Maestral, Laurie Joshua (consultant), Katya Maynzyuk and Yuriy Dzygyr from FISCO id, and Philip O’Keefe from the World Bank for sharing their experience and knowledge. Finally we would also like to thank Judith Streak Gomersall from the University of Adelaide and Sarah Holloway, PFM and Accounting Specialist for their comments on the paper.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS II

TABLE OF CONTENTS III

ACRONYMS IV

INTRODUCTION 1

SECTION 1: PROPOSALS FOR THE FINANCIAL BENCHMARK 1

Section 1.1 General design parameters 1

Section 1.2 What child protection expenditures will be included? 3
  A. Demarcating child protection for the benchmark 3
  B. Direct and indirect costs 6

C: Funders and expenders of child protection expenditure 7
  C. The level of the state for which expenditure will be included 15
  D. What expenditure data will be used? 18
      Summarising the process for selection of relevant expenditure 20

Section 1.3 Proposal for construction of the benchmark 21
  Core benchmark content 21
  Extended benchmark 25

Section 1.4 Presentation of the indicator 25

Section 1.5 Summary of proposed benchmark 31

SECTION 2 A PROCESS FOR PILOTING THE BENCHMARK 32

Section 2.1 Objectives of the pilot 32

Section 2.2 Scope of the pilot 33

Section 2.3 Piloting Process 34
      Phasing 34
      Team requirements 35
      Time lapse 35

ANNEX 1: BIBLIOGRAPHY  Annex p1

ANNEX 2: MID-ASSIGNMENT REPORT (LITERATURE REVIEW)  Annex p3

ANNEX 3: DISCUSSION OF OPTIONS FROM OPTION PROPOSAL  Annex p84
# Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADB</td>
<td>Asian Development Bank</td>
</tr>
<tr>
<td>CP</td>
<td>Child Protection</td>
</tr>
<tr>
<td>CPWG</td>
<td>Child Protection Working Group (of the IASC Protection Cluster)</td>
</tr>
<tr>
<td>CRC</td>
<td>Convention on the Rights of the Child</td>
</tr>
<tr>
<td>CWI</td>
<td>Child and Youth Well-being index</td>
</tr>
<tr>
<td>EAPRO</td>
<td>East Asia and Pacific Regional Office</td>
</tr>
<tr>
<td>GDP</td>
<td>Gross domestic product</td>
</tr>
<tr>
<td>GNI</td>
<td>Gross national income</td>
</tr>
<tr>
<td>HDI</td>
<td>Human Development Index</td>
</tr>
<tr>
<td>INCAF</td>
<td>OECD DAC International Network on Conflict and Fragility</td>
</tr>
<tr>
<td>ICP</td>
<td>International Comparison Programme</td>
</tr>
<tr>
<td>MTEF</td>
<td>Medium Term Expenditure Framework</td>
</tr>
<tr>
<td>OBI</td>
<td>Open Budget Index</td>
</tr>
<tr>
<td>ODA</td>
<td>Official Development Assistance</td>
</tr>
<tr>
<td>OMBI</td>
<td>Ontario Municipal Benchmarks Initiative</td>
</tr>
<tr>
<td>PEFA</td>
<td>Public Expenditure and Financial Accountability (framework)</td>
</tr>
<tr>
<td>PFM</td>
<td>Public Financial Management</td>
</tr>
<tr>
<td>PPP</td>
<td>Purchasing Power Parity</td>
</tr>
<tr>
<td>SP</td>
<td>Social Protection</td>
</tr>
<tr>
<td>SPEXP</td>
<td>Social protection expenditure</td>
</tr>
<tr>
<td>SPI</td>
<td>Social protection index for Asia</td>
</tr>
<tr>
<td>TORs</td>
<td>Terms of Reference</td>
</tr>
<tr>
<td>UNHCR</td>
<td>United Nations High Commission on Refugees</td>
</tr>
<tr>
<td>UNICEF</td>
<td>United Nations Children’s Fund</td>
</tr>
<tr>
<td>WGI</td>
<td>World Governance Indicators</td>
</tr>
</tbody>
</table>
Introduction

1. This is the third output in a project to develop proposals for a financial benchmark for child protection (CP), commissioned by UNICEF. The financial benchmark is intended to provide a comparable measurement of expenditure on child protection across countries, as well as an indicator of the adequacy of expenditure. The purpose of this output is to develop the preferred option for the construction of an indicator further, and discuss a process for piloting the benchmark.

2. The paper draws on the first two outputs, namely the Mid-Assignment report, which was a literature review, and the proposals on options for the benchmark. These are included as Annex 2 and 3. The text in the annexes from these earlier outputs in the assignment has not been substantially altered -- besides corrections and clarifications from the comments received -- as they function as discussion papers towards this final output. The proposals presented in this paper draw on the first two outputs, and are framed within the preferences indicated by the UNICEF Child Protection Financial Benchmark reference group, in response to these outputs.

3. This paper is set out in 2 main sections, besides this introduction. Section 1 summarises the parameters for the benchmark; and Section 2 discusses a process to pilot the benchmark.

Section 1: Proposals for the Financial Benchmark

4. As discussed in the mid-assignment report, the design of a UNICEF child protection financial benchmark would need to address three core questions:
   - What child protection expenditure data will be included / excluded?
   - What benchmarking methodology will be used, where methodology refers to the architecture of the benchmark and its application?
   - What interpretive perspectives are needed as part of, or to frame, the expenditure measure(s) used in the benchmark?

5. The UNICEF preferred financial benchmark architecture/construction is a simple unitary summary indicator, expanded through the interpretive information in its presentation.

6. Following the setting out of general design parameters, the sub-sections below sets out the design for the indicator against each of these main questions.

Section 1.1 General design parameters

7. The child protection financial benchmark is aimed at providing a comparative measure of expenditure on child protection actions and services across countries, for in-country and cross-country use as an analysis and advocacy tool. Any country-specific decisions within the framework for the benchmark described in this report, should at all times be guided by this purpose. The design should allow for the measure to be applied to countries over a period time while still providing comparability, as the intent is not to create a global assessment programme that runs periodically, but rather that individual countries (or regions) would undertake assessments at regular intervals.
8. The same benchmark methodology for the core benchmark should be used in all countries.

9. The benchmark is not policy-neutral: the benchmark is aimed at providing an analysis and advocacy tool against specific global concerns and recognized good practices with regards to child protection. These concerns and practices are informed by Article 19 of the Convention on the Rights of the Child, associated documentation such as the Committee on the Rights of the Child General Comment No.13 (CRC 2011) and the UN Secretary-General’s Study on Violence Against Children (UN Secretary General 2006). The benchmark is therefore primarily exogenously defined in relation to individual countries. However, as its design must allow fair and equal comparison between countries, it must allow for variations in how countries respond to child protection risks given different socio-cultural values and political and administrative traditions.

10. The benchmark will assess expenditure on formal child protection systems. This formal system consists of all efforts by all levels of the state, international organizations and local NGOs (including community- and faith-based organizations) involved in providing child protection that are recognized or endorsed by and subject to supervision and regulation by a state. The benchmark will not attempt to assess expenditure in informal systems, mainly associated with the practices used and initiatives undertaken by communities and families for protecting children.

11. The core benchmark is intended as an analysis and policy tool in UNICEF’s policy advocacy with its government partners. While assessing formal systems, the core benchmark is therefore is state-centric or of what is deemed public expenditure, in other words a measure of the extent of the state’s response (direct state actions and facilitation of others to act) to child protection risks and needs. The discussion below includes a definition of which expenditures would be considered part of the state’s response, important for calculating the benchmark in countries that receive official development assistance (ODA).

12. An extended benchmark will be assessed in exceptional circumstances where the state is for the time being not able or not available to act -- or historically not acting -- as the primary duty bearer for fulfilment of the rights of the child under CRC Article 19. The proposals below include a set of criteria for when a country would move into the ‘exceptional’ category, and a methodology for assessing expenditure in such circumstances. The extended benchmark of course can be calculated for any country, should the UNICEF Child Protection country team wish to invest the additional resources and effort to assess expenditures by other actors.

13. When used as absolute numbers, all child protection expenditure and other financial data will be standardized to one currency, to enable the comparability envisaged. The proposal is to use a USD purchasing power parity (PPP) conversion, utilizing the International Comparison Programme (ICP) PPP estimates. Where indicators of child protection expenditure are ratios and percentages, this standardization would not be as necessary.

14. As the benchmark will not be applied in a regular global process, assessing all participating countries at once, it is necessary to define an assessment window relative to the assessment being undertaken for comparisons to be valid. The proposal is comparison to other assessments within 3 years of the previous assessment having been undertaken. Also, in the interest of limiting the cost of benchmark application, unless the benchmark construction
choice requires a different approach, expenditure for one year only will be assessed, except in the case of capital or investment expenditure, which may spike in any one year as a bulky investment is made. For this expenditure a three year average should be used.

15. Overall the parameters for the benchmark in a situation of scarce data in the exact format required – namely emphasizing cross-country comparability, the desire to count both national and sub-national expenditure, and take into account exceptional circumstances while taking cost and effort of application into account – result in a relatively complicated methodology, with significant compromises. It is proposed that the presentation of the methodology should make these explicit, and present the limitations at all times.

**Section 1.2 What child protection expenditures will be included?**

16. The mid-assignment report identified the following parameters for identifying the expenditures that will be included in the benchmark:

- **A. Demarcating child protection for the benchmark.** Key questions were which definition of child protection will be used, what demarcation rules will be applied to include and exclude expenditures, and whether these rules will be the same across countries.
- **B. Direct and indirect expenditure on child protection services.** The key question was whether only the direct cost of providing a service would count (e.g. the cost of conducting background checks, or running inspecting services), or also the indirect cost of developing policies, drafting regulations, monitoring services, publishing data etc.
- **C. The funders and expenders included.** Key questions were whether (i) both domestically and internationally financed expenditure would count, (ii) whether only government financed or also non-government financed local expenditure would count, and (iii) whether it would matter who delivers the service, government actors, NGOs, private providers or families. Again a key question was whether these rules would remain the same across countries.
- **D. The level of the state at which services are provided.** The key question was whether sub-national (regional and local) expenditures would be included.
- **E. The expenditure data that will be used.** This question unpacked in two main sub-questions, namely (i) whether budgeted, disbursement, actual outturn or audited expenditure data will be used; and (ii) whether the benchmark will collect and reflect data disaggregated by economic classification.

This sub-section sets out the benchmark parameters against each of these parameters.

**A. Demarcating child protection for the benchmark**

17. The focus of the benchmark will be on child protection defined as: “prevention of and response to violence, abuse, exploitation and neglect of children” (UNICEF 2014).

18. Countries will be provided with an identification matrix for core, targeted prevention and protections services for which expenditures must be included in the financial benchmark. A preliminary proposal for the matrix is provided overleaf, developed from the initial list provided in the Terms of Reference (TORs), responses received on the mid-assignment and options reports, the UNICEF Core Toolkit to Map and Assess Child Protection Services (UNICEF
2010) and the UN Secretary General Study on Violence against Children (2006). The matrix provides two axes along which services are to be identified by assessors\(^1\) at the country level, namely (i) groups of children in need of protection (or a list of global, specific child protection risks/harms) and (ii) types of services (or a list of typical prevention and response services).

All services in-country that are aimed at children falling in any one or more of the groups of children in need of protection and which are of the types of services identified, must be included – i.e. services which can be mapped onto the matrix.

19. In addition, for instances of services which the assessors have identified as potentially child protection expenditures, but which does not fall clearly in a cell in the matrix, the decision rule is that the expenditure has to be made deliberately and specifically to prevent and respond to child protection concerns, with these concerns demarcated by the types of harm identified. This means that the benchmark is exogenously defined, by the risks / types of harm that are seen as by UNICEF as global priorities.

20. The same matrix will be used in all countries, whether the child protection system is operating under normal or exceptional circumstances. Provision however is made for variation in responses between countries, insofar as a set of expenditures can be included that is not supporting one of the pre-defined types of services, but which nonetheless meet the deliberately and specifically to prevent and respond to the types of harm identified -- i.e. it must fit under the columns of the matrix, even if it doesn't correspond to any of the rows.

21. Expenditure on services that do not fit under the columns of the matrix is excluded. Specifically, for the core indicator, expenditures on social protection in general should be excluded, unless they are for these children specifically. For example, income supplements or other entitlement payments to families are excluded, unless they are for children specifically, such as foster care grants and child support grants. Other support services for individual families, unless they are child-specific, are also excluded.

\(^1\) We use the term ‘assessor(s)’ to refer to the person(s) who will be applying the financial benchmark methodology at country level.
Table 1  Child protection expenditure identification matrix

<table>
<thead>
<tr>
<th>INCLUSION REQUIREMENT</th>
<th>CHILDREN NOT REGISTERED AT BIRTH</th>
<th>CHILDREN IN LABOUR AND OTHER WORK THAT IS HARMFUL</th>
<th>CHILDREN SUBJECTED TO HARMFUL CULTURAL PRACTICES, e.g. CHILD MARRIAGE, FGM/C, GENDER DISCRIMINATION</th>
<th>ABUSED CHILDREN (PHYSICAL, SEXUAL, EMOTIONAL)</th>
<th>NEGLECTED CHILDREN</th>
<th>CHILDREN WITHOUT ADEQUATE FAMILY CARE</th>
<th>CHILDREN ON THE MOVE DUE TO MIGRATION, KIDNAPPING AND TRAFFICKING</th>
<th>CHILDREN WHO ARE SEXUALLY EXPLOITED COMMERCIALY</th>
<th>CHILDREN IN CONTACT WITH LAW</th>
<th>CHILDREN AFFECTED BY EMERGENCIES</th>
<th>CHILDREN AFFECTED BY ARMED CONFLICT AND VIOLENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDENTIFICATION GUIDE:</td>
<td>ACTION/SERVICE IDENTIFICATION GUIDE</td>
<td>↓</td>
<td>Prevention</td>
<td>Prevention</td>
<td>Prevention</td>
<td>Prevention</td>
<td>Prevention</td>
<td>Prevention</td>
<td>Prevention</td>
<td>Prevention</td>
<td>Prevention</td>
</tr>
<tr>
<td>Action/service MUST be aimed at</td>
<td>Action/service MUST be aimed at</td>
<td>Action/service MUST be aimed at</td>
<td>Action/service MUST be aimed at</td>
<td>Action/service MUST be aimed at</td>
<td>Action/service MUST be aimed at</td>
<td>Action/service MUST be aimed at</td>
<td>Action/service MUST be aimed at</td>
<td>Action/service MUST be aimed at</td>
<td>Action/service MUST be aimed at</td>
<td>Action/service MUST be aimed at</td>
<td>Action/service MUST be aimed at</td>
</tr>
<tr>
<td>Prevention</td>
<td>Public education &amp; community mobilisation</td>
<td>Birth registration</td>
<td>Life skills, youth civic engagement (e.g. child-friendly spaces)</td>
<td>At-risk children &amp; families identification</td>
<td>Background checks &amp; codes of conduct for those working with children</td>
<td>Individual family support, e.g. income supplements, mediation, entitlement assistance, service access, respite entitlement, legal aid, parenting groups</td>
<td>Reporting/Complaints mechanisms</td>
<td>Response</td>
<td>Verification, investigation &amp; assessment</td>
<td>Referral, best interest determination &amp; gate keeping procedures</td>
<td>Sensitive health, police, judicial, social work interventions (e.g. counselling, case management)</td>
</tr>
</tbody>
</table>

5
B. Direct and indirect costs

22. While this sets out the types of services relative to types of harm and typical services, it does not identify which components of a service would count. As identified in the TORs and discussed in the mid-assessment report, the benchmark should work with the notion of a child-protection system, rather than an issue-based approach which only looks at the actions to prevent and or respond to individual issues or problems.

23. The Chapin-Hall diagram and its interpretation in the UNICEF Toolkit to map and assess child protection systems (UNICEF 2010) systematically unpack a service into its different system components. Table 2 below reflects this unpacking, noting which components of the system are likely to carry explicit costs that are easier to assess, which should then be assessed in the application of the financial benchmark.

24. While the team sought to follow the exclusion rule proposed in the mid-assessment report to exclude expenditures that are not easily quantifiable, this was found to be risky for the usefulness of the benchmark. Expenditure on functions (governance, management and enforcement), and on accountability (data collection, monitoring and evaluation, quality standards, research, analysis, communication) would be amounts that are not necessarily discrete and which would be difficult to quantify. However, given the importance of these expenditures for child protection outcomes, excluding them would limit the standardized interpretive data that can be used in interpreting the indicator. For example, if country x spending 0.5 per cent of GDP on child protection of which none is on accountability, this may have a lesser effect on child protection outcomes than country y that spends 0.4 per cent of GDP, but spends 5 percent of that on ensuring accountability for expenditure and results. For this reason, expenditure on all components of the system must be quantified, using the expenditure estimation tool proposed.

25. UNICEF however may still wish to consider excluding these two sets of expenditure after the pilot assessments, given that they will require additional work to quantify and include, making country processes to calculate the benchmark more complex and costly. The key trade-off is between the usefulness of being able to comparatively assess how countries are delivering child protection through spending on ‘backroom’ functions, versus the lower complexity of calculating a cross-country comparable measure of direct expenditure on services only.

Table 2 Tracking components of child protection systems

<table>
<thead>
<tr>
<th>Proposed indicator marker</th>
<th>Chapin-Hall system description</th>
<th>UNICEF Toolkit unpacking</th>
<th>Use in Financial Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mapping of child protection services in Table 1 matrix</td>
<td>Relationships between systems, components and actors</td>
<td>Overview of systems, components and actors</td>
<td>No cost assessment (therefore excluded)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Useful step to identify relevant budgets</td>
</tr>
<tr>
<td>Development of child protection institutions (where institutions refer to)</td>
<td>Functions (governance, management, enforcement)</td>
<td>Laws, policies, standards &amp; regulations</td>
<td>While policy development and legislative functions on child protection is not easily quantifiable(^2), expenditure on standards and regulations, cooperation, coordination and collaboration will be. Given the likely impact of</td>
</tr>
</tbody>
</table>

\(^2\) A proxy cost could be determined, for example using an adjusted staff cost for the units responsible for this.
the formal and informal rules that manage child protection) cooperation, coordination and collaboration these functions on the quality of child protection expenditure, it would be important to include, even if negligible compared to implementing services, particularly if the composition of expenditure is part of the benchmark. The costs are therefore included.

<table>
<thead>
<tr>
<th>Capacity building functions</th>
<th>Capacities (human resource, infrastructure, funding)</th>
<th>Capacity Building</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Capacity building is a key factor in determining whether policies and budgets are translated into effective services, both in terms of human resources and infrastructure. It can also carry a sizable cost. Expenditure on capacity building and infrastructure development within any of the qualifying services should therefore be included.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Service delivery functions</th>
<th>Continuum (promotion, prevention, response) and process of care (identification, reporting, referral, investigation, assessment, treatment, follow up) Communication, education and mobilization for social change Service and service delivery mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>This is the core of on-going expenditure on undertaking actions and delivering services for child protection. Identifying which actions and services are included and excluded in the continuum of care and process of care, is the subject of the identification matrix and the expenditure selection methodology. Qualifying services are included.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Accountability functions</th>
<th>Accountability (data collection, monitoring and evaluation, quality standards, research, analysis, communication) Accountability mechanisms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>While part of the overhead costs, these are important investments in ensuring the effectiveness and efficiency of expenditure on child protection actions and services. Costs could be significant, can be quantified, and would differ between countries, and should therefore be included.</td>
</tr>
</tbody>
</table>

26. In summary, for the qualifying services in Table 1 capital expenditure on infrastructure and other capital goods, as well as recurrent expenditure on the development of child protection institutions; capacity development; service delivery, and on accountability functions (e.g registering and inspecting facilities) are included.

These demarcation rules in terms of child protection and types of actions related to child protection, will be applied consistently across countries.

**C: Funders and expenders of child protection expenditure**

27. This is the main parameter against which the benchmark will make a distinction between ‘normal’ and ‘exceptional’ circumstances in countries. Where child protection is delivered under ‘normal’ country circumstances, only the core benchmark will be state-centric or of deemed public expenditure. Under exceptional circumstances, the calculation of an extended benchmark is advised, which will also include services funded by other actors, such as local and international donors, NGOs and the private sector.

28. The assessed benchmark in countries with exceptional circumstances would still calculate a core benchmark assessing public expenditure, namely expenditure on state financed or managed services (see below for definitions), but would also calculate an extended benchmark, collecting data for other funders. This will allow cross-country comparisons on the core benchmark with all countries, but also comparison on the extended benchmark, with other countries operating in similar exceptional circumstances. This – and other parameters for comparison -- is discussed further in the section on presentation of the benchmark below.
29. We discuss first a proposal to identify in which category a country falls, before discussing which the expenditure associated with which funders and expenders will count.

**Categorisation of countries**

30. Three factors have emerged that designate countries as exceptional, requiring assessment of child protection expenditure beyond that of government. It is proposed that if a child protection system operates in any of these circumstances, it would be considered exceptional circumstances, triggering the calculation of the extended benchmark.

31. **Factor 1: The provision of child protection services in emergencies / complex emergencies.** The proposal is to consider countries affected by emergencies as exceptional circumstances, as emergencies affect or may overwhelm the capacity of the state to deliver services. Under such circumstances it would be useful for the UNICEF Country Office or Regional Office to have data on child protection expenditure by all actors, to support advocacy and analysis. It may also be that the benchmark assessment cannot be undertaken in some emergency contexts due to the specific circumstances.

32. **Definition:** UNICEF defines an emergency as “a situation which threatens the lives and well-being of large numbers of a population and in which extraordinary action is required to ensure their survival, care and protection. Emergencies may be created by natural or technological disasters, epidemics or conflicts” (UNICEF 2007). A complex emergency is defined as “a humanitarian crisis in a country, region or society where there is a significant or total breakdown of authority resulting from internal or external conflict and which requires an international response that extends beyond the mandate or capacity of any single agency” (UNICEF 2007).

33. The proposal is to include countries affected by Level 2 or 3 emergencies – defined as situations where the humanitarian needs are of sufficiently large scale and complexity that significant external assistance and resources are required and a multi-sectoral response is needed with the engagement of a large range of international humanitarian actors -- but exclude Level 1 emergencies which are circumstances in which Government(s) affected is/are generally able to respond to the needs, but may request specialized assistance.

34. These definitions are drawn from Office for the Coordination of Humanitarian Assistance definitions, which categorises emergencies by criteria associated with the emergency itself. UNICEF also refers to Level 1,2 and 3 emergencies, but distinguishes between them with reference to the demands they place on UNICEF, and how UNICEF would respond. Level 2 emergencies for example would be emergencies where the magnitude of the emergency is such that a Country Office needs additional and prioritized support from other parts of the organization (HQ, RO and COs) to respond and that the RO must provide leadership and support, while Level 3 emergencies require a corporate response and are cases in which the UNICEF Executive Director has activated the Corporate Emergency Activation Procedure (CEAP).
35. **Identification**: While there are cross-agency procedures to identify Level 3 emergencies through the Inter-Agency Standing Committee Transformative Agenda Protocols (IASC, 2012), procedures with regards to Level 2 emergencies are region specific. It is therefore proposed that the Benchmark should use the UNICEF definitions as a proxy. In other words, countries where UNICEF has declared a Level 2 or Level 3 emergency would be considered exceptional circumstances for the duration of the emergency and up to 3 years after moving out of the L2 status, given time required for recovery. In some cases countries would move into emergencies in between assessments, which would require the calculation of an extended benchmark. Emergencies may also affect only a part of a country, in such cases too, an extended benchmark would be calculated, but only for the non-state financing that goes into the emergency. If providers of services cannot clearly delineate expenditure on the zone experiencing an emergency from the expenditure in the rest of the country, a proportional calculation must be done.

36. **Factor 2: The provision of child protection services in situations of fragility.** A second set of countries that will be considered exceptional, are the states operating in situations of fragility. As the state is not in a position to provide comprehensive child protection services in situations of fragility, the benchmark would take into account additional sources of external and domestic financing. It follows that UNICEF’s advocacy targets for child protection in such circumstances would include non-state actors.

37. **Definition**: There is no internationally-agreed definition of ‘fragile states’, or ‘fragility’. However, it is most often principally defined as a fundamental failure of the state to perform functions necessary to meet citizens’ basic needs and expectations. States operating in situations of fragility are commonly described as incapable of assuring basic security, maintaining rule of law and justice, or providing basic services and economic opportunities for their citizens (Governance and Social Development Resource Centre, 2014). The OECD DAC characterises fragile states as: ‘unable to meet [their] population’s expectations or manage changes in expectations and capacity through the political process’. The Centre for Research on Inequality and Social Exclusion, for example, defines fragility as describing a country that is failing or at high risk of failing in three dimensions: (i) authority failures: the state lacks the authority to protect its citizens from violence of various kinds; (ii) service failures: the state fails to ensure that all citizens have access to basic services; (iii) legitimacy failures: the state lacks legitimacy, enjoys only limited support among the people, and is typically not democratic.

38. Fragility affects a wide range of countries and economies. As set out in the 2014 OECD Report on Fragile States (OECD 2014), it includes countries that are recovering from conflict and embarking on peace and statebuilding processes (e.g. Liberia, Myanmar and Timor-Leste). It also includes countries that are experiencing long-term insecurity, recurrent crises or localised conflict (e.g. the Central African Republic, Guinea-Bissau and Yemen), or high levels of criminality and violence (e.g. Pakistan). It is found in a range of situations where governments have strong administrative structures but where political exclusion combined with lack of economic opportunities are fuelling tension and violence. This is the situation in countries such as Egypt, Libya and North Korea, as well as across entire regions such as the Maghreb and the Sahel (OECD, 2014, p16).
39. **Identification**: Every year the OECD DAC International network on conflict and fragility (INCAF) compiles a list of countries and economies that are considered to be fragile\(^4\), combining the latest harmonised list of countries on the World Bank, Asian Development Bank and African Development Bank lists, with countries that have a Failed State Index above 90 on the Failed States list developed by the Fund for Peace. It is proposed that this list is used to identify countries that are considered to be providing child protection services under conditions of fragility, in other words, where the state is not able to provide services.

40. It is expected that countries in which international development partners are unable to disburse funding to government would be considered to be either countries affected by emergencies or state fragility.

41. **Factor 3: The provision of child protection services in circumstances where non-governmental actors are the historical providers.** UNICEF has identified the need to treat countries differently in which non-governmental actors have historically provided child protection services, *with funding for such services flowing directly to these actors*, rather than to or through the state. Only considering expenditures associated with state-financed or managed services in circumstances where funding traditionally does not flow through the state, would present a skewed picture of child protection and would not equip the UNICEF Child Protection country team to advocate for services with all key stakeholders.

42. **Definition and identification**: The definition of “non-governmental actors as historical providers of child protection services” will be benchmark specific. The key judgment to make is whether the state is historically the channel through which such services are financed, notwithstanding whether the services or service providers are recognized and regulated by the state. It is proposed that this judgment should be made in the application of the benchmark, on a country-by-country basis.

43. The methodology for making the judgment on whether to apply the extended benchmark, requires

i) *An assessment of the proportion of all child protection services financed by the state.* When child protection services are mapped using the Table 1 matrix, the country assessment team should make a preliminary assessment based on pre-assessment evidence, whether for each service mapped the understanding is that it is predominantly and traditionally financed through or managed by the state (using the definition of ‘managed by the state’ as proposed below). It would be important to consider reach and coverage of state services in this assessment, as in some countries the team may find that whereas the state finances services for some parts of the country, services in difficult to reach regions may be financed traditionally by non-state actors, without any state involvement.

ii) *An assessment of the usefulness of extended benchmark data:* If for more than 50 per cent of services mapped for the state as a whole or a specific region, the financing is not traditionally by the state, or does not traditionally flow through the state the country Child Protection team should consider the purpose of calculating the benchmark. Is it predominantly to advocate for greater financing of services

---

\(^4\) This is done in the annual fragile states report, see for example OECD INCAF (2014). Fragile States 2014: Domestic revenue mobilisation in fragile states. Paris, OECD.
notwithstanding government involvement in the flow of funding, or would a contrast with or calculation of non-state financed or managed expenditure assist in advocating for more financing from government for services? If the answer to either these questions is yes, the team should consider calculating the extended benchmark. The Country Office should however weigh the additional cost of an extended benchmark against the benefit of having the information and provide the team undertaking the assessment with instruction.

Of course, the questions under (ii) would guide any country in deciding whether to calculate an extended benchmark or not.

**The core benchmark: qualifying funders and spenders**

44. The core benchmark will include public expenditure on child protection, deemed to be *all expenditure on qualifying services that is funded internally*, ie by tax revenues, levies, fees and charges in terms of country legislation, regardless of who undertakes the expenditure. It will also include all externally financed expenditures (by local and international donors), notwithstanding who delivers the actual service, as long as the expenditure is managed by government.

45. **Internally financed**: Internally financed services refer to services financed by government revenues. This would by definition include co-payments (for example for birth registration), insofar as internally funded includes all fees and charges levied by the state for services. This should in most cases not involve additional data collection, as benchmark data will be collected from the expenditure side of the budget. Even when co-payments are retained at the facility level and not deposited as central revenue, co-payments in most countries still need to be approved and the commensurate expenditure appropriated by the legislature. Budgets traditionally therefore show gross projected co-payments on the revenue side, and gross projected expenditure on the expenditure side of approved budgets. A standard check in the process of applying the benchmark however would be to check this. If co-payments are netted out, data on them need to be collected and added to the budgeted/actual expenditure figure used, to ensure comparability with countries where co-payments are reflected gross on budget and in reports.

46. **Managed by the state**: The proposal is also to include services that are externally financed (by local and international donors), but only if managed by government. Defining whether a service is managed by the state is not straightforward.

47. Firstly, 'managed by' requires further definition. For example, one can envisage a justice-sector child protection programme that is donor funded, and for which the funds are disbursed to government, and which is fully managed through government’s own public financial management and audit systems. On the other hand, one could also envisage a child protection programme that disburses to a project implementation unit which is housed in government, but which manages its own bank accounts, accounting and procurement system and so forth. However, it is not impossible that the former set of arrangements may be a programme which was introduced by the donor, whereas the latter is the result of an initiative for which government sought external financing to complement its internal financing. This set of issues is further complicated by the variety of ways in which donors use country systems, depending on donor-specific factors such as their domestic legal framework for aid, and
appetite for fiduciary risk. This may mean that a programme which is funded by donor A will not be included, as the donor’s legal / policy framework does not allow use of country systems, but the same programme by donor B is, as donor B’s frameworks do allow the use of country systems.

48. A second complication is that ‘the state’ can refer either to the public sector, or more narrowly to general government (whether at the central, state or local government level). According to the System of National Accounts, as applied to Government Finance Statistics (GFS), the Public Sector comprises two main branches, public corporations and general government. General government in turn comprises institutional units (albeit ministries, departments or agencies) that carry out the functions of government as their primary purpose (IMF 2001). This demarcation (and sub-branches) is depicted in the diagram below.

**Figure 1. General government and public corporations**

Source: IMF (2001)

49. The proposal is that for expenditure on an *externally financed service* to count:
   - The service must be delivered by a general government unit, on the logic that all services delivered by government units (i.e. not outsourced to the private sector, non-government sector or households) are services by the state even if they are financed externally, and the financial flows not managed by country systems to a more or less significant degree; or
   - The service must be managed by a general government unit, meaning that it must be an initiative of government (therefore have government ownership) and at least to some extent under control of government, signified by the account from which the expenditures are paid being in the name of government, and requiring the (co-) signature of a designated government representative to release funds.
50. In summary thus far, for the core benchmark expenditures will be deemed to be public if:
   - The qualifying expenditure is financed by domestically raised taxes, levies, fees and charges
   - The qualifying service is delivered by a general government unit, as defined by the IMF GFS (even if the financial flow is not managed by government systems)
   - The service is financed from an external source, but managed by a general government unit, meaning that it must be an initiative of general government and be disbursed to an account in the name of a general government unit, even if in a commercial bank

51. This proposal is depicted in the diagram below:

**Figure 2. Inclusion: Services delivered and funded by whom?**

![Diagram showing inclusion of services delivered and funded by whom]

The extended benchmark: qualifying funders and spenders

52. The extended benchmark can be calculated for any country. However, it is advised that an extended benchmark is calculated for countries: (i) that are experiencing emergencies or have experienced an emergency in the three years before calculating the benchmark (identified by UNICEF as level 2 or 3 emergencies); or (ii) that are governing in circumstances of fragility (identified by their inclusion on the OECD annual list in the year of assessing the benchmark); or (iii) in which funding is traditionally not by the state or channeled through the state for child protection.

53. The extended benchmark would be for all expenditure on the formal child protection system, whether financed by government, domestic and international donors, local non-governmental actors (including NGOs, community and faith-based organisations) or the private sector. Formal systems exclude practices used by or initiatives undertaken by
communities and families for child protection, but includes all efforts within the Table 1 matrix that are recognized or endorsed by the state, or are or should be subject to supervision and regulation by government.

54. The extended benchmark would need to be calculated on the basis of a representative estimation of additional financing of expenditure over the core benchmark within a country. A methodology for this is proposed in Box 1 below.

**Box 1. Proposed methodology for collecting data on additional financing of child protection expenditure**

The methodology proposed is based on the extended benchmark taking into account additional financing of child protection expenditures, as a proxy for the additional expenditure by non-state service providers over and above their expenditures that are financed by or managed through government. The collection of data therefore is from the additional main funders, rather than providers of formal services, although providers or NGOs managing funding for services on behalf of funders will also be useful sources of information. It is deemed that collecting expenditure information from all providers would be highly costly, and that sampling would not be an option, since one would need an assessment of the population before you could sample and extrapolate reliably, and in most countries collecting information on the population of service providers and their expenditure, would be if not impossible, very costly.

**Step 1: Using the matrix and a network approach to map all funders of expenditure:** Once the UNICEF Child Protection country team has agreed the need to calculate the extended benchmark, the benchmark team should map all financiers of expenditure on child protection against the matrix in Table 1. On the assumption that child protection is a specialized area, and that all child protection actors in a country would be more or less connected through a network, this would require the team starting interrogating UNICEF on (i) its financing of child protection services that is not channeled through or managed by the state, and (ii) other financiers of child protection services, both local and international. The benchmark team would then need to follow up with additional funders or managers of funding identified, asking the same questions. The process should be repeated until no additional funders are being identified, by existing child protection actors.

In countries where sub-national locations are being sampled, the team should follow a similar approach to map additional funders at the sub-national national level.

**Step 2: Surveying funders of child protection expenditure:** The team must then distribute a survey to the identified funders or managers of funding for services, or interview the funders or managers. The survey should request information on (i) the quantum of financing *disbursed* by provider, in the same year for which the core benchmark is being calculated, and (ii) the purpose of the financing, categorized within the matrix and by whether it is on the regulatory functions, capacities, continuum of care or accountability components of service delivery (refer Table 2 above), and (iii) the source of funds. It would be critical to get information on the source of funding reported, in order to avoid double counting.

As funders of services would not be able to consistently provide a regional breakdown of in which geographical area a service is being delivered, or the economic classification of the expenditure, the extended benchmark would not allow disaggregation against these categories.

If financial information is provided in currencies other than the local currency, the team must convert to the local currency, using the average exchange rate for the year being assessed.

**Step 3: Collating information, and netting out funding included in the core benchmark.** The final step is to compile total extended benchmark expenditure in the local currency, taking care to net out funding that was already included in the core benchmark (this would only be expenditure financed by international and local donors, but managed by government).
Where data for the extended benchmark was collected at the local level, the team would take the average proportion of the extended benchmark.

Calculating an extended benchmark would be resource intensive, as it would require first mapping, then surveying and finally data cleaning and quality checks for all funders of child protection expenditure. As ODA donors do not classify their expenditure at a level of disaggregation that is useful for the exercise (the only Creditor Reporting System Code that is directly relevant is 15261 Child soldiers -- prevention and demobilization), it would also require an investment of time by the funders to complete the survey. UNICEF Child Protection country teams should therefore only calculate the extended benchmark if highly necessary for advocacy and analytical purposes.

### C. The level of the state for which expenditure will be included

55. Countries differ significantly in the distribution of services across levels of government and the underlying intergovernmental fiscal arrangements to finance services. There is therefore consensus that within the above rules around flow of funds for services, services at all levels of government will be included.

56. However, pragmatically, in some cases, the proportion of expenditure at sub-national levels of government may be so small, that the marginal difference it makes to the total expenditure assessed, may not warrant the cost of collecting the data from all locations, particularly when data sets are not accessible without visiting every location. The methodology for assessing sub-national expenditure in Box 2 below therefore starts with an assessment of the need to collect sub-national information. Note that the methodology is based on child protection being a specialized area of public policy and service delivery, and therefore likely better knowledge of the system by the network of child protection stakeholders in a country.

57. Furthermore, when sub-national expenditure is deemed to be significant, it would not be possible to collect the information from all regions or local governments. If the functional or programmatic classification of government expenditure provided standard codes to identify all expenditures in detail, and if sub-national data were generally available at national level, it may have been possible. However, since it is likely that most child protection expenditure will not be identified by a specific expenditure classification code, but will require proxy estimations, and in many countries sub-national data would not be readily accessible at the national level, collecting the universe of sub-national data for a country would not be possible without visiting every location. Box 2 below therefore sets out a methodology for sampling localities, and extrapolating the information for an estimation of all sub-national expenditure. Again, this methodology is based on accessible pooling of knowledge of the child protection system among key stakeholders at national level.

58. Note that the methodology as a whole would need to be used when the state is fiscally decentralized, namely where revenues are transferred to elected sub-national governments through fiscal transfers, or these governments raise their own revenue to finance expenditure. In addition, the second step onwards of the methodology would be used when services are devolved, namely when regional or local offices of national governments undertake or manage expenditures and sufficiently disaggregated data on these expenditures (e.g. how much of the total expenditure on civic services office of government is spent on birth registration) are not available at the national level.
59. The methodology proposed uses stratified sampling to guard against a bias towards poorer or richer sub-national localities. Identifying criteria for stratifying a sample for example is also proposed in the PEFA guidance on sub-national PEFAs (Doex 2013). A two-phased sampling is proposed, in order to minimize the cost of sub-national data collection, if possible, given that in some countries the number of sub-national localities may stretch into hundreds. In a first sample a smaller number of localities are assessed, to develop the data for these localities and to check on variation between localities. If the variation is significant, a second sampling round would be required in order to estimate full sub-national expenditure based on the sample with confidence. The first sample will also assist the team to identify additional criteria by which the second sub-national sample may need to be stratified, in order to not bias the calculation towards a specific set of sub-national localities.

**Box 2. Methodology to assess sub-national expenditure**

<table>
<thead>
<tr>
<th>This methodology applies to the calculation of the core benchmark, which by its nature assesses the on-budget expenditure on child protection, plus expenditure that is deemed managed by the state even if the cash does not flow through the central treasury.</th>
</tr>
</thead>
<tbody>
<tr>
<td>The sub-national level of government may refer to expenditure either by regions or by local governments in a fiscally decentralized state, or by regional or local offices in a unitary state.</td>
</tr>
<tr>
<td>The methodology would be used for child protection services that are provided at the sub-national level, and financed either by an unconditional grant from a higher level of government or locally collected revenue. For conditional grants that are child-protection specific (e.g. specifically to finance the identification of children at risk in schools), the total transfers per region would be included and it would not be necessary to collect data at the local level. If conditional grants also finance non-child specific services, the methodology would need to be used in order to arrive at an estimate of the proportion of the total transfer per locality that is used for child protection. In applying the methodology, the assessment team would need to take care not to double count expenditure.</td>
</tr>
<tr>
<td>The methodology would be applied once all services provided in the country and within the matrix are mapped, and sub-national expenditure identified.</td>
</tr>
</tbody>
</table>

**Step 1: Assessing whether to collect data on sub-national expenditure (use for fiscally decentralized states):** Data for sub-national locations should be collected if the UNICEF country Child Protection team believes that more than 10 per cent of child protection expenditure mapped in the matrix and financed or managed by the state is used or disbursed by sub-national units of government, and this assessment is triangulated with other key child protection actors, such as the key ministry(ies) responsible.

**Step 2: Preliminary assessment of sub-national expenditure:** In the preliminary assessment a stratified sample of sub-national locations are examined and/or visited to collect expenditure data. In order to undertake the sampling in a way that takes into account that revenue and expenditure flows differ between regions, the benchmarking team would respectively order the full distribution of regions and municipalities by general government expenditure per capita, using the latest expenditure data collected by the ministry of finance or local government ministry and census data. If expenditure data is not readily available, localities should be ordered by proportion of the region or locality living below the national poverty line as a proxy for different revenue and expenditure flows, using the latest census data. If fewer than 15 regions, the team should select the mid-point of the bottom half of regions, and mid-point region of the upper half of regions in the list. If more than 15 regions and for cases of fewer than 100 municipalities, the team should select the region or municipality at the mid-point (or first of the two mid-points if an even number of localities per tertile) of the first, second and third tertiles in the distribution. If more than 100 municipalities, the team should select a number of sample municipalities equal to 3 per cent of the total but no more than 10, using the same ordered distribution methodology, and selecting the mid-points of appropriately sized portions of the distribution (e.g. quartiles if 4 municipalities, percentiles if 10 municipalities). Once the sample regions or municipalities are selected, the team should collect data for the sample localities at the national level (if appropriate, e.g. when national data is kept), and if necessary then visit the localities. The data collected should comprise (i)
A survey of all localities would provide the most reliable data on sub-national expenditure. This first component of the exercise is to determine what further work is required to collect a sample that will provide reliable information. The stratification of the sample is to allow a mix of ‘rich’ and ‘poor’ regions and localities. This stratification may be modified if first mapping of child protection services already indicate different means of production in different localities. Stratification could then usefully be by different means of production of sub-national services and relatively ‘rich’ and ‘poor’ regions.

**Step 3: Decision to survey all localities, increase the sample, or extrapolate:** In the next step the benchmarking team would need to decide, based on their observations in the sample localities, (i) whether it is possible to survey all localities through the appropriate national government ministry to collect data for more localities; (ii) whether data for the sample can be extrapolated to a total for all localities based on results in the sample and total expenditure for all localities; or (iii) whether data for additional localities need to be collected to be able to extrapolate.

**Surveying all localities:** A survey of all localities would provide the most reliable data on sub-national expenditure. This would be possible if (i) the survey can be done through a national government ministry; and (ii) the means of production of a service is sufficiently constant across the sample localities so that the benchmarking team is able to isolate exactly which expenditure information to collect through a sample with manageable risk of uneven reporting (for example, requesting all regional civic affairs offices to report on the full time equivalent number of personnel involved in birth registration services as a percentage of full time equivalent staff, in order to estimate total expenditure on birth registration out of the expenditure per office). Even if only a proportion of localities other than the ones sampled report, it would still provide better data on which to base an extrapolation than just the sample localities.

**Extrapolation of expenditure:** Extrapolation at this point would only be possible if the means of production of child protection services and proportion of child protection expenditure to total local expenditure is constant in the sample, to extrapolate the expenditure to all localities using the proportion of child protection expenditure to total expenditure. In most cases additional sampling or a survey would be required.

**Additional sampling:** If it is not possible to survey the remaining localities or extrapolate at this point in the benchmark assessment, the benchmarking team in collaboration with the UNICEF country Child Protection team must decide a country-specific methodology for further sampling in order to allow extrapolation. This methodology should be noted in the presentation of the benchmark.

**Step 4: Extrapolation and collation of data, and noting of confidence levels:** In the final step the benchmarking team would extrapolate the full sample of data collected for the population of localities, collate the data into the national benchmark, and note the confidence levels associated with the assessment.

There may be countries in which general government primary expenditure by region and locality is not available at the central level, either from the finance ministry, local government ministry or central bank. In such cases it would not be possible to identify the sample locations by public expenditure per capita, or to extrapolate the child protection expenditure from the sample locations to the population of locations. The alternative would be to use average child protection expenditure per capita for the sample locations, and extrapolate to all locations. If the sample included more than one location per portion of the distribution of regions per capita, the average for that portion can be extrapolated to the rest of the regions or municipalities in the portion. The assumption underlying this extrapolation – that the average expenditure per child in the sampled locations – would be similar to the average for the population would mean that the confidence level with which the extrapolation can be made would be significantly less, and would require noting this deviation from the standard methodology in the presentation of the benchmark.

---

5 In countries like India and China where local governments run into thousands, the methodology should be applied to cities only if a country-wide benchmark is calculated. The presentation of the benchmark should make clear this deviation from the standard methodology.
D. What expenditure data will be used?

60. The standard preference for the core benchmark will be to use audited outturn data. However, where audited outturn data is not available, is older than 18 months at the time of the assessment, or not available in useful formats, for

i) **domestically financed expenditure** outturn, internal disbursement and then budget data will be used, in that order of preference. Note that internal disbursement data may not be specific beyond the budget holder, and that the default may be budget data if outturn data is not available; and

ii) **externally financed expenditure**, where the programme or project is not included in the budget documentation, development partner disbursement data will be used.

In countries where budget or development partner disbursement data is used as a proxy, this should be noted in the report and in any comparison with other countries. Furthermore, the assessment should provide a view on budget credibility in the country, within the benchmark architecture (insofar as different types of benchmarks use different ways of providing interpretive data) to assist in the interpretation of the benchmark. Note that the proposal is that the choice of data should be consistent across services within a country.

61. Countries should use the same expenditure data set over different assessments, so as to avoid comparing different sets in the presentation of the benchmark. When an expenditure set previously used is no longer available within the time-lapse limits set, or when a better set becomes available within these limits and the assessment is that this will be the case in future, countries may need to switch to a different data set. In such cases this must be acknowledged clearly, and comparisons qualified, particularly if the country experience significant variance between datasets.

62. It is possible that for some services, in some countries, there would be no official expenditure data available. In such cases the proposal is first again to assess the likely size of the expenditure compared to overall expenditure on child protection actions and services, and for expenditures beyond an estimated threshold proportion (5% is proposed) to be estimated using a costing methodology, which is set out in Box 3 below.

**Box 3. Costing services when no expenditure data is available**

<table>
<thead>
<tr>
<th>Given the low transparency – internal and external – of budget and expenditure data in many countries, it is likely that benchmarking teams would need to resort to zero-based costing, i.e. activity-based costing of services where the starting point is not actual expenditure, but an estimate of likely costs based on service-specific research and assumptions. This would require for every service listed in Table 1 and for which expenditure data is not available,</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Identifying cost-bearing activities (undertaking inspectorate visits to places of care), appropriate costing units (or cost drivers) for cost-bearing activities, and the inputs required per unit (including for example staff, transport, office overheads and accommodation costs);</td>
</tr>
<tr>
<td>2) Estimating a cost per unit and the number of units for the year of assessment. Note that the cost per unit would need to be estimated on the best pricing information available (e.g. average public sector salary for the pay level involved), and the estimated number of units per year based on evidence gathered for the level of service provision of the specific service;</td>
</tr>
</tbody>
</table>

---

6 A discussion on different expenditure datasets is available in the Mid-Assignment report in Annex 3.
3) Researching likely capital cost for the service, e.g. the number of vehicles purchased by the inspectorate service and average market price per vehicle;

4) Calculating the cost per service for the assessment year, against the main economic classification items (personnel cost, goods and services cost, transfers and capital cost);

5) Comparing the result of the zero-based costing against expenditure data that are available, and assessing the likelihood of its reliability as an estimate. For example, if the estimated cost of a service turns out to be 20 per cent of the total budget for a ministry, checking with stakeholders whether this is a feasible percentage given other activities undertaken by a ministry. If the estimated cost is significantly out of proportion, the team would need to re-examine and adjust their costing assumptions and data.

As this methodology is resource-intensive, it should only be used in circumstances when a reliable estimate based on available expenditure information using the methodology in Box 4 cannot be derived.

63. Given the specificity of child protection expenditures, together with the likelihood that protection services to children may be bundled in with protection services to the population, it is likely that in many cases the country budget classification system/chart of accounts will not identify child protection expenditure specifically. The benchmark will need to specify accepted methods for apportioning expenditure classified at higher levels, or in cases where services are bundled.

Box 4. Methodology for apportioning expenditure when child protection service-specific data is not available

It is likely that very few child protection expenditures will be classified in budgets and actual expenditure outturns specifically. For most services therefore, benchmarking teams would need to decide a basis for apportioning expenditures within the country's budget structure, to estimate expenditure on the child protection service specifically.

The following rules are proposed:

i) Where child-protection services form part of a larger grouping of services for which expenditures are discrete in the country budget structure, and non-financial performance data is available linked to expenditures (e.g. beneficiary, performance audit or output data) expenditure should be apportioned by the proportion of the beneficiaries or cases that are child-protection specific. This rule would apply to for example apportioning a country’s total expenditure on courts between child courts and other courts by the number of courts that are child courts, compared to other courts. Or total expenditure on mental health services, between child protection response mental health services and other health services by the number of children treated compared to total case loads.

ii) Where such data is not kept, but relevant statistical data series are – such as population or household data – assumptions could be made about the number of beneficiaries that are children or cases that are child-specific as a proportion of all services / cases bundled together. For example, if a country has a unit in the police focused on prevention and response to the trafficking of humans across a country’s borders, and no data is available on the number of cases that involve children, one way of estimating the proportion of expenditure on child protection, is to use the proportion of children in the population to apportion expenditure on the unit.

iii) In many cases however, child-protection specific services would be one of many services funded in a budget category, specifically for countries that do not use a programme budget classification system. For example, the identification of children and families at risk may be one of many activities undertaken by a local social welfare office. In such cases, even if case data is available, it would not allow an apportionment of expenditure as there would be no guarantee that one instance of identifying a family or child at risk, would equal the expenditure of providing a weekly visit to one pensioner. In such cases – and where relevant performance information data is not available -- the proposal is to use staff costs as the proxy for recurrent expenditure on personnel and goods and services, plus
estimated actual capital investment or transfers. The benchmarking assessor would need to interview key respondents (e.g. service providing ministries, departments and agencies (MDAs)) to obtain estimates of the proportion of time spent by staff on child protection actions and services (including all components of the services) compared to other services, and of capital expenditure and transfers.

**Summarising the process for selection of relevant expenditure**

64. The process to identify expenditure that qualify for the core and extended benchmarks, is set out in Figure 3 and Figure 4. Figure 3 shows the sequencing of decisions on whether to calculate the core benchmark only, or also calculate an extended benchmark. Figure 4 sets out the decision tree that assessors will use under these proposals to identify child protection services to include in the benchmark, and to determine the methodology to collect best possible expenditure data, within reasonable cost.

**Figure 3.** Decision tree for determining whether only the core, or also the extended benchmark should be calculated
Section 1.3 Proposal for construction of the benchmark

Core benchmark content

65. The UNICEF reference group for the child protection financial benchmark indicated a preference for a **simple unitary summary indicator**, expanded through the use of interpretive information in its presentation. This section presents options for the indicator content, if a simple unitary summary indicator.

66. A simple unitary indicator (option 2 in the Options Report) is a single indicator but summary in nature, usually a simple numerator/denominator construction with 2 to 4 integer
variables. Its benefits compared to the other construction options presented in the Options Report are that it allows multi-country comparability, ease of presentation, is relatively easy to understand for stakeholders, has a relatively simple methodology, and beyond the cost of collecting data on expenditure (which would be the same for all construction options), will be relatively practical and less costly to apply. Its single greatest drawback is that it is limited in the degree to which it allows child protection expenditure in countries to be compared relative to their context. However, if presented in a standard way with appropriate presentation information, this information can compensate for this shortcoming. Readers are referred to the Options Report discussion of indicator construction options in Annex 3.

67. The table overleaf presents 4 options for indicator content. In all four cases key factors in selecting other variables besides the quantum of child protection expenditure, were:

- **Data availability**: the availability of reliable data for most countries is a key criterion. For all data series used, the data will be available in country, either from the finance ministry and the official budget tables or from the national statistics office. In addition, for all data series proposed data are also available from UNStats, the World Bank development indicator database, or from IMF standardized reports. The preference would be to use country level data from the national statistics office. If there is a significant difference between this data and international data series, the pilot team will need to investigate the reason for the difference, and decide which series to use given the purpose of the benchmark.

- **Presenting child protection expenditure relative to country financial capacity**: Cross country comparisons of expenditure on child protection would be meaningless without also taking into account the likely resources available to governments. The options include two main proposals on presentation relative to country available resources, namely expressing child protection expenditure relative to GDP or relative to government primary expenditure. Each of these in turn can be sharpened in terms of presentation relative to likely government resources, by expression in per capita terms, which takes the wealth of countries relative to their populations into account.

- **Presenting child protection expenditure relative to likely need**: Options against this factor include expressing expenditure on child protection relative to the child population, children in poverty or malnourished children. The proposals below proposes only the child population as a useful statistic for the core benchmark itself, as (i) it would be available for the year of assessment in all countries, even if an estimate, or from UNStats, and (ii) it does not make assumptions about whether across countries children in poverty or malnourished children are more likely to be vulnerable to all of the harms identified equally. Also, it avoids data and comparability complexities related to different measures of poverty across countries.

The team’s assessment is that using expenditure on child protection per child (0-18) relative to government primary expenditure per capita would be the best option for the indicator. It is still relatively simple and easy to interpret, contains no assumptions that would need to be defended, provides a sharp presentation of expenditure relative to country available resources, provides information on countries relative prioritization of child protection services.

---

and will result in more sizable percentages for presentation. Its use however depends on preliminary research on the availability of data to calculate consolidated general government expenditure, or of the indicator itself, at the central level holding true for most countries.
## Table 3  Core indicator content options

<table>
<thead>
<tr>
<th>Indicator option</th>
<th>Calculation</th>
<th>Comments</th>
<th>Data availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Child protection expenditure as a percentage of GDP</td>
<td>Child protection expenditure in local currency units (LCU) nominal/ GDP in local currency units nominal x 100</td>
<td>This indicator will provide a cross-country comparison that takes into account difference between countries in access to resources. The indicator will present small percentages, for many probably less than 1 percent of GDP.</td>
<td>Child protection expenditure as assessed for the benchmark GDP from the country statistics office World Bank Development Indicators also provides GDP in LCU</td>
</tr>
<tr>
<td>2. Child protection expenditure per child as a percentage of GDP per capita</td>
<td>(Child protection expenditure as assessed in LCU/ children 0-18 population estimates for year of assessment) / (GDP in LCU nominal / population estimates for year of assessment) x 100</td>
<td>This indicator will provide a cross-country comparison that takes into account differences between countries (i) in access to resources relative to their populations and (ii) expenditure on child protection relative to the population of potential beneficiaries. The indicator is likely to present small percentages, again for many probably less than 1 percent of GDP.</td>
<td>Child protection expenditure as assessed for the benchmark GDP from the country statistics office World Bank Development Indicators also provides GDP in LCU Population statistics from the country statistics office</td>
</tr>
<tr>
<td>3. Child protection expenditure as a percentage of primary expenditure</td>
<td>Child protection expenditure as assessed in local currency units / (government expenditure – interest payments in local currency units) x 100</td>
<td>This indicator will provide a cross-country comparison that takes into account differences between countries in spending capacity of the state, while also indicating relative prioritisation of expenditure on child protection. Using primary expenditure rather than government revenue minus interest payments, means that donor grants are included in the denominator, thereby providing a below the line figure that can be compared with child protection expenditure including donor expenditure managed by government. The indicator will provide higher percentage numbers than option 1 and 2.</td>
<td>Child protection expenditure as assessed for the benchmark Primary expenditure in LCU is an internationally standardised fiscal framework calculation provided in government budgets and central bank national accounts data² It is also available in IMF Article IV reports, and can be calculated from the World Development Indicators database which provides data on expenditure and interest payments in LCU (but not for all countries for all years)</td>
</tr>
<tr>
<td>4. Child protection expenditure per child, as a percentage of primary expenditure per capita</td>
<td>(Child protection expenditure as assessed in LCU / children 0-18 population) / (government primary expenditure per capita) x 100</td>
<td>Same as for indicator 3; in addition the indicator takes into account cross-country differences in the spending capacity of the state relative to its population.</td>
<td>Same as for indicator 3, plus population statistics from the country statistics office.</td>
</tr>
</tbody>
</table>

---

² In countries where data to calculate consolidated general primary public expenditure are not available at the national level, the average expenditure per locality as a percentage of regional or local gross domestic product or income would be used to extrapolate. This deviation would need to be noted in the benchmark presentation.
**Extended benchmark**

68. As the extended benchmark takes into account society wide expenditure, notwithstanding funder or expender, it would not be appropriate to simply calculate the higher expenditure number as a percentage of government primary expenditure per capita, as all financial flows considered in the numerator, will not be considered in the denominator.

69. For the extended benchmark therefore, the proposal is to use the total extended expenditure on child protection per child, as a percentage of GDP per capita, which is a broader indicator for the denominator. Standard country comparisons will not be part of the extended benchmark presentation, as the reason for calculating the extended benchmark would differ across countries, making comparisons misleading.

**Section 1.4 Presentation of the indicator**

70. The indicator content proposed already provides information on the adequacy of the expenditure relative to potential need, and expenditure relative to country available resources per capita. Further proposals below are standard data that should be presented with the benchmark, in order to frame better stakeholders’ understanding of what it means.

71. The benchmark should be presented with:

72. **A Summary sheet:** The benchmark for all countries should be presented as a summary sheet containing the information set out in Table 4 below. The summary sheet presents the benchmark for the country; key complementary indicators; key dis-aggregations of the benchmark; country relative performance against other countries assessed; as well as key limitations on the calculation of the benchmark that are specific to the country. The benchmark report would include the key limitations of the methodology as a standard box.

73. **Standardised radar graphs:** The summary sheet should be presented in the first year, with single line radar graphs that records the country’s performance.
   - *Radar graph 1:* records performance relative to the top performer of all countries considered if that performance is equal to 100 (radar graph 1, in other words a graphic relative performance index for the country compared to other countries, for each selected indicator). Previous assessments can be shown in subsequent assessments, allowing changes in the country’s position relative to other countries to be shown.
   - *Radar graph 2:* records country performance on public expenditure on child protection aggregate, and for key disaggregations in constant local currency units (LCU). This graph will function as a graphic representation of the quantum and composition child protection expenditure in the country, and progress over time against the country itself. Table 4 indicates which information provided in the summary sheet could be included in which radar graph (see Figure 5).

74. **A Benchmark report:** The benchmark report is the narrative discussion that accompanies the summary benchmark. A tentative table of contents for the report is proposed in Table 5. The purpose of the narrative discussion is to interpret the country data further against (i) particular child protection risks and challenges in the country, and (ii) against other available data for the country, thereby deepening the analysis and strengthening the benchmark as a country-specific policy advocacy tool.
<table>
<thead>
<tr>
<th>Country information</th>
<th>Explanation</th>
<th>Radar graphs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The core benchmark</strong></td>
<td>Country public expenditure on child protection per child aged 0 to 18, as a percentage of government primary expenditure per capita For example: 0.5%</td>
<td>The core benchmark</td>
</tr>
<tr>
<td><strong>Public child protection expenditure</strong></td>
<td>Country public expenditure on child protection in LCU nominal for the year assessed to the nearest thousand For example: LCU 3 405 thousand In countries with significant sub-national expenditure, this can be disaggregated by national level, and sub-national level expenditure</td>
<td>The benchmark expenditure number provided for information purposes.</td>
</tr>
<tr>
<td><strong>Growth in child protection expenditure</strong></td>
<td>From the second assessment onwards, average annual change in constant expenditure on child protection (with the first assessment year as base year for calculating expenditure in constant LCU) as a percentage of average annual change in constant government primary expenditure. For example: 80%</td>
<td>Indicating change in relative prioritisation, if smaller than 100 indicating that expenditure on child protection is growing slower than government available resources, if higher than 100, indicating that it is growing faster.</td>
</tr>
<tr>
<td><strong>Core benchmark disaggregated by prevention and response</strong></td>
<td>Percentage share of public expenditure on child protection on prevention and response services Both numbers can be provided, although the one would be the complement of the other. Radar graph 2 could show both.</td>
<td>Provides information on the composition of public child protection expenditure.</td>
</tr>
<tr>
<td><strong>Core benchmark expenditure on non-service delivery components</strong></td>
<td>Percentage share of public expenditure on child protection on public service functions in the system that support the quality of service delivery, e.g. institutional development, capacity building and accountability functions defined in Table 2</td>
<td>Provides information on public expenditure to regulate, build capacity for and hold government and providers to account for child protection services.</td>
</tr>
<tr>
<td><strong>Assessment of the likelihood of expenditure translating into service delivery</strong></td>
<td>If an A score on PEFA is equalled to 4, a B to 3, a C to 2 and D to 1, with + scores adding 0.5, presentation of the average score on the following PEFA indicators: PFM outturn level (PI-2 Composition of Expenditure Outturn compared to the original approved budget); Measures of the quality of (policy-led) budgeting (PI-5 Classification of the budget, and PI-12 Multi-year perspective in fiscal planning, expenditure policy and budgeting); Measures on the quality of financial management (PI-16 Predictability in the availability of funds for the commitment of expenditures; PI-20 Effectiveness of internal controls for expenditure),</td>
<td>Provides a summary assessment on the likelihood of expenditure translating into service delivery given the quality of de facto public financial management in the country. All scores will be in a range of 1 to 4, with 4 indicating a higher likelihood and 1 indicating a lower likelihood due to issues in public financial management. An alternative would be to use the country’s Country policy and institutional assessment (CPIA) quality of budget and financial management rating from the World Bank.</td>
</tr>
</tbody>
</table>

9 Note that the summary sheet proposed does not include a disaggregation by type of harm. This is in response to the concern that the benchmark should not encourage a return to a silo issue-based approach to child protection.
- Measure on procurement (PI-19 Competition, value for money and controls in procurement); and
- Measures on the functionality of the accounting, reporting and oversight system (PI 25- quality and timeliness of annual financial statements, and PI-28 Legislative scrutiny of external audit reports).

<table>
<thead>
<tr>
<th>Country relative to other countries</th>
<th>Explanation</th>
<th>Radar graph</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country position in comprehensive league table</td>
<td>“Country position of all countries assessed (core benchmark only) in three years prior to the country assessment, where the highest benchmark equals 1” out of “all countries assessed (core benchmark only) in the three years prior to the country assessment. For example: Ranked 7 out of 12 countries assessed</td>
<td>Provides a comparison for the core benchmark against all countries. An option is to differentiate reporting the country’s position by status, with countries positions in alternative tables for countries experiencing emergencies and fragile states indicated or indicated in addition.</td>
</tr>
<tr>
<td>Country position in region</td>
<td>Using UNICEF regions, if available, country position relative to other countries in the region for most recent year of assessment</td>
<td>As the benchmark construction and content already adjusts for differences in per capita income, the next most relevant country grouping is by region, on the assumption that countries within a region would have similar cultures and approaches to child protection.</td>
</tr>
<tr>
<td>Public child protection expenditure relative to average for all countries assessed</td>
<td>Country public expenditure on child protection per child in USD PPP for year of assessment, as a percentage of the average on child protection per child in USD PPP (year of assessment) for all countries assessed in the three years prior to the benchmark. For example: Country public expenditure per child on child protection is 80% of average public expenditure per child on child protection of the 12 countries assessed in same period.</td>
<td>While the league table position provides information on the ranking of the country, this indicator in addition would indicate its relative expenditure compared to other countries.</td>
</tr>
<tr>
<td>Public child protection expenditure relative to other assessed countries in region</td>
<td>Same as above, but using only countries in UNICEF region and for the most recent year assessed.</td>
<td>Shows expenditure per child relative to other countries in the region.</td>
</tr>
<tr>
<td>Public child protection expenditure relative to other assessed countries in income group</td>
<td>Same as above, but using only countries in same income group for the most recent year assessed.</td>
<td>Shows expenditure per child relative to other countries in income group.</td>
</tr>
</tbody>
</table>

Extended Benchmark (only in countries where an extended Benchmark is calculated)
<table>
<thead>
<tr>
<th>Measure</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total child protection expenditure as a percentage of GDP</strong></td>
<td>Total child protection expenditure (public and by additional funders) on children aged 0-18 as a percentage of GDP in local currency. For countries where only a part of the country was considered to be an exceptional circumstance, only expenditure by additional funders for these parts will be included. For example: Total child protection expenditure is 0.04% of GDP.</td>
<td>The extended benchmark</td>
</tr>
<tr>
<td><strong>Total child protection expenditure</strong></td>
<td>The amount of total child protection expenditure as defined in the row above</td>
<td>The total child protection expenditure amount provided as reference</td>
</tr>
<tr>
<td><strong>Public child protection expenditure as a proportion of total child protection expenditure</strong></td>
<td>The amount of public child protection expenditure on children aged 0-18 as calculated for the core benchmark, as a percentage of total child protection expenditure on children aged 0-18 as calculated.</td>
<td>This line would allow comparison of government and total expenditure.</td>
</tr>
<tr>
<td><strong>Growth in total child protection expenditure</strong></td>
<td>From the second assessment onwards, average annual change in constant total expenditure on child protection (with the first assessment year as base year for calculating expenditure in constant LCU) as a percentage of GDP. For example: 80%</td>
<td>Indicating change in relative prioritisation, if smaller than 100 indicating that expenditure on child protection is growing slower than government available resources, if higher than 100, indicating that it is growing faster.</td>
</tr>
<tr>
<td><strong>Total child protection expenditure disaggregated by prevention and response</strong></td>
<td>Percentage share of total expenditure on child protection on prevention and response services Both numbers can be provided, although the one would be the complement of the other.</td>
<td>Provides information on the composition of public child protection expenditure.</td>
</tr>
<tr>
<td><strong>Total child protection expenditure on non-service delivery components</strong></td>
<td>Percentage share of total expenditure on child protection on public service functions in the system that support the quality of service delivery, e.g. institutional development, capacity building and accountability functions defined in Table 2</td>
<td>Provides information on public expenditure to regulate, build capacity for and hold government and providers to account for child protection services.</td>
</tr>
<tr>
<td><strong>Memorandum items</strong></td>
<td>Notes on calculation Any data limitations and adjustments to the standard methodology utilised in country. Explanation of sample size and associated reliability of the sub-national expenditure assessment</td>
<td>Include both in constant LCU in extended benchmark radar graph</td>
</tr>
<tr>
<td><strong>Memorandum items</strong></td>
<td></td>
<td>Include total child protection expenditure in extended benchmark radar graph (see paragraph 75) in constant LCU</td>
</tr>
<tr>
<td><strong>Memorandum items</strong></td>
<td></td>
<td>Include public child protection expenditure in extended benchmark radar graph in constant LCU</td>
</tr>
</tbody>
</table>
Figure 5. Example radar graphs

**Radar graph 1: Country performance relative to other countries (index score)**

![Radar graph 1](example_radar_graph1.png)

**Radar graph 2: Country Public Expenditure on Child Protection (LCU million constant)**

![Radar graph 2](example_radar_graph2.png)

75. For the extended benchmark a version of radar graph 2 should be presented, using the total expenditure on PE, rather than public. A fifth point should be added, namely public expenditure on child protection.
| Table 5  | Indicative table of contents for Benchmark Report |
|---------|------------------------------------------------|---|
| Content header | Content |
| Introduction | Presentation of the summary sheet and standard radar graphs  
Rationale for undertaking the benchmark in country  
History of assessments  
Parameters of current assessment (core and extended or not, year assessed, information on assessors and key government counterparts) |
| Child protection in country | Discussion of key child protection risks  
High level summary mapping of child protection approach and system – legal framework, key actors and relationships between them, roles and responsibilities. In some countries this mapping will be available from previous studies on child protection, and should be re-used if still up to date.  
Key challenges of system (resources, capacity, balance between prevention and response etc) |
| Child Protection Benchmark methodology | Summary chapter on methodology, with full discussion in annex. Standard table on expenditure included and excluded for the country, method of recording or estimating expenditure at national and sub-national level, discussion of key limitations and challenges |
| Analysis of Child Protection Benchmark | Discussion of change in the benchmark between assessments, and drivers of change.  
Discussion of the rankings in the summary sheet, and drivers of why country performance differs from other similar countries; discussion of further rankings as relevant.  
Further disaggregation and analysis of the benchmark, e.g. by types of harm for which services are provided. This disaggregation and analysis can look at  
- government financed expenditure (excluding externally financed expenditure) as a percentage of government tax revenue minus interest payments – this would separate donor financed from government financed expenditure,  
- the benchmark for different disaggregations of the numerator (by types of harm (if useful), by region if primary public expenditure per capita per region is available etc),  
- relative growth rates in child protection public expenditure (absolute and per capita) by type of harm, by region, by service function,  
- if gender disaggregated performance or expenditure incidence information is available for some or all types of expenditure, analysis of percentage share in total expenditure for specific services by gender, or change over time.  
- analysis of the data in terms of the effectiveness and efficiency of public expenditure in preventing and responding to incidence of harm to children (e.g. balance between expenditure on regulatory and service delivery functions, or investment in human resource capacity vs infrastructure etc)  
- discussion of PFM limitations in country, and key factors inhibiting the translation of public expenditure into appropriated targeted service delivery, particularly with regard to the difference between de jure and de facto PFM systems.  

Analysis of Extended Benchmark | For countries that have calculated an extended benchmark, this chapter should be added.  
Discussion of change in the benchmark between assessments, and drivers of change.  
Further disaggregation and analysis of the benchmark. This disaggregation and analysis can look at  
- Discussion of the difference between public expenditure and expenditure by other donors on child protection (extended benchmark).  
- Expenditure by donors as a percentage of total expenditure, for the donors for which the team has data on total expenditure.  
- Discussion on differences between total child protection expenditure on prevention and response actions, and how additional donor expenditure adds to this comparative to public expenditure.  
- Relative growth rates in total child protection expenditure (absolute and per capita) by type of harm, by region, by service function, and how additional donor expenditure adds to this comparative to public expenditure.  
- Analysis of additional donor expenditure in terms of the effectiveness and efficiency of public expenditure in preventing and responding to incidence of harm to children (e.g. balance between donor expenditure and donor expenditure as a percentage of total expenditure) |

\[\text{10 The data in these graphs are fictional and for graph illustration purposes}\]
expenditure on regulatory and service delivery functions, or investment in human resource capacity vs infrastructure etc)

- Discussion of issues in the management of additional donor funding in country, and key factors inhibiting the translation of donor expenditure into appropriated targeted service delivery.

Conclusion

Summary of main findings, key recommendations to address main issues.

Annex

Assessment methodology and standard limitations of methodology
Detail on assessment methodology for each service / function included, and specific limitations for the country

76. These presentation proposals are for the presentation of the benchmark at country level. World-wide, or at the regional level, UNICEF can design a standardized dashboard that presents data from all assessments to date. The dashboard could include:

- The league table for the core indicator for all countries assessed, with disaggregation by region, income level, and status (presenting countries affected by emergencies up to three years prior to the assessment and fragile states at the time of assessment separately from countries in which the child protection system delivers services under normal circumstances).
- Country profiles, providing the summary sheets for each country assessed for each assessment, radar graphs and benchmark reports.
- Scatter graphs that plots countries' performance on the core indicator against relevant other indicators where these are available, for example against percentage of the population living on less than US$2 per day, indicators of malnutrition, child mortality under 5, life expectancy, country HDI ranking. This aspect would only be added to the dashboard once enough countries have undertaken the assessment for the validity of the plots as indicators of correlation patterns.

Section 1.5 Summary of proposed benchmark

77. The table below summarises the proposals for the UNICEF Child Protection Financial Benchmark.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Proposal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose of the benchmark</td>
<td>To enable cross-country comparisons of child protection expenditure as an advocacy and analysis tool for UNICEF country Child Protection Teams</td>
</tr>
<tr>
<td></td>
<td>To enable measurement within country of progress over time</td>
</tr>
<tr>
<td></td>
<td>Potential use as a UNICEF strategic performance measure</td>
</tr>
<tr>
<td>Core benchmark</td>
<td>Public expenditure on child protection per child (0-18) as a percentage of primary public general government expenditure per capita.</td>
</tr>
<tr>
<td>What would count as public expenditure</td>
<td>Government financed expenditure, notwithstanding who provides the service + expenditure financed by local and international actors if government manages the expenditure, notwithstanding who provides the service.</td>
</tr>
<tr>
<td></td>
<td>Audited outturn data, or closest available for year of assessment</td>
</tr>
<tr>
<td></td>
<td>Expenditure for all levels of government, with subnational expenditure calculated if assessed to be more than 10% of total expenditure on child protection at the national level, and estimated through a sampling methodology.</td>
</tr>
<tr>
<td>What would count as primary public expenditure</td>
<td>Consolidated central and sub-national government public expenditure – consolidated interest payments</td>
</tr>
<tr>
<td>What would count as child protection expenditure</td>
<td>Expenditure on all services in the formal child protection system, that can be located within the list of types of harm in the benchmark matrix</td>
</tr>
<tr>
<td></td>
<td>Expenditure on services that could be seen as child protection sensitive, or child protection specific services that are for a type of risk or type of service not included in the matrix, only if the service/expenditure is for children specifically</td>
</tr>
</tbody>
</table>
• Expenditure to be estimated if service not clearly delineated in country budget classifications using a standard methodology
• Expenditure to be estimated using zero-based costing and research price information for the service if no expenditure data available for a service

What happens in countries affected by emergencies, conflict, limited government capacity etc – the exceptional circumstances

• An extended benchmark is calculated, in addition to the core benchmark. The extended benchmark is the system-wide expenditure on child protection per child (notwithstanding funder and service provider) as a percentage of GDP per capita.
• Three criteria, any of which would signal the need to calculate the extended benchmark,
  1) If the country has been designated by UNICEF as experiencing a level 2 or 3 emergency in the three years prior to the benchmark assessment;
  2) If the country is on the OECD DAC list of states affected by fragility
  3) If formal child protection services in the country is traditionally not financed by government, as assessed by the benchmark team in collaboration with the UNICEF country office

What interpretive information will be presented with the benchmark

For the core benchmark:
  1) Country ranking against all countries assessed, and against countries in the UNICEF region
  2) Country expenditure on child protection per child relative to expenditure per child in all other countries, countries in the region, and countries in the same income group
  3) Percentage share of public expenditure on child protection by prevention and response services
  4) Percentage share of public expenditure on child protection of expenditure affecting the quality of services provided (on the development of institutions, capacity building and accountability functions)
  5) Relative growth of public expenditure on child protection, compared to primary public expenditure (2nd assessment onwards)
  6) Summary indicator of PFM capacity (average of key PEFA indicators expressed as numbers)
  7) Memorandum note on key limitations to the benchmark calculation specific to the country

• For the extended benchmark
  1) Comparison of public expenditure on child protection with total expenditure on child protection
  2) Percentage of total expenditure on child protection by prevention and response services
  3) Percentage share of total expenditure on child protection of expenditure affecting the quality of services provided (on the development of institutions, capacity building and accountability functions)
  4) Relative growth of total expenditure on child protection, compared to GDP (2nd assessment onwards)

What additional presentation measures?

• Graphic representation of country performance relative to other countries over time, and graphic representation of country performance for the country itself for the core benchmark
• Graphic representation of country performance relative to earlier country performance for the extended benchmark
• A benchmark report

Section 2 A process for piloting the benchmark

78. This section sets out important considerations for piloting the benchmark.

Section 2.1 Objectives of the pilot

79. The primary purpose of the pilot is to test the methodology for the benchmark, as set out in Section 1 above. The pilot should serve to demonstrate how straightforward, or otherwise, the methodology is to apply at country level; identify methodological adjustments and additional guidance/clarification required; and to
determine whether the resulting indicator is meaningful, can be compared cross-country, and has the potential to serve as an advocacy tool to push for more appropriate fiscal space for CP.

80. At the same time, the pilot should serve to elucidate the requirements of applying the methodology in country, in terms of the time required, necessary expertise, financial cost, and a feasible reporting cycle.

81. The pilot will also provide the opportunity to develop a handbook for the benchmark, and test the handbook, for use in subsequent assessments.

Section 2.2 Scope of the pilot

82. Given that a primary object of the CP financial benchmark is cross-country comparison in order to demonstrate what is affordable and feasible in CP spending, the benchmark has been designed to be applicable across a broad range of country contexts. Testing this broad-based applicability through a pilot would ideally involve a relatively large sample of countries. Pragmatically, this must be limited by the cost implications (although the costs of rolling out an insufficiently tested benchmark should also be considered).

83. The sample does not need to be representative of UNICEF operational universe, instead it needs to reflect the range of country characteristics which are likely to have a significant impact on the indicator. In the pilot set, the following country characteristics should be represented:
   - Include both middle income and low income countries;
   - Include countries that have already undertaken a mapping and assessment of their child protection system;
   - Varying by (UNICEF) geographic region;
   - Varying by type of CP systems (drawing for instance on the Connolly, Katz et al report on a typology of child protection systems);
   - Include at least one fiscally decentralized country;
   - Include at least one country deemed in the methodology to be “exceptional” requiring the calculation of an extended benchmark;
   - Varying by degree of PFM capacity and transparency, to include for instance a country where budget execution data is unreliable or unavailable;
   - Varying by country office size and expertise: one purpose of the pilot is to determine what expertise is required from UNICEF Country Office staff to apply the benchmark, and therefore individuals working on applying the benchmark in pilot countries reflect an array of skill sets, including child protection and social policy specialists;
   - Willingness of the Country Office to pilot the benchmark.

84. The review of comparator benchmarks and indicators under the first output of this assignment served to underline the importance of close cooperation with Government counterparts for sustainability, to ensure access to (sometimes unpublished) budget data as well as to facilitate the impact of recurring analyses on policy and budget frameworks. Therefore we are proposing that only countries with good working relationships with the Government (Ministry of Finance and lead ministry for child protection in particular) and prepared to share budget data be included in pilot. An exception to this may be the “exceptional circumstance” country pilot, which is likely to be an emergency or post conflict situation, where a working relationship with the state may not be viable, and where the public availability of accurate budget data is likely to be highly limited. However the absence of such relationships under these circumstances is unlikely to derail the construction of the indicator given that it will be more geared to non-Government (donor and CSO) funds anyway.
85. To sufficiently cover all these aspects, the set of pilot countries is unlikely to be fewer than three, and would ideally be more. The pilots would not need to occur at the same time; rather it might make sense to have 2 or more phases of piloting, where after each phase the findings and lessons learnt are synthesized and the methodology adapted accordingly, before being piloted in a second phase, in a set of countries with different characteristics.

86. The pilot should also develop a handbook for the benchmark. This handbook should include:

- Guidance to Country Offices on setting up a country assessment, including generic terms of reference for the pilot team (see paragraph 88 below for a discussion on the composition of the team);
- Guidance to the pilot team on undertaking the assessment, based on the proposals in this report. The Guidance should set out clearly the scope of the benchmark; the criteria for the key decisions on calculating an extended benchmark and including sub-national expenditure; the process to be followed; methodologies for identifying public expenditure, estimating child protection expenditure if clear allocations are not present, for identifying additional donor expenditure if an extended benchmark is calculated, and estimating sub-national expenditure; stakeholders to be consulted; and for presenting the benchmark and preparing a benchmark report.
- Tools for undertaking the assessment, including a database tool for collecting data, and guidance on using the tools.

87. The piloting process set out in Table 8 provides for the development of this Handbook.

Section 2.3 Piloting Process

88. Table 8 below offers a tentative indication of the potential phases and duration of the piloting process. The table is based on the following assumptions on the requirements for the pilots:

**Phasing**

- There will be a global preparation phase before piloting in any one country starts. This phase will be used to agree pilot countries, contract the global piloting team, prepare tools for the pilots and prepare the pilot team and UNICEF Country Offices for the pilots.
- When all pilots are complete, there will be a global finalization phase. In this phase the lessons from the pilots will be synthesized and the final Financial Benchmark methodology, guideline, training materials and data collection instruments prepared.
- If the piloting approach is iterative, time would need to be factored in to adjust the methodology and guidance in line with the lessons learnt in between ‘waves’ of country case studies. While this will take more time, it would ensure that the adjustments are tested as well, leading to a more robust end product. The table below assumes an iterative approach with adjustment of tools at least once in the process. Figure 6 below depicts the phasing of the piloting process.
**Team requirements**

- The assumption is that the UNICEF Country Office Child Protection team will not be undertaking the technical work for the pilots, but that a team per country will be contracted. This team will work closely with a global team lead consultant, who will work with UNICEF HQ to develop the methodology and materials for the pilot, and oversee the work of the country pilot teams. This global team lead consultant together with the country pilot teams are referred to as the pilot team. In more detail, the pilot team is envisaged as comprising
  - A global team leader, who will take responsibility for developing core benchmark tools (such as guidelines and database tools), preparing the inception and synthesis reports and adjusting and finalizing guidelines, training materials and tools for use in roll-out. The global team leader will also oversee the work of the country pilot teams.
  - Country teams, which will pilot the Financial Benchmark in pilot countries. A country team will comprise one international consultant, and one national consultant. The global team leader might also act as the international consultant in one or more countries, but it would be important for country international consultants to have prior experience in and knowledge of the country’s budget system and/or child protection system. The national consultant would need to be well versed in the country’s budget and public financial management system, and have good knowledge of Child Protection in the country.
  - If sub-national sampling is likely to be required in country, a country team should be extended by more national consultants to enable parallel sampling of sub-national localities. If these are not appointed sampling will need to be done in series by the two person team, lengthening the time the case study will take. The additional national consultants would not need to attend the global training workshop at the start. They will be appointed based on the initial analysis of the country, and will receive training from the country team leader in-country.
  - Table 7 and Table 8 does not include the additional time lapse or consultant days for a second round of sub-national sampling. If this is required an extra 3 weeks (time lapse) will be needed, and 3 consultancy days per sub-national sample locality will need to be factored in, as well as an extra half day for the international consultant per sub-locality to review and process data. UNICEF may elect to request UNICEF Child Protection or Social Policy Advisors to work alongside a country team if required.

**Time lapse**

- The total time lapse for the pilots and consultancy days required are set out in Table 7 below.
### Table 7  Pilot time lapse and consultancy days required

<table>
<thead>
<tr>
<th>Phase</th>
<th>Time lapse</th>
<th>Consultancy days estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 0</td>
<td>10 weeks</td>
<td>12 days global team leader&lt;br&gt;2 days per additional international and lead national consultant, in other words 4 days per pilot country</td>
</tr>
<tr>
<td>Phase 1-4</td>
<td>12 weeks per country</td>
<td>35 days per international consultant&lt;br&gt;35 days per lead national consultant&lt;br&gt;14 days per additional country consultant&lt;br&gt;Assuming 2 additional country consultants per country, a total of 98 days per country.</td>
</tr>
<tr>
<td>Phase 5</td>
<td>2 weeks</td>
<td>5 days global team leader</td>
</tr>
<tr>
<td>Phase 6</td>
<td>12 weeks per country</td>
<td>As for phase 1 – 4</td>
</tr>
<tr>
<td>Phase 7</td>
<td>4 weeks</td>
<td>7 days global team leader&lt;br&gt;42 weeks (assuming countries run parallel in two country piloting phases)&lt;br&gt;24 days global team leader&lt;br&gt;98 days per case study country (mix international and local consultants)</td>
</tr>
<tr>
<td>Phase</td>
<td>Description</td>
<td>Who</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>-----</td>
</tr>
<tr>
<td><strong>PHASE 0: Global preparation (preparation for all country pilots)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.</td>
<td>Country selection; identification of team and preparation of core instruments</td>
<td>0.1 Initial phase of pilot countries to be selected along the lines of criteria suggested above, and their willingness to cooperate confirmed. Exceptional country(ies) designated and agreed by HQ.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.2 Piloting team to be identified (see under 5 below) and contracted</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.3 Pilot team to prepare standard methodology instruments and guidance, including expenditure database in which to record expenditure, methodology for each set of expenditure information collection, key assumptions etc and which would standardize calculation of benchmark (the Handbook)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4 UNICEF reviews and agrees materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5 Methodology workshop for contracted Pilot team and country UNICEF support officers to familiarize themselves with instruments and methodology Further development of methodology as required</td>
</tr>
<tr>
<td><strong>PHASE 1 – 4 First country pilot(s)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>Country Preparation</td>
<td>1.1 Preliminary mapping of child protection services in the country in the matrix</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.2 Preliminary assessment of the required information that is existing against the services mapped, and what needs to be collected, along the numerator, denominator and interpretation data required. The denominator and interpretation data should be available publically. The budget structure and classification, and need for sub-national data collection, number of initial sub-national pilots; likelihood of a sub-national survey and sub-national budget data availability should be assessed.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.3 Preliminary assessment of ease with which methodology can be implemented, likely methodology adaptions and finalization of country fieldwork plan in an inception report for each country assessed. The inception report may adjust timeframes for fieldwork, given country circumstances</td>
</tr>
<tr>
<td>Phase</td>
<td>Description</td>
<td>Who</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>2. In-country data collection</td>
<td>For CP expenditure data:</td>
<td>Piloting team for country</td>
</tr>
<tr>
<td></td>
<td>2.1 Confirm preliminary mapping of child protection system, and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>identification of ministries with responsibility for implementing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>child protection related activities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.2 Ascertain donor data to be collected and assessed if not on budget;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.3 Collect budget documentation, to begin with corresponding to the whole</td>
<td></td>
</tr>
<tr>
<td></td>
<td>budget cycle (but noting the preference for which audited outturn data) for</td>
<td></td>
</tr>
<tr>
<td></td>
<td>the ministry.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.4 From the budget documents of each ministry and department identify</td>
<td></td>
</tr>
<tr>
<td></td>
<td>expenditure programmes or units relevant to the services identified (</td>
<td></td>
</tr>
<tr>
<td></td>
<td>according to inclusion /exclusion criteria)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.5 From donor data ascertain donor disbursements to include, using inclusion</td>
<td></td>
</tr>
<tr>
<td></td>
<td>/ exclusion criteria</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.6 Interview key respondents in relevant ministries to ascertain information</td>
<td></td>
</tr>
<tr>
<td></td>
<td>on expenditure and means of producing the service;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.7 Apply methodology to apportion expenditure where necessary / apply</td>
<td></td>
</tr>
<tr>
<td></td>
<td>zero-based costing methodology if necessary</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.8 Enter data into preconfigured database spreadsheet, noting sources</td>
<td></td>
</tr>
<tr>
<td></td>
<td>For decentralised governments:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.9 Apply sampling strategy detailed in the body of the report</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.10 Repeat above steps for subnational sample budgets</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.11 Decide on next step (survey, further sampling, extrapolation)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.12 Enter data into preconfigured spreadsheet (3) [assumes the spreadsheet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>will be able to extrapolate results for a whole of country estimate]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>For &quot;exceptional&quot; country pilot(s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.13 Collect information on funding for child protection from all funders</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.14 Enter data into preconfigured spreadsheet (4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>For denominator/ interpretative aid data:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.15 Collect information on spending on other government programs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.16 Enter data into preconfigured spreadsheet (5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>For country pilot(s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.17 Collect information on spending on other country pilot(s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.18 Enter data into preconfigured spreadsheet (6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>For “exceptional” country pilot(s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.19 Collect information on spending on other “exceptional” country pilot(s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.20 Enter data into preconfigured spreadsheet (7)</td>
<td></td>
</tr>
</tbody>
</table>

For “exceptional” country pilot(s):
2.13 Collect information on funding for child protection from all funders
2.14 Enter data into preconfigured spreadsheet (4)

For denominator/ interpretative aid data:
For country pilot(s):
2.17 Collect information on spending on other country pilot(s)
2.18 Enter data into preconfigured spreadsheet (6)
<table>
<thead>
<tr>
<th>Phase</th>
<th>Description</th>
<th>Who</th>
<th>Where</th>
<th>Estimated duration of task</th>
<th>Estimated days contracted</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.15</td>
<td>Verify/identify the data series required and enter into spreadsheet (2). This spreadsheet will also identify likely data sources.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Data cleaning and Calculation Procedures</td>
<td>3.1 Completion of data inputting and cleaning (e.g. correcting double counting in subnational expenditures, or between donor disbursement and country expenditure data) including follow-up through CO to fill in any gaps. 3.2 After the data has been inputted, the calculation of indicator is automatic, and will be displayed on a summary tab.</td>
<td>Pilot team for country</td>
<td>In-country for national consultant(s) Remotely for international consultant</td>
<td>3 weeks</td>
<td>12 days for international consultant 12 days for lead national consultant 2 days for each additional national consultant for follow-up at sub-national level</td>
</tr>
<tr>
<td>4. Reporting</td>
<td>4.1 Preparing the summary sheet, graphs and Benchmark Report. 4.2 Sharing of summary sheet, graphs and Benchmark Report with Country Office, Headquarters, Regional Office and key country counterparts for review. 4.3 Pilot country team produces a short report/aide memoire focusing on the process of applying the methodology (what worked/didn’t, usefulness of guidelines, timeframes, etc)</td>
<td>Pilot country teams and CO</td>
<td>In-country for national consultant Remotely for international consultant</td>
<td>1 week for drafts 2 weeks for review and finalisation</td>
<td>7 days for international consultant 5 days for lead local consultant</td>
</tr>
<tr>
<td>Phase 5: Mid-pilot adjustment of methodology</td>
<td></td>
<td></td>
<td></td>
<td>2 weeks</td>
<td>-</td>
</tr>
<tr>
<td>5. Mid-pilot methodology adjustment</td>
<td>5.1 Global Team leader adjusts methodology, guidance and data instruments based on lessons learnt in first pilots 5.2 HQ reviews</td>
<td>Global team leader</td>
<td>Remotely</td>
<td>2 weeks</td>
<td>5 days for global pilot team leader</td>
</tr>
<tr>
<td>Phase 6: Second pilots</td>
<td>6.1 Country pilot teams undertake phase 4 processes</td>
<td>Pilot country teams and Cos</td>
<td>As in phase 4</td>
<td>As in phase 4</td>
<td>As in phase 4</td>
</tr>
<tr>
<td>Phase 7: Global adjustment of methodology</td>
<td>7.1 Pilots and aide memoires synthesised, drawing lessons and puts together recommendations for methodology review. 7.2 Methodology reviewed. 7.3 Methodology, guidelines and data tools finalised (finalization of the Handbook for the benchmark).</td>
<td>HQ, Regional Offices, pilot teams, Consultants</td>
<td>Pilot team leader works remotely. A lesson learning workshop could be useful</td>
<td>1 month</td>
<td>7 days for global pilot team leader</td>
</tr>
</tbody>
</table>
## Annex 1: Bibliography

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Title</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian Development Bank</td>
<td>2013. The Social Protection Index: Assessing Results for Asia and the Pacific</td>
<td></td>
</tr>
<tr>
<td>Connolly, M., I. Katz, et al.</td>
<td>(2014). Towards a typology for child protection systems. Melbourne, Commissioned through the University of Melbourne by UNICEF, Save the Children</td>
<td></td>
</tr>
<tr>
<td>CRC</td>
<td>(2011). The Right of Child to freedom from all forms of violence, Committee on the Rights of the Child. General Comment no 13 (2011) CRC/C/GC/13</td>
<td></td>
</tr>
<tr>
<td>De Renzio</td>
<td>2009, Taking Stock: What do PEFA Assessments tell us about PFM systems across countries?</td>
<td></td>
</tr>
<tr>
<td>Fernandes</td>
<td>2010, A review essay on child well-being measurement: uncovering the paths for future research</td>
<td></td>
</tr>
<tr>
<td>Foundation for Child Development</td>
<td>2013, NATIONAL Child and Youth Well-Being Index (CWI)</td>
<td></td>
</tr>
<tr>
<td>Helgason, S.</td>
<td>1997 International Benchmarking: Experiences from OECD Countries. OECD.</td>
<td></td>
</tr>
<tr>
<td>Kauffmann et al.</td>
<td>2007. The Worldwide Governance Indicators Project: Answering the Critics</td>
<td></td>
</tr>
<tr>
<td>Kauffmann et al.</td>
<td>2010. The Worldwide Governance Indicators: Methodology and Analytical Issues</td>
<td></td>
</tr>
<tr>
<td>Mackie and Caprio</td>
<td>2011, Assessing the Impact of the PEFA Framework: A Study for the PEFA Steering Committee</td>
<td></td>
</tr>
</tbody>
</table>
Maestral International, 2013, Data Collection on the Structure and Composition of Child Protection Budgets, the Stock and Flow of Cases, and the Profile of Preventative Programmes: Briefing Note #1


OMBI 2011, Performance Management Report
OMBI 2012a, Performance Management Report
OMBI, 2012b, Developing a Survey Instrument for Local Governments: OMBI CMT
OMBI, date unknown. Overview: Ontario Municipal Benchmarking Initiative

PEFA, 2009, Issues in Comparison and Aggregation of PEFA Assessment Results Over Time and Across Countries
PEFA, 2011, PEFA - PFM Performance Measurement Framework
PEFA, 2012, Good Practices in Applying the PFM Performance Measurement Framework
PEFA, 2013. Supplementary Guidelines for the application of the PEFA Framework to SNGs

Save the Children, 2010, Budget for Children Analysis

Seifert et al, 2013, The Open Budget Index (OBI) as a Comparative Statistical Tool

Timothy Besley, Robin Burgess, and Imran Rasul, 2003; Benchmarking Government Provision of Social Safety Nets


UNESCO 2007, Education Counts: Benchmarking Progress in 19 WEI Countries


Annex p2
Annex 2: Mid-Assignment report (literature review)

Introduction

The paper provided in this Annex is the first output in a project to develop proposals for a financial benchmark for child protection (CP), commissioned by UNICEF. The financial benchmark is intended to provide a comparable measurement of expenditure on child protection across countries, as well as an indicator of the adequacy of expenditure. The purpose of this mid-assignment report is to record key perspectives from the literature review undertaken by the team, noting the characteristics of a framework for financial benchmarking and criteria for assessing options against the framework. It is important to note that research work for the benchmark is on-going, and that this report serves to capture information gathered to date, and discuss a framework for the approach to design options for the benchmark.

This Annex is set out in 3 main sections, besides this introduction. Section 1 establishes the team’s understanding of the purposes and likely users of the benchmark; Section 2 sets out the perspectives from the literature review pertinent to the project, particularly on the definition of child protection, how to demarcate a financial benchmark and on good practices in financial benchmarking and comparative assessment; and Section 3 discusses the framework and criteria for assessing options.

Annex 2 Section 1: Intended use of the Child Protection Benchmark

Annex 2.1. The main intended users of the benchmark are UNICEF Child Protection staff, particularly at the country level. However, it is expected that a child protection financial benchmark methodology would have applicability across the sector, and could be used by non-UNICEF stakeholders.

Annex 2.2. The purpose of the benchmark is to provide a standardized approach to measuring expenditure on child protection, in order to assess and monitor the volume of expenditure on child protection, and identify gaps. The aim of the work is to contribute to better child protection outcomes through policy influence with key actors, particularly the state as the primary duty holder in terms of the right of the child. While the focus of the work is on longitudinal assessment of progress across time within a country, a standardized approach will provide cross-country comparability, useful for in-country analytical work and evidence-based advocacy by UNICEF staff with their government counterparts.

Annex 2.3. A related use of the benchmark is to potentially provide a base for a future indicator against which UNICEF could measure its progress in the Child Protection Programme (programme 6). In UNICEF’s Strategic Plan (2014-2017) the overarching outcome for this programme is “Improved and equitable prevention of and response to violence, abuse and exploitation of children”, articulated as eight sub-outcomes around: sexual violence, disciplinary violence, child detention, female genital mutilation/cutting, child

---

11 Interview K Gregson and J Chai
labour, children in residential care, child marriage, and violations against children in armed conflicts. A standardized way of measuring expenditure on child protection across settings would create the possibility of measuring UNICEF’s policy advocacy effectiveness, insofar as expenditure increases are a measure of greater commitment of in-country actors to child protection.

Annex 2.4. The benchmark could also contribute to building the capacity of UNICEF Child Protection teams to engage on budgets for child protection at country, regional and global level. This is an important benefit: while such capacity could be built in the absence of a benchmark and a benchmarking methodology, the benchmark and is research, assessment/calculation and application at the country level will help build knowledge on the financing of child protection, budget/medium term expenditure framework (MTEF) cycle, budget policy and public financial management. At the same time it can provide an entry point for Child Protection teams into these processes at the country level.

Annex 2.5. As the ToRs acknowledge, full benchmarking of child protection cannot be limited to financial data, and a financial benchmark on its own is insufficient to assess the whole child protection system. However, the particular purpose of the benchmark will be to comprehend the financial resources available for child protection and enabling an understanding of how this limits or enables child protection outcomes. At the same time, data in the benchmark could provide a base for understanding the financial implications of reforms in child protection.

Annex 2 Section 2: Perspectives from the Literature Review

Focus of the literature review
Annex 2.6. The team sourced literature along three related axes: (i) the nature and scope of the child protection sector; (ii) issues arising from (public) expenditure and expenditure measurement; and (iii) good practices in benchmarking. A full bibliography is provided in Annex 1 above. In the first group the team prioritized UNICEF documentation, insofar as the primary users of the benchmark would be UNICEF Child Protection teams and the benchmark would need to be relevant to the interpretation of child protection by UNICEF. However, the team also reviewed literature from other sources, including the papers submitted at the New Delhi Conference on Child Protection Systems.

Annex 2.7. On the second theme the team drew on its internal knowledge base, as well as seminal documentation in the field. In the third area, in an effort to draw out key lessons for the child protection (CP) financial benchmark and generate design ideas, a series of comparative benchmarks and indicators have been reviewed, including established benchmarks in the CP and Social Protection (SP) sectors, as well as measurements in the fields of governance, public management and public policy and services.

Annex 2.8. Annex 2 Table 1 below presents a brief description of the form and function of the different comparative benchmarks and measurements which have been included in the review. It highlights the particular aspects of each which are relevant for the design of the child protection financial benchmark, given the envisaged use of the benchmark as set out in section 2 (namely, in-country monitoring and diagnostics).
Annex 2.9. By looking at the purpose, construction and presentation of these benchmarks, as well as their documented strengths and weaknesses, it was possible to extract a number of good practices and key lessons which may inform the child protection financial benchmark. These are discussed against the relevant themes in the subsections below, but a more systematic analysis of the comparative benchmarks can be found in Annex 2.2.

Annex 2 Table 1 Summary of Comparative Benchmarks and Measurements

<table>
<thead>
<tr>
<th>Comparative benchmark / measurement</th>
<th>Description</th>
<th>Relevant aspects for CP Financial Benchmark</th>
</tr>
</thead>
</table>
| Core Indicators for Governance of CP Systems UNICEF East Asia and Pacific Regional Office | A framework of core indicators developed by the UNICEF East Asia and Pacific Regional Office to support the monitoring and assessment of the enabling environment or ‘governance’ of national child protection systems. The set of 37 core indicators are spread across six domains of governance, including PFM, human resources and information management, legislative and policy frameworks, as well as public attitudes to CP. The framework is primarily intended for in-country monitoring and diagnostic purposes, to assess the impact of policies and programmes on child protection systems over time, thereby developing an evidence base for the strengthening of national child protection systems. Some regional comparison is also possible. | Scope of services considered under child protection  
Focus on part of CP system other than results (in this case governance processes, but potential lessons for our focus on financial inputs)  
Use of PFM indicators as they relate to CP (potentially to be used as an interpretation aide in the financial benchmark)  
Treatment of sub-national levels of government  
Flexible approach to application of the framework at country-level. Disaggregation potential |
| (Revised) Social Protection Index for Asia (SPI) Asian Development Bank | Summary measure of national SP programmes and their broad impact on the poor and vulnerable in Asia and the Pacific. The index can be interpreted directly as social protection expenditures per potential beneficiary as a percentage of per-capita poverty line expenditures. | Denominator options  
Scope of services considered under social protection  
Treatment of sub-national levels of government  
Construction of denominator  
Spheres of expenditure (focus on public expenditures)  
Disaggregation potential |
| (Original) Social Protection Index for Asia (SPI) Asian Development Bank | Composite index encompassing four indicators which cover Social Protection expenditure, coverage, distribution, and impact on incomes of the poor. It was designed to support monitoring and diagnostic purposes in country, and to facilitate international comparisons of Social Protection provision. | Scope of services considered under child protection and social protection  
Spheres of expenditure (inclusion of non-public expenditures)  
Weighting of composite indicators  
Disaggregation potential  
Denominator options |
| Human Development Index (HDI) UNDP | Summary measure of three dimensions of human development; life expectancy, access to education and standard of living, as measured by 4 indicators. The HDI’s primary purpose is cross country comparison and ranking. | Weighting of composite indicators  
Utilisation of existing data sources  
Breadth of coverage |
and it sets out to shift attention away from preoccupation with growth as the singular objective of development.

| **Child and Youth Wellbeing Index (CWI)**<br>Foundation for Child Development and Duke University | Composite summary index measuring overall wellbeing in children and youth in America over time. Composed of 28 indicators grouped into 7 wellbeing domains including economic, risky behaviour, social relationships, emotional well-being, community engagement, education and health. The objective of the CWI is to measure changes in overall well-being of children and youth in the United States over time, using 1975 as the base-year. | Utilisation of existing data sources Emphasis on longitudinal data and showing a trend over time |
| **Public expenditure and financial accountability (PEFA)**<br>Multi-donor partnership | The PEFA consists of 28 high level composite indicators measuring operational performance of key elements of PFM systems, processes and institutions, against an ordinal scale of A-D. The PEFA’s primary purpose is in-country monitoring, to inform measures to strengthen PFM systems and to track implementation progress. However a number of research methods have emerged to enable aggregation and therefore cross-country comparison. | Treatment of sub-national levels of government Quality assurance Breadth of coverage Use of PFM indicators as they relate to CP (potentially to be used as an interpretation aide in the financial benchmark) |
| **Worldwide Governance Indicators (WGI)**<br>World Bank | Six composite indicators measuring broad aspects of governance based on several hundred underlying variables taken from perception surveys, and public, private and NGO experts. The purpose of the WGI is to facilitate comparison between countries (including ranking) and over time. | Utilisation of existing data sources Weighting of composite indicators |
| **Open Budget Index (OBI)**<br>International Budget Partnership | Summary index measurement of public access to information on the receipt and expenditure of public funds, developed from the answers to 95 Open Budget Survey questions on the amount of budget information that is publically available. OBI is used to measure the current state of budget transparency in a country and changes over time. Countries are ranked and categorised according to their score. | Weighting of composite indicators Quality Assurance |
| **Ontario Municipal Benchmarks Initiative (OMBI)** | Opt-in collaboration between municipalities in Ontario Alberta and Manitoba to measure performance in municipal service delivery, for the purpose in measuring progress over time comparison with other municipalities for the purpose of sharing best practice. Framework of more than 850 indicators spanning 37 service areas, covering impact, services delivery, quality of services and efficiency. | Self-measurement methodologies and tools Breadth of coverage |
| **Benchmarking Government Provision of Social Safety Nets**<br>World Bank | In an effort to address the question of how much governments should spend on social safety nets, this paper presents an analysis of 63 countries spending patterns from 1972-1997 using a comparative benchmarking methodology, based on three different rankings; a ranking of unconditional expenditures on safety nets; a ranking conditioning on structural features of the economy; and a ranking also conditioning on the quality of institutions within the country. The intended use is for cross country-comparison, to generate a picture as to how states are performing relative to international expenditure norms. | Disaggregation potential Benchmarking public expenditures Using cross-country comparison as an indication of feasibility constraints and what countries ought to be able to do (as opposed to “optimisation” approaches). |
| **Child Budgeting** | The primary purpose of budget analysis from a child rights perspective is to create effective advocacy strategies for seeking more commitment from the state, holding it | Scope of expenditures considered |
The ACPF construct a composite performance to compare and categorise the performance of African countries in terms of their efforts to allocate the maximum amount of available financial resources in sectors that benefit children. Analysis of effectiveness and efficiency of expenditure

Spheres of expenditure (focus on public expenditures)

### World Education Indicators (UNESCO)

The WEI programme helps to assess progress and challenges in national education systems from an international perspective by facilitating international comparison (in particular, comparison of 19 countries participating in the WEI programme against OECD averages). A framework of indicators across 5 thematic areas including: Sources and flows of education expenditure, and levels and uses of education expenditure.

Spheres of expenditure

### OECD Social Expenditure Database (SOCX)

SOCX record OECD social spending, grouped spending by the nature of provision into public, mandatory private and voluntary private social expenditure, across nine different social policy areas.

Scope of services under social protection

Time period of collection

### OECD Child Well Being Indicators

Child well-being indicators measure the quality of children’s lives. These data compares 21 policy-focussed measures of child well-being in six areas, chosen to cover the major aspects of children’s lives: material well-being; housing and environment; education; health and safety; risk behaviours; and quality of school life. Each dimension is a composite of several indicators, which in turn have been selected in part because they are relatively amenable to policy choices. Cross-country comparison is intended to highlight areas of weakness and strength to draw best practice and inform future policy changes.

Presentation and construction


Annex 2.10. The remainder of this section sets out findings from the literature review against key themes for the design of a benchmark, namely

- Child protection, and the scope of the child protection expenditure
- The structure and financing of the child protection sector
- Key issues in measuring expenditure: data validity, availability and classification
- From expenditure to services: the interpretation of expenditure data
- Benchmarking and comparative measurement

### Child protection and the scope of child protection expenditure

Annex 2.11. One of the challenges of a financial benchmarking methodology for child protection is to define child protection and demarcate the related scope of expenditures that will be included in the measurement across different contexts. Overall it is expected that the scope of expenditures to be included would be determined by two factors, firstly whether they are considered child protection expenditures, and secondly pragmatic issues of data collection and reliability. This sub-section of the paper discusses perspectives on the definition of child protection, and the scope of child protection expenditure.

Annex p7
protection and the scope of the sector, and its implications for demarcating expenditures to be included in a benchmark in different settings. The following sub-sections also pay attention to the second aspect.

**Defining child protection**

Annex 2.12. Internationally there is significant convergence in the definition of child protection, circumscribed by Article 19 of the Convention on the Rights of the Child (CRC), which reads:

89. “1. States Parties shall take all appropriate legislative, administrative, social and educational measures to protect the child from all forms of physical or mental violence, injury or abuse, neglect or negligent treatment, maltreatment or exploitation, including sexual abuse, while in the care of parent(s), legal guardian(s) or any other person who has the care of the child.

90. 2. Such protective measures should, as appropriate, include effective procedures for the establishment of social programmes to provide necessary support for the child and for those who have the care of the child, as well as for other forms of prevention and for identification, reporting, referral, investigation, treatment and follow-up of instances of child maltreatment described heretofore, and, as appropriate, for judicial involvement.” (CRC, 1989).


Annex 2.14. The UNICEF Strategic Plan (2014 to 2017) itself targets social protection as the “prevention of and response to violence, abuse, exploitation and neglect of children” (UNICEF 2014). This definition is shared by for example Save the Children (UK) and World Vision (Goldman 2012), as well as the Global Child Protection Working Group for child protection in emergencies (CPWG 2012). However, the inclusion of ‘neglect’ is a recent addition to the UNICEF definition. The previous strategic plan (UNICEF 2005) referred to only the “violence, exploitation and abuse”, as the area of work is still described on the UNICEF website (UNICEF, 2014).

Annex 2.15. Given the latest strategic plan content however, and the inclusion of neglect in the CRC Article as well as the definitions used by other international actors, the proposal is to use the following definition for the purpose of constructing a benchmarking methodology:

“Child protection comprises the prevention of and response to violence, abuse, exploitation and neglect of children.”

**Scope of activities included in child protection**

Annex 2.16. While this definition is clear, defining which services and related expenditures are to be included and excluded poses a methodological challenge, as noted in the ToRs.

Annex 2.17. One of the key issues emerging from this definition is the implications of “prevention” of harm to children (where harm is used as a catch-all phrase for violence, abuse,
exploitation and neglect). The point of departure to demarcate actions or services that respond to harm are instances of harm themselves. This means that defining the scope of child protection from this perspective requires clarity primarily on what would count as instances of harm and tracing responses. Prevention of harm however, is more elastic as a definitional device for child protection – it requires additional demarcation rules besides what counts as harm. Prevention of harm is related to risk, and the degree to which any action or service is likely to prevent the perceived risk. For the purpose of demarcating child protection expenditure for the benchmark, the methodology would need to define the point at which the link from an action undertaken or a service offered is too indirect for the service to be considered part of the child protection sector. For example, a campaign to raise awareness on child abuse is clearly an action to prevent harm to children, as is educating children on their rights or on how and where to report sexual abuse. Social security transfers however, is much less clearly so, but it could still be argued that it prevents neglect of children, by providing incomes to families. However, adult education, skill development and employment-intensive public works programmes, may be considered to fall outside of the scope, even if they are also aimed at raising family incomes.

Annex 2.18. The review of benchmarking approaches identified a number of different approaches to similar demarcation problems. The Asian Development Bank (ADB) Social Protection Index (SPI)\textsuperscript{12} for Asia (in both its iterations) sought to manage the boundary issue by excluding expenditures primarily associated with the health, education or rural development sectors, to focus exclusively on “direct” expenditures associated with social protection. In addition to being clear-cut and therefore simple to follow, this approach has the reported added value of “facilitating the acceptability of the results by national Governments and International Financing Institutions” (ADB, 2008), presumably as a result of their propensity to plan and budget sectorally. However, exclusively focusing on “direct” expenditures ignores a variety of programs and activities that have clear impacts on the protection of children. A strong case can be made for including both direct and indirect expenditures, also because child protection is less clearly a sector on its own than social protection.

Annex 2.19. The OECD Social Expenditure (SOCX) database groups social spending by the nature of provision into public, mandatory private and voluntary private social expenditure, across nine different social policy areas. Clearly, the potential for double counting between these categories is significant. In an effort to avoid this, OECD provides guidance as to which services fall under which categories; and furthermore, required countries to report at programme level. The detailed nature of expenditure data in SOCX constitutes an important form of quality control as the high level of transparency associated with detailed recording limits the scope for inappropriate recording (including double counting) of spending items in SOCX.

Annex 2.20. Under the UNICEF EAPRO regional core indicators for measuring governance of child protection systems the scope of expenditures considered differs by indicator, depending on what makes intuitive sense under each. For example when measuring the existence of a multi-year perspective in fiscal planning, the focus is on overall expenditure totals for all sectors of spending which are usually prepared in an MTEF, as what matters for the sustainability of child protection expenditure is the political willingness and technical

\textsuperscript{12} Social Protection is subject to the same complexity over definitions of boundaries.

Annex p9
capacity of the government to engage in medium-term budgeting and the credibility of budget expenditures for child protection. When measuring the credibility of budget expenditures for child protection, a sector wide assessment is made of all sectors with relevance to child protection, whilst the measurement of the transparency of intergovernmental fiscal transfers for child protection only covers transfers that are linked to child protection (UNICEF EAPRO, 2012). To a degree such an approach could be adopted to the financial benchmark; for example a measure of the quantum of child protection expenditures will need to cover child protection related expenditures only, but an associated measure of the quality of PFM systems, and their ability to translate expenditure into service, could encompass the budget more broadly.

Annex 2.21. At the same time, some common principles for expenditure parameters can be identified, and the literature review provided some potential exclusion criteria (particularly deriving from the Asia SPI). Excluding activities not amenable to quantification (such as legislation on labour standards, children’s rights, as well as consciousness-raising initiatives) was found to significantly simplify the methodology. Furthermore, focusing on large child protection programmes and excluding the smaller, piecemeal ones, helps contain the level of effort required with minimal influence on summary results (ADB, 2011).

From child protection issues to child protection systems

Annex 2.22. Until recently defining which expenditures should be included, and which excluded may have been less of a problem for the construction of a benchmark, as the planning, prevention and response measures to violence, abuse, exploitation and neglect of children have been structured around individual issues (TORs and UNICEF EAPRO 2012). Child protection was focused on monitoring the risks to child survival and well-being, and on specific measures and services to prevent children’s vulnerability to those risks, and respond to their exposure to risks and harm (see the box below for an example from The Gambia on how an issue-based approach can translate into an expenditure mapping).

Annex 2.23. The UNICEF Child Protection Strategy (UNICEF 2008) offered a list of such issues arising from emergency and non-emergency contexts, including: sexual abuse and exploitation; trafficking; hazardous labour; violence; living or working on the streets; the impact of armed conflict, including children’s use by armed forces and groups; harmful practices such as female genital mutilation/cutting (FGM/C) and child marriage; lack of access to justice; and unnecessary institutionalization, among others. As set out in paragraph Annex 2.3, UNICEF’s current strategic goals in the Child Protection Programme are similarly set around specific issues affecting child well-being and survival.
Annex 2 Box 1. Issue-based mapping of child protection budgets

A useful example of how the mapping of expenditure related to specific issues can work is provided from the Gambia, where the UNICEF Country Office in partnership with the Government has undertaken a recent data collection exercise on the structure and composition of child protection budgets, related to the introduction of MTEF and performance-based budgeting. This began with a review of the composition of public spending (wages and allowances, goods and services, subventions and development) on child protection across the relevant key ministries where responsibility for these services fall, including:

- Ministry of Health and Social Welfare, as the primary ministry responsible for child protection,
- Ministry of Justice (Trafficking Unit, Legal Aid Unit)
- The Judiciary (Child Court unit),
- Ministry of Interior (Child Welfare Police, and Juvenile Detention, Immigration)
- Ministry of Basic and Secondary Education (scholarships, subsidies for uniforms for poor children),
- Ministry of Trade, Regional Integration and Employment (Labour Inspection and Child Labour).

This “top-down” approach was complemented by a bottom-up assessment of total number of cases dealt with by various Ministries, Departments and Agencies that make up the child protection system. These service categories include:

- Residential services for all children in any institutions that fall under the mandate of the Ministry of health or justice;
- Community support – including a day care services, child trafficking, child labour services, day care service for children with disabilities, and non-custodial juvenile offending;
- Preventative services – including services that aim to provide safe environments for all children and services that focus on those children who are potentially at risk of harm;
- Family based alternative services – including kinship care, foster care, guardianship, wardship, and adoption;
- Commissioned services – this includes the management and administrative costs of any child protection services that are paid for, or commissioned by, government agencies or municipalities but are provided by local Non-Governmental Organisations (NGOs);
- Civil registration services – including compilation and certification of public documents e.g., for citizenship, Birth registration, and inheritance.

Source: Maestral International, 2013

Annex 2.24. However, it is also realized that the risks to child protection are multiple and inter-related, requiring integrated responses that cut across the sectors of national economies, as well as across sectors of government. Different risks are often inter-linked insofar as vulnerability to violence, exploitation, neglect or abuse in one area, may lead to vulnerability in another area. Specific child protection risks are also related to larger social risks, such as poverty, unemployment and loss of shelter. Taking a narrow issue-based lens to define child protection responses, therefore runs the risk of addressing one issue only without addressing other related or underlying risks; of fragmented, multiple and overlapping responses; and of not addressing systemic risks (such as human resource capacity or adequate strategic and budgeting frameworks in the state to ensure that expenditure is sustainable) that cut across issues affecting the sustainability of any one set of responses or improved child protection outcomes.

Annex 2.25. Over the last decade therefore, important documents and events have signaled a move to a systems approach to child protection, including the UN Secretary-General’s Study on Violence against Children in 2006; the 2007 UNHCR Executive Committee Conclusion on Children at Risk; and the Global Child Protection Systems Mapping Workshop held in 2008 (UNICEF, UNHCR et al. 2012).
Annex 2.26. In UNICEF itself the 2008 Child Protection Strategy introduced an approach to child protection focused on the creation of a protective environment and the strengthening of child protection systems, beyond addressing specific child protection risks. The Strategy described national child protection systems as the

“laws, policies, regulations and services, capacities, monitoring, and oversight needed across all social sectors – especially social welfare, education, health, security, and justice – to prevent and respond to protection related risks.” (UNICEF 2008, p4).

Annex 2.27. The Strategy set out actions that will “reduce children’s exposure to harm by accelerating actions that strengthen the protective environment in all settings” (op cit, p2). It framed a protective environment – comprising the national (i.e. government) protection systems and social norms and behaviours in terms of eight elements, namely,

- **Governmental commitment to fulfilling protection rights**: includes social welfare policies, adequate budgets, public acknowledgement and ratification of international instruments.
- **Legislation and enforcement**: includes an adequate legislative framework, its consistent implementation, accountability and a lack of impunity.
- **Attitudes, traditions, customs, behaviour and practices**: includes social norms and traditions that condemn injurious practices and support those that are protective.
- **Open discussion, including the engagement of media and civil society**: acknowledges silence as a major impediment to securing government commitment, supporting positive practices and ensuring the involvement of children and families.
- **Children’s life skills, knowledge and participation**: includes children, both girls and boys, as actors in their own protection through use of knowledge of their protection rights and ways of avoiding and responding to risks.
- **Capacity of those in contact with the child**: includes the knowledge, motivation and support needed by families and by community members, teachers, health and social workers and police, in order to protect children.
- **Basic and Targeted Services**: includes the basic social services, health and education to which children have the right, without discrimination, and also specific services that help to prevent violence and exploitation, and provide care, support and reintegration assistance in situations of violence, abuse and separation.
- **Monitoring and oversight**: includes effective systems of monitoring such as data collection, and oversight of trends and responses.

Annex 2.28. This framing of a child protection system is echoed by definitions employed by UNICEF’s key global partners in the sector. For example World Vision, a key partner in the 2012 New Delhi Conference on Child Protection Systems, define these systems as “… a set of coordinated formal and informal elements (laws and policies, services, capacities, coordination, accountability, circle of care, child resilience) working together to prevent and respond to abuse, neglect, exploitation and other forms of violence against children”. The Save the Children definition also refers to legal and policy frameworks, budget allocations and coordination, besides services, but also includes oversight and monitoring systems (Goldberg, 2012).
Annex 2.29. The 2009 UNICEF-commissioned study by Chapin-Hall to develop the concept of child protection systems stated that child protection systems operate at several levels, involve several nested contexts, and relies on different actors, including children, the family, the community, and the state, and they can operate at one or more levels. Structures, functions and capacities are the “building blocks” of a child protection system, while “the continuum of care delineates the specific ways in which the system responds to rights violations, whereas the process of care specifies the procedures that are followed when the system engages children, families and communities... the process of care also incorporates assessment strategies, case planning, treatment, and follow up, with the specific processes shaped by whether the underlying services are promotion, prevention, or response.” (UNICEF 2010, p4) Figure 1 below illustrates the multiple components, actors and contexts of a child protection system.

Annex 2 Figure 1. Mapping child protection systems

[Diagram of child protection system]

Source: UNICEF 2010

Annex 2.30. This conceptualisation of the child protection system is applied in the 2010 UNICEF child protection system mapping and assessment toolkit, which sets out to identify the main country child protection risks and to examine how the child protection system might mitigate against these risks (UNICEF 2010). Useful for understanding how this model can be applied to child protection financial benchmarking is the inclusion of a risk profile in the general country information component of the toolkit, as well as the distinction made between a systems overview (a mapping of government and non-government services and structures related to child protection against the risks) and the continuum of care (mapping the protective environment for children from attitudes and customs, through children’s life skills knowledge to protective/preventative actions and services, and responses and interventions). The approach in the mapping toolkit would be useful to identify different factors contributing to a child protection system in a country, and distinguishing between different types of factors some of which may be more important to include in a benchmark than others.

Annex 2.31. Taking a systems approach to child protection in the design of a benchmark, rather than through the lens of specific risks and their related prevention and response...
services, would require the design to consider including expenditures on actions and services in the broader protective environment, in addition to expenditure on direct actions and services to mitigate specific risks or respond to specific harm. It would also require a consideration of the core systems to plan, budget and manage overall public and specific child protection expenditure, to monitor the effectiveness, efficiency and sustainability of such expenditure. At the same time however, the model developed by Chaplin-Hall provides conceptual distinctions within a child protection system that could aid in making choices about which components to include, and how.

Annex 2.32. A key option in this consideration may be about a systems approach providing the lens to identify a broader set of expenditures, but also allowing distinguishing in the treatment of such expenditures within the benchmark between actions and services that promote child well-being (and in doing so prevent harm) or sustain the child protection system, and actions and services that prevent harm more directly, and respond to harm, similar to the distinction made in the EAPRO Child Protection governance indicators discussed in paragraph Annex 2.20 above.

Annex 2.33. There is an additional consideration to be made regarding the scope of expenditures to be considered, given that the intended purpose of this benchmark is to contribute to better child protection outcomes through policy influence with key actors, particularly the state. A likely point of advocacy arising from this benchmark will be to push for bigger fiscal space for CP, by demonstrating what is feasible and affordable. A broad definition that for example includes all costs associated with prevention may undermine this objective.

**Different settings manifest different risks and different responses**

Annex 2.34. In their discussion on designing a universal typology for child protection systems Connolly, Katz et al (2014) place great emphasis on the need for such a typology to be capable of describing different approaches to child protection and different stages of development of child protection systems. The discussion challenges the fit, nature and scope of formal statutory child protection systems to contend with violence, abuse and neglect not committed by family members, including institutional abuse, trafficking, cyber abuse and child labour.

Annex 2.35. While the fact that different settings result in different vulnerabilities and risks is an obvious point in the context of the UNICEF Child Protection Programme which operates in low-income or emergency settings, it is nonetheless an important point for choices in the design of a benchmarking methodology. A demarcation of the child protection expenditures which are issue-focused (relating for example to the UNICEF Strategic Plan list of priority risks) for a benchmark would tell users how much are being spent in different countries on those risks, but not how much are being spent on child protection overall, as the greatest risk in any one specific country may not be included in an issue-specific demarcation.

Annex 2.36. A second important point about child protection systems for a benchmarking methodology from the Connolly, Katz et al discussion is the degree to which systems are regulated by the state and other authority figures, in other words whether systems are more or less formal (reflecting the formal / informal dimension of the Chaplin-Hall model in Annex 2 Figure 1 above). Particularly they refer to difficulties in applying a definition based on typical systems in high-income countries – statutory systems focused on the identification and
A Financial Benchmark for Child Protection

protection of individual children who are abused, neglected, and exploited — to middle and low-income country contexts where such a formal system may not be accessible to children and families, who rely more on community based endogenous practices.

Annex 2.37. A child protection system thus comprises elements that go beyond government. Child protection is not the exclusive remit of a statutory, government-action centric child protection system, but is increasingly conceptualised as being ‘whole of society’ responsibility (Connolly, Katz et al. 2014). This means that protection responsibilities — in a normative and positive sense — also extend to the non-government sector, including civil society, private sector organisations, community organisations and families.

Annex 2.38. Within a whole of society approach most systems also “include less and more formal aspects that co-exist and interact — statutory elements of the system that are regulated by the State and more community-based elements that are derived from custom” (UNICEF, UNHCR et al. 2012, p7). For a financial benchmarking methodology this is an important demarcation, insofar as expenditure on the state and other authority figure regulated elements can be measured, whereas community-based “informal” elements may be far more difficult, even if they form an important component of child protection in any one setting. It may be important however, to take note of more informal community-based responses in one or another way in a benchmark, in some contexts.

Annex 2.39. A final related point is the degree to which child protection actions and services are driven by societal norms and values, and how that affects both the risks that are considered important and the response to those risks within societies. As discussed by Connolly Katz et al (2014, p 5): “There has been considerable development in recent years with respect to the theoretical understanding of child protection systems, much of which complicates the current typologies. Developments include…a much more nuanced awareness of the link between national culture, professional cultures and the nature of child protection in each jurisdiction”.

Annex 2.40. For the design of a financial benchmarking methodology this is important in two ways: firstly insofar as different values may lead to different conceptions of risk, and the recognition that a benchmark designed around specific risks would define child protection exogenously, even if closely associated with international Human Rights instruments. Secondly, societal values shape the orientation of child protection responses. The initial discussion paper on a typology of child protection systems identified the orientation of child protection systems — related to the overall approach of the system to the child in his/her family -- as a key dimension in which systems differ across settings (Connolly, Katz et al. 2014, p 45, 46). It draws on earlier characterisations of systems as oriented between child-protection, community-caring and family service/support, and combines it with whether actions and services are prevention or response oriented, to develop a typology of seven orientations ( punitive, moral instruction/rescue, welfare, communal harmony, child protection, family support, and rights-based child focused). The point for a financial benchmarking typology is not the specific orientations, but rather that in different contexts different kinds of actions and services may count as part of the child protection system, depending on whether a society is oriented towards supporting the individual child and his/her rights, or prioritises families and communities and their rights in their response.
**Child Protection in emergency and post-conflict settings**

Annex 2.41. Child Protection expenditure in emergency and post-conflict settings may require separate consideration in the design of a benchmarking methodology.

Annex 2.42. The humanitarian response in these settings includes specific child protection actions, as a subset of the social protection areas of responsibility agreed under the Inter-Agency Standing Committee (IASC) structures and processes. The Global Protection Cluster of the IASC defines child protection as; “the prevention of and response to abuse, neglect, exploitation and violence against children” (CPWG 2012, p13). While this definition is the same as the definition of child protection in non-emergency contexts, the nature of emergency contexts – namely “the serious disruption of the functioning of a community or a society involving widespread human, material, economic or environmental losses and impacts that exceeds the capacity of the community or society to cope using its own resources” – implies that the regular child protection system cannot cope with the effect of the emergency on child protection, and requires urgent action by external actors. What this action would be (and the risks faced by children) is determined by the emergency and the context (including factors such as the organization and stability of the state prior to the emergency). Important for the design of a financial benchmark is (i) whether specific and new risks arise in an emergency context in any given locality; (ii) what responses would count as child protection expenditure; and (iii) whether and how the expenditure by external actors would be measured as part of the benchmark for an affected country.

Annex 2.43. While (i) and (iii) raise no particular context related point separate to the discussion above of different risks and different actors being relevant in different contexts, what responses would count as child protection expenditure in emergency context does require further discussion. As noted by Murray and Landry (2013) in their paper for the Global Protection Cluster of the IASC, protection – and by extension child protection – can be defined broadly; “Some relate it to the fundamental delivery of humanitarian assistance in accordance with the essential survival needs (food, water, health, shelter) of vulnerable populations. Others place protection within the framework of international legal instruments where the monitoring and recording of violations of international humanitarian and human rights law is used as a tool to confront those responsible in an effort to cause change. Institution building, governance and judicial programmes and deployment of peacekeeping troops are further examples of actions also categorized as protection activities” (op cit, p 4).

Annex 2.44. An alternative to benchmark child protection expenditure in a humanitarian setting would be “a narrower definition of protection, namely activities which have protection as their specific and primary purpose, within the definitional boundaries developed by the IASC” (op cit, p 5). The Minimum Standards for Protection on Humanitarian Action (CPWG, 2012) expands further on these definitional boundaries, offering two instruments by which to identify child protection expenditure in emergency settings: firstly it states “child protection is not the protection of all children’s rights, but refers instead to a sub-set of these rights”. Secondly, it states that “child protection in emergencies includes specific activities by child protection actors, whether national or community based, and or by humanitarian staff supporting local capacities. It also includes activities in other humanitarian sectors that have the effect of improving children’s safety, even where this is not their specific purpose” (op cit, p 13).
Annex 2.45. These two definitional devices (the protection of a sub-set of rights, and including activities in other sectors that have the effect of improving children’s safety even where this is not their specific purpose) have applicability beyond the humanitarian context for the design of a financial benchmark, insofar as it addresses boundaries with regards to what could be considered child protection services.

Annex 2.46. From an alternative perspective, it could be argued that there no need to make a programming distinction between emergency/non-emergency when determining the finances to be included. While the weight of different child protection programmes and how they are implemented may be different in an emergency setting, the overall programme objective is unlikely to change significantly, retaining a focus on services including family tracing and reunification, alternative care, birth registration, detention etc. Any minor changes in emphasis will be sufficiently captured where the country is given due scope to define a context specific definition of child protection.

**Individual children, as well as groups of children**

Annex 2.47. A final aspect around the definition of child protection and the scope of the sector to take note of in the design of a benchmarking methodology, is that while risk or harm may come to individual children and that a child protection system may be oriented primarily to prevent or respond to such risks, such systems may not be set up to respond to risks that face groups of children. For example, research in China noted that there are tens of millions of children who are affected by urban migration. These children are either left in rural areas to be looked after by relatives or family friends, or must accompany their parents and live in inappropriate situations with little access to services (Katz, Shang and Zhang (2011) as cited by Connolly, Katz et al. 2014). Response in such situations based on the identification of individual children would be very challenging, as would be prevention. While the risks presented are clearly about violence, abuse, exploitation and neglect of children, the prevention and response actions may be systemic interventions (such as addressing the migration patterns themselves) that in other circumstances would not be considered child protection actions or services.

Annex 2.48. A briefing note for the Department of Social Welfare in the Ministry of Health and Social Welfare in the Gambia, also makes this distinction between pro, referring to idiosyncratic (or micro) shocks (or risks) which affect only some children, like an abusive household, and covariant (or macro) shocks which affect all children in a group of households or all households at the same time, such as drought. The briefing notes that there are often more mechanisms available to cope with idiosyncratic risks, than covariant risks, as the latter may leave households with nowhere to turn to.

Annex 2.49. Relatedly, there are many references in the literature to specific groups of children being more or less vulnerable to risks related to neglect, violence, abuse and exploitation. Attention is often drawn to children with disabilities, children of migrant families, children in poverty and children from indigenous or minority ethnic groups. As noted by Connolly, Kats et al (2014), recently there is more awareness of the “diversity within most countries and the particular issues that apply to the protection of Indigenous, minority ethnic and migrant children, as well as children with disability”.

Annex p17
Conclusion

Annex 2.50. The sub-sections above highlighted key issues raised by the literature reviewed with regards to the definition of child protection and the scope of the sector:

- The child protection sector comprises actions and services that prevent and respond to violence, abuse, exploitation and neglect of children. These actions and services can be conceptualized as a system, with different components that include the actions and services, but go beyond to include the environment within which the services operate.
- Demarcating child protection expenditures for a financial benchmark presents a definitional and methodological problem, as child protection is both sectoral and inter-sectoral.
- In practice child protection actors have taken an issue-based approach to child protection, with the issues comprising common significant risks faced by children.
- However, the risks faced by children relevant to child protection differ between locations, and between groups of children within any one country.
- Risks are faced by individual children, but some groups of children may be more vulnerable to risks. Prevention and response in such cases may be actions or services which would ordinarily not be considered core child protection actions and services.
- Across countries the child protection differs in orientation – linked to predominant values and the country’s child protection history and current circumstances -- with approaches being more or less child-centric and responses being more or less statutory and regulated by the state. Within countries there may be variations in how different groups in different localities are serviced in a child protection system.
- Child protection in emergency contexts is circumscribed by the definitional boundaries of the IASC cluster system. However, within these boundaries issues around defining the scope of the sector are similar to issues encountered in non-emergency contexts. However, the risks faced by children in emergency situations are particular, and require specific responses.
- In different countries there are different financiers and different providers of child protection actions and services, including the state, international development actors, the private sector and local non-governmental organisations.

The structure and financing of the child protection sector

Annex 2.51. Child protection is financed and delivered by multiple actors in multiple sectors, and does not comprise a clearly demarcated sector in the functions of government. Annex 2 Table 2 below reflects the list of core child protection services drafted by UNICEF in 2008, and where components of such services are commonly placed in sectors across countries.

---

13 The Classification of the Functions of Government -- developed by the United Nations – classifies government expenditure in the national accounts. The framework designates a separate Social Protection division, within which there is a sub-function for family and children, but this sub-function is not broken down further and will not include for example expenditure related to spending in the Public Order and Safety division on child protection services.


Annex p18
Annex 2.52. Within the state child protection actions and services may be financed or undertaken by different levels of government. The General Comment No 13 (2011) by the Committee on the Rights of the Child noted that “References to “States parties” relate to the obligations of States parties to assume their responsibilities towards children not only at the national level, but also at the provincial and municipal levels” (CRC 2011, p4). Which functions are financed and delivered by which level of government, would differ from country to country. While, civic, policing and justice services are often financed, managed and delivered by national governments, social welfare services may be of a local nature due to the size of their benefit area.

Annex 2.53. Different actors are involved in financing and delivering child protection services, resulting in a potentially complex web of money flows within a country, as illustrated in the diagram below.
Annex 2 Figure 2.  Funders and providers of child protection services

Source: Authors

Annex 2.54. For a financial benchmark identifying the funders of child protection related actions and services may be as important as identifying specific actions and services and their related expenditures. For example, if a pure state-centric approach is taken it would exclude the provision of services financed by non-governmental domestic actors or international development partners, the latter particularly in low and middle income settings where development partners are active, as well as in emergency or post-conflict settings. On the other hand the methodology may take the provision of services by the state as a point of departure. In such a case all state-provided services would be included, notwithstanding the funder.

Annex 2.55. Pragmatically – while the methodology of a benchmark would need to ensure that double counting does not occur – it may very well comprise a combination of collecting data on state-financed actions and services by expenditure, and development partner or domestic non-state actor financed services by disbursement of the financing. Again, a methodology would need to distinguish between these flows (state and non-state), and different providers of actions or expenditures (state and non-state) of expenditure from these flows, and may treat them differently in the indicator.

**Lessons from existing benchmarking methodologies**

Annex 2.56. Levels of expenditure included: As noted in the UNICEF forthcoming guidance on child budget tagging, most child-related budget analyses to date have focus on the central or national budget, despite the fact that larger programs that support children often cross multiple levels of government (UNICEF, forthcoming). The literature review has uncovered a number of means by which the different benchmarks have dealt with sub-national levels of Government, but ultimately this is constrained by the availability of budget information and the feasibility of consolidating expenditures across various administrative levels.
Annex 2.57. As a guiding principle, subnational budgets should be included in the analysis only if the child protection activities implemented by subnational governments are significant. In this vein, the revised ADB Social Protection Index (SPI) in its guidance recommends a sampling approach to determine the levels of Government to be included on a country-by-country basis.¹⁵ In countries where the administration of some SP programmes are decentralised (and inadequate information is collected centrally), it was recommended that researchers visit to 2 or 3 local administrations to collect a sample of information on the types of programmes administered locally. Judgment must then be made as to whether these programmes are large enough to affect the summary results, and if they are, additional visits and extrapolation is used to produce national estimates of expenditures and beneficiary numbers.

Annex 2.58. The experience derived from the original SPI indicates that SP programs financed by provincial governments are rarely significant except in the largest countries (ADB, 2011); however that may not be true for child protection. It is possible that a similar sampling approach could be employed under the child protection financial benchmark, on a case by case basis. It should be noted that compared to the SPI, the level of effort associated with sub-national monitoring is lower because the focus is exclusively on expenditures, rather than expenditures and beneficiaries.

Annex 2.59. The UNICEF EAPRO regional core indicators for measuring governance of child protection systems are designed to be applied to the national level, but this is complimented by a few indicators which specifically assess the effectiveness of decentralization arrangements salient for child protection outcomes. Whether a similar approach would work for the financial benchmark (for example, by focusing on national level child protection expenditures and inter-governmental transfers, whilst including an indicator that assesses the horizontal allocation of transfers linked to expenditures on child protection among sub-national governments) would depend on the degree of decentralisation and therefore the incidence of responsibility for child protection services within a Government, which will vary by country.

Annex 2.60. **Spheres of expenditure included**: In considering whether non-public finances should be captured in the child protection financial benchmark, the literature review offers some useful comparisons. Some indicators, such as the Open Budget Index (OBI) and the Ontario Municipality Benchmarking Initiative (OMBI) are deliberately restricted to the public sphere, as is fitting for their purpose of improving the accountability of public bodies. The core indicators in the PEFA focus in on public financial management under the auspices of the government, however there is an additional set of indictors on donor practices intended to capture elements of donor practices which impact the performance of country PFM system (PEFA, 2011).

Annex 2.61. UNESCO’s World Education Indicators (WEI) are intended to cover the entire national education system regardless of the ownership or sponsorship of the institutions concerned and regardless of education delivery mechanisms. However at the time the report

¹⁵ Notably, however, in the 2011 data collections process for the SPI, funding was inadequate to finance national consultants to engage directly with subnational units, such as state or local governments. (ADB, 2013)
was published, fewer than one-half of WEI countries provide data on private expenditure on educational institutions, and only six can do so by level of education. This information gap meant that reported expenditure on education can be greatly understated in countries where the private education sector is sizeable or tuition fees are common.

Annex 2.62. The original SPI included Government, private sector and NGO expenditures in the calculation of the component summary indicator on social protection expenditure (SPEXP), but presents problems associated with data collection from highly disparate and disbursed entities (ADB, 2008). In its revision the SPI restricts its assessment to government or government-supported social protection programs, but includes international development agency funding to those programmes (ADB, 2011). The motivation for this change in approach was the assumption that governments should take ownership of their social protection systems and be held accountable ultimately for the success or failure of those systems (ADB, 2013). Going down such a route would require a clear definition of what constitutes a “government-supported programme”, and might risk excluding significant volumes of child protection expenditures in countries where aid operations operate largely outside of Government systems and oversight (for example, in some fragile countries).

Measuring expenditure, key issues
Annex 2.63. Child protection budgets do not equal expenditure, and expenditure does not equal service delivery to the expected target group. A key issue in this regard is which expenditure data is kept in what form by different funders and providers of child protection services and therefore the data that is sensible to collect.

Status of expenditure data
Annex 2.64. Four options with regards to the status of the expenditure data collected present, as set out in the table below. The table analyses each option against the validity of its use for financial benchmarking, and against the likelihood of this type of expenditure information being readily available to UNICEF country offices. For this analysis we make use of publicly available data for 79 countries from the Public Expenditure and Financial Accountability assessment methodology, as recorded in their latest assessments by the end of 2012. The table provides a calculation of the average assessment across these 79 countries as a proxy for the likelihood of data being available in different settings, where 1 would equal a D score in the PEFA methodology (i.e. the worst possible score) and 4 an A score (i.e. the best possible score).

Annex 2 Table 3 Types of expenditure information and their use in financial benchmarks

<table>
<thead>
<tr>
<th>Type of expenditure data</th>
<th>Validity for financial benchmarking</th>
<th>Likelihood of its collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Budget data: Planned expenditure as reflected in the budget documents of governments</td>
<td>Budget data would compare the stated intent of different funders of child protection services, but not whether the services are actually funded. The average PEFA score against indicator PI-2 (which assesses variance of PEFA data of some nature is the most likely to be available for government, development partner and non-government sector financing. For government funded programmes for example, at the national level the average PEFA score for the 79 countries with publicly available PEFA Assessments by the end of 2012, on Indicator P16, which measures the coverage and quality of the executive’s</td>
<td>Budget data of some nature is the most likely to be available for government, development partner and non-government sector financing. For government funded programmes for example, at the national level the average PEFA score for the 79 countries with publicly available PEFA Assessments by the end of 2012, on Indicator P16, which measures the coverage and quality of the executive’s</td>
</tr>
</tbody>
</table>

Annex p22
<table>
<thead>
<tr>
<th><strong>Actual expenditure to the composition of budgets</strong> for the 79 countries is 2.3, reflecting a score just above the C level. Overall over 20% of countries scored below the C level, indicating significant lack of budget credibility.</th>
<th>budget proposal was on average just above a B score. Overall, 60 countries scored at a B or A level. While this does not imply that the information is publicly accessible, it does imply that on average in most countries there is some budget information with which to work.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Disbursement data</strong></td>
<td>For child protection services provided by subnational units and/or non-governmental organizations, disbursement data from funders (government or other sources) may be the most reliable as a proxy for expenditure data. For government funds, disbursements from government ministries and agencies to organizations, structures or units actually providing services will equal actual transfer data (part of actual expenditure data). For development partners, disbursement information would need to be collected donor by donor and programme by programme at the country level, as disbursements for child protection would not necessarily be reflected in existing aggregate donor reports or the Government’s aid management reports.</td>
</tr>
<tr>
<td><strong>Actual expenditure (outturn data)</strong></td>
<td>Actual expenditure information is far more difficult to collect. It would require that country accounting and reporting systems are timely, comprehensive and provide good quality information to the centre of government for consolidation. A proxy PEFA indicator on countries’ capacity in these core financial management functions is Indicator PI 24, which measures the quality and timeliness of in-year budget reports. For the 79 countries the average score is 2.5, or a C+. Again, this does not mean that actual expenditure reports are publicly available — data may need to be collected from individual ministries or the finance ministry. However, about a fifth of countries scored below a C, indicating that actual expenditure data are not readily available.</td>
</tr>
<tr>
<td><strong>Audited expenditure</strong></td>
<td>Audited expenditure data would be the most valid type of expenditure data to use. Audited expenditure data would reflect whether the funds were actually expended, rather than just reported to be expended. For child protection services provided by government this would require the availability of audited appropriation accounts, rather than just audited consolidated financial statements of government, as the latter may not disaggregate by administrative units or sectors. Even audited financial statements may be challenging: for the 79 countries the average PEFA score on the timeliness of submitting audited statements to the legislature is 2.5, with 23 countries scoring below a C, indicating high likelihood that audited information will not be available.</td>
</tr>
</tbody>
</table>

Source: Author and PEFA data (PEFA, 2014)

Annex 2.65. The UNICEF EAPRO Child Protection governance indicators take account of the differences between budgeted and actual data, in the PFM indicators which applies the PEFA PI 2, comparing budget and outturn data in key child protection sectors (see paragraph Annex 2.20 and Annex 2.2).
Annex 2.66. The status of expenditure data to be considered could differ by country, according to a scale of preference; wherein audited expenditure is the preference, followed by outturn, then disbursement and lastly budget estimates, which would be qualified according to the latest PEFA score.

**Classification of expenditure data**

Annex 2.67. The assessment in Annex 2 Table 3 does not take into account whether the classification of expenditure will allow specific child protection services to be tracked. A robust classification system allows the tracking of spending on the following dimensions: administrative unit, economic, functional and programme. Within each of these dimensions greater or lesser degree of details may be present, as illustrated in Annex 2 Table 4 below:

### Annex 2 Table 4  Government expenditure classification dimensions

<table>
<thead>
<tr>
<th>Design and purpose</th>
<th>Economic Classification</th>
<th>Functional classification</th>
<th>Programme Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Determined by country, offering classification of expenditure by country administrative structures. Naming of levels determined by country. A ministry may cover more than one sector, or any one sector may be covered by more than one ministry.Used for expenditure control and often for budget allocation</strong></td>
<td>Government Finance Statistics (GFS) set international standard, can be aggregated across government administrative, functional and programme classifications. Used for economic policy purposes and analysis In some countries used for budget allocation and control</td>
<td>UN COFOG sets international standard, also included in GFS May cut across government administrative units and programmes Can be aggregated across government administrative units and programmes Used for policy and analysis purposes, and the national accounts</td>
<td>Determined by country, offering classification of expenditure within a ministry, department or agency in line with objectives Unique to every ministry and cannot be aggregated across Used for budget allocation and expenditure control in some cases, in others only a presentational device</td>
</tr>
</tbody>
</table>

| Typical disaggregation | |
|------------------------| |
| **Ministry** | Main category (compensation of employees, use of goods and services, transfers, capital) | Function (division), e.g. | Programme |
| **Department** | Category (E.g. in goods and services: office expenses, transportation, materials and supplies etc) | Sub-function | Sub-programme |
| **Sub-departments, divisions or sections** | Sub-category (e.g. in office expenses: uniforms, gifts, telephone etc) | | Project |
| **Units or cost-centres** | Class, sub-class and details | | |

Source: Authors and Jacobs, Hélis et al. (2009)
Annex 2.68. The budget will be presented in a format that reflects the most important classifications for either budget control purposes, or policy purposes. Traditionally this is by administrative classification (budget control), combined with economic classification. Increasingly countries allocate to programmes rather than administrative units, or in addition to allocating by administrative units, they present budgets by programmes and functions, to reflect better the policy intent of the budget. If such budget classifications are not also embedded in the chart of accounts (e.g. a programmatic classification in some cases which is not also in the chart of accounts for accounting purposes) transactions cannot be reported in accordance with the classifications used in the budget document.

Annex 2.69. In order to collect expenditure information on child protection, budget and/or outturn information would need to be presented in more detail than just the main administrative units. This may in some cases be offered by a programme classification, which would provide information collected in line with specific objectives, or by a detailed administrative classification which would allow expenditure on specific units of government to be isolated.

Annex 2.70. In a recent study UNICEF assessed the public budgets of 109 countries, classifying the detail in which information is provided in five categories, namely by

- Ministry (which may cover several sectors)
- by individual Sector
- by sub-ministry/agency level or sub-sector
- by individual programme or project
- by individual programme or project, but also with expected beneficiaries.

Annex 2.71. It found that 59 countries reflected expenditure at the ministry or sector level only, or not for all of government, or do not provide information at all. The pie chart below provides the information for the 109 countries reviewed.

Annex 2.72. The study then tracked broad areas of child-focused spending in the 80 countries for which all information was available. Annex 2 Figure 4 below shows how many of the three areas could be identified in each country, and the level of detail of the reporting. In brief, UNICEF found that it would have been possible to track health spending in all 80 countries; however, this number falls to 53 once child-related education is included; and to only 35 when child-focused social protection information is needed. When more specific child-focused spending is required – ie when broad allocations to groups of beneficiaries which may include children or children at risk need to be broken down to specific child protection expenditure – the number of countries that could be tracked dropped to five countries, namely those who provided programme-level information with information on beneficiaries and outcomes.
Annex 2 Figure 3. Comprehensiveness and level of detail in country budgets (UNICEF study)

Source: UNICEF (forthcoming)

Annex 2 Figure 4. Tracking broad child-focused spending areas

Source: UNICEF (forthcoming)
**Time period of collection**

Annex 2.73. A final issue to consider is whether annual expenditure data will be collected for only one year, and how to reconcile different fiscal years across different budget cycles. Expenditure on child protection is by its nature data on financial flows. However, the benchmark can total or average expenditure over a number of years, to compensate for the risk of cross-country or longitudinal comparisons selecting an outlier year for any one country in either the numerator or denominator of an indicator (used on its own or as component in a benchmark). Alternatively, it can assess expenditure against a benchmark for any proportion of a set number of years, for example, to mitigate against the same effect in a measure that scores or indexes against a standardized benchmark across countries, or within a country.

Annex 2.74. The PEFA methodology for example, uses a combination. Most of the indicators are calculated for one year, but indicators PI-1, PI-2, PI-3 and D-1 require data for three years as a basis for the assessment. The data should cover the most recent completed fiscal year for which data is available and the two immediately preceding years. The assessment is based on the performance in two out of those three years i.e. allowance is made for one year to be abnormal (and not contributing to the score) due to unusual circumstances such as external shocks. The OBI on the other hand, standardizes across countries on one year.

Annex 2.75. If the intent is to undertake cross-country assessments at the same time, the standardization of fiscal years across different budget cycles could become an issue. The OBI – which undertakes the research for all the participating countries at the same time – manages the issue by designating the fiscal year for different budget cycles. Other benchmarks simply use the latest year measured. Enabling cross-country comparison within a defined time period (such as five years) might be a pragmatic way forward.

Annex 2.76. This discussion highlights the methodological problems presented for a child protection specific financial benchmark, given the sophistication of government budget classification in practice, and the cross-sector (and cross-ministry) nature of child protection services. This would be relevant for the level of complexity in terms of demarcating expenditure that can be built into a workable benchmarking methodology.

**From expenditure to services: the interpretation of expenditure data**

Annex 2.77. The degree to which comparable amounts of expenditure will reliably and sustainably translate into the provision of quality child protection services targeted at the right population group is not consistent across settings. Just an expenditure number would not provide sufficient information on whether the expenditure is efficient, effective, reaches the right target group and is sustainable.

Annex 2.78. The ToRs however are clear in specifying the scope of work to be the development of a financial benchmark for child protection, whilst recognising that this on its own would be insufficient to assess the whole child protection system. The core issue is that a financial benchmark should focus on child protection inputs and not results. This is similar to the UNICEF EAPRO child protection system governance measuring framework which
focuses on processes relevant to child protection and not on outputs of outcomes such as prevalence of maltreatment.

Annex 2.79. Focusing on the child protection inputs however does not necessarily mean foregoing non-financial and financial information in a benchmarking methodology that will facilitate interpretation of the financial information. The design of the benchmark can ensure that it is primarily a financial benchmark, but that the amount of expenditure is interpreted with the provision of additional information on

- the quality of expenditure, such as the balance between investment in human resources, goods and services and capital investment, to ensure an able workforce or sufficient non-personnel funding so that services can be delivered;
- the effectiveness of expenditure in relation to objectives and target groups, with distinctions as the design requires between expenditure in relation to the target beneficiaries, or the number of actual recipients reached;
- the governance of expenditure, namely the existence and functioning of systems to plan, budget, disburse, control, expend, account for, monitor and evaluate expenditure on child protection and overall.

Annex 2.80. Of course, information just on the functionality of the PFM system is only aspect of the broader governance environment for child protection. The UNICEF EAPRO child protection governance indicators look at several domains of governance, namely the policy framework; PFM; human resources for child protection; statistics and information; child protection systems coverage; surveillance, gate keeping, referral and quality assurance; and public attitudes and values. However, of importance to a financial benchmark would be the governance of financial resources, which means a focus on the quality of the underlying PFM system. The UNICEF EAPRO child protection governance framework measures the following PFM aspects:

- Credibility of the budget allocations across child protection sectors (likelihood of budgets being executed as planned)
- Robustness of the budget classification (functionality of the budgeting system to allocate resources to objectives)
- Transparency and realism of intergovernmental fiscal relations (the governance of sub-national revenue and expenditure)
- Multi-year perspective in budgeting (the functionality of the budget system to prioritise resources over the medium term, and be sustainable)
- Policy-based budgeting (the functionality of the policy-budget link, enabled through [sub] sector resource-linked planning
- Conduciveness of public financing framework (the incentives offered by the framework for financing child protection expenditure in any one country)

Annex 2.81. From a practical perspective for UNICEF Country Offices, and for the application of the benchmark, including the assessment of the public financial management system within which child protection expenditure is utilized would provide an entry point for

---

16 Whilst this is a deliberate and defining feature of the framework, it has been "received with caution" as a result, and it would be reasonable to expect a similar reception for the financial benchmark.

Annex p28
Child Protection Staff into the budget process through the acquisition of knowledge and insight on the PFM system. This would be of importance for the utilization of the benchmark, as it is the budget process within which advocacy for increasing expenditure – or linking expenditure better to child protection policy objectives and outcomes – should take place.

Annex 2.82. The Africa Child Policy Forum presents a framework for budget analysis from a child rights perspective which incorporates an assessment of the efficiency and effectiveness of budgets in terms of bringing about concrete child wellbeing outcomes. Ideally, efficiency and effectiveness analysis would consist of an in-depth assessment of the multitude of systemic, management, capacity, corruption and attitude-related factors that cause inefficiency in the utilisation of budgets. However, such an assessment would require extensive data and information on administrative and process-oriented aspects, which is more often than not unavailable in the African countries for which the framework was developed. As an alternative, budget expenditures can be evaluated for efficiency and effectiveness in terms of objective child wellbeing outcomes achieved and services delivered. In the absence of results-based budgeting in most countries, an approximation of this is derived from a comparison of budget expenditures against objective child wellbeing outcomes, collecting from existing data sources. This provides an indication, albeit in crude form, of whether budgets for children are well targeted and efficiently utilised to result in concrete outcomes reflected on children themselves.

Annex 2.83. Key options in the design of the indicator for ensuring the interpretation of the expenditure data relative to country context and outcomes are discussed in the next subsection.

Benchmarking design and application
Annex 2.84. A final set of issues from the literature review is the design of benchmarks (how are they constructed), and their application (what are the processes for their calculation and use), in order to draw good practice lessons for the child protection financial benchmark. The following aspects are reviewed:

- Overall construction
- Use and choice of denominators
- Other interpretive mechanisms
- Disaggregation potential
- Flexibility for and coverage of different country contexts
- Scope / practicality considerations
- Quality assurance

Overall application
Annex 2.85. The benchmarks reviewed fall into two main categories: benchmarks that are constructed for in-country use, and therefore utilizes a country-determined time-table of data collection and assessment/calculation, and benchmarks that are constructed to enable ranking of countries, which for the most part are applied in a universal, cross-country process at the same time, to support comparability between countries. A consideration of the approach for the UNICEF child protection financial benchmark – with higher emphasis on in-country monitoring than cross-country ranking -- would be a key input in the identification and assessment of design options.

Annex p29
**Overall construction**

Annex 2.86. A financial benchmark can be constructed in a number of ways, as demonstrated in the table below:

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>UNITARY BENCHMARKS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unitary indicator</td>
<td>A single numerical integer</td>
<td>Expenditure on social protection in US dollars</td>
</tr>
<tr>
<td>Unitary summary indicator</td>
<td>A single indicator but summary in nature.</td>
<td>A numerator/denominator construction, either simple (expenditure on social protection as a % of GDP, e.g. the WB benchmarking of social safety net provision) or complex (revised SPI)</td>
</tr>
<tr>
<td>Unitary index</td>
<td>A single statistical measure of change/ relative standing to a selected benchmark</td>
<td>Expenditure on social protection as a % of GDP expressed as an index of the average across similar countries, or with a specific year as a base point for all countries.</td>
</tr>
<tr>
<td>Unitary composite index / indicator</td>
<td>A summary indicator calculated from a standardised set of components</td>
<td>Components can be an average of indexes (HDI, the Child Well Being Index), or a series of quantitative assessments with a standardised scoring methodology (OBI), or a series of unitary summary indicators (the original SPI). The calculation of the composite index or indicator can weight different components equally, or not.</td>
</tr>
<tr>
<td><strong>FRAMEWORK OF MULTIPLE INDICATORS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Framework of standardised indicators/indexes,</td>
<td>A framework of indexes/indicators, with a standardised scoring methodology against each component, but not summarised into a single value</td>
<td>Another option is to use a framework of different indexes or indicators, but these are not summarised. However, the indicators are calculated in a standardised scoring methodology. The indicators can be numerical (WGI) or an ordinal scale (PEFA).</td>
</tr>
<tr>
<td>Framework of indicators/indexes, but the scoring methodology is not standardised</td>
<td>A framework of indexes and indicators, that measure in a standardised way for each indicator, but different scoring values and calculation methods are used for each indicator. Some may be numerical and others not.</td>
<td>Examples are the Ontario Municipal Benchmarking Initiative (OMBI) and the EAPRO child protection governance indicators</td>
</tr>
</tbody>
</table>

91. Source: Authors based on Annex 2.2

Annex 2.87. The analysis in Annex 2.2 suggests a trade-off in the choice between unitary measures and a framework of measures between suitability for cross-country and in-country use; and between ease of interpretation for cross-country or longitudinal comparisons, and richness of analysis for diagnostic purposes. While a composite unitary measure can provide more breadth of measurement for in-country and diagnostic purposes, it is difficult to find a
credible basis for weighting of the different components in the summary measure, undermining its credibility. On the other hand framework measures can result in high cost and effort of data collection, depending on their size and the data required for each indicator. Composite frameworks that score in a standardized way are also easier to interpret and track change more easily, than indicator frameworks that allow variability in scoring.

Annex 2.88. The next output in this assignment will undertake more thorough analysis of the different design options, against the criteria agreed based on this first output, and with content proposals based on the child protection sector.

**Use and choice of a denominator**

Annex 2.89. The TORs highlighted the choice of a denominator as a key issue for the assignment. It is of course particularly pertinent to unitary measures that use a or more than one numerator/denominator-based indicators. Furthermore, should such (an) indicator(s) feature in the design, the choice of denominator would be an important interpretive mechanism, for example to frame the amount of expenditure against the country context, and build comparability between countries.

Annex 2.90. An obvious choice for a denominator would be to relate an amount of expenditure to GDP, or to overall government expenditure, thereby making different quanta of expenditure relative to country capacity to fund. The costing of child protections systems done to date however, has shown expenditure to be very small percentages of budget and even smaller of GDP, limiting the interpretability of the indicator as the spread is likely to be narrow across countries.

Annex 2.91. Furthermore, as noted in the ToR, a central drawback of using a denominator such as GDP/GNI for a benchmark such as this, is that it is too far removed the target group and so does not reflect the impact on the child or family. A similar criticism was levied at the original SPI, wherein the expenditure component indicator (SPEXP) compared the magnitude of SP benefits to the size of GDP instead of the number of beneficiaries, hence only giving an indirect indication of the impact of the expenditures on human well-being. It also lacks normative significance as no one would argue for increasing child protection spending to the levels of GDP (ADB, 2008 & 2011). In the revised Asia SPI, the regional poverty line is used as the denominator, so that the index can be directly interpreted as SP expenditures per potential beneficiary as a percentage of threshold of total expenditures that each person needs to exceed to be considered non-poor (ADB, 2011).

**Other approaches to interpretive information**

Annex 2.92. Choices on the design of the benchmark would be important to allow additional information to the amount of expenditure to be provided in a consistent way across different contexts, thereby aiding interpretation.

Annex 2.93. The choice of PFM indicators in the UNICEF EAPRO child protection governance indicators provide an example of adapting the tested PEFA methodology for a light-touch application to child protection sectors specifically, to assess the quality of the underlying PFM systems. While budget and public financial management functionality is not sufficient to ensure the delivery of quality child protection services, it is necessary component of the overall system. If the paucity of reliable, timely actual or audited child-focused
expenditure data forces the use of budget data, having comparable information to assess the likelihood of the expenditure being executed as planned, would be particularly critical. Also, from a UNICEF Country Office perspective, including budget system assessment in the benchmark design itself, would at the same time engage UNICEF child protection staff in budget processes and issues, which is the key system within which advocacy for higher and better quality expenditure on child protection would need to occur.

Annex 2.94. The World Bank safety net benchmarking exercise (2003) offers another example of providing a core indicator, but with additional interpretative information. The methodology is composed of three interrelated ranking of public expenditures on social safety nets (one of Government spending, one of spending controlled for structural-economic factors, and one controlling for structural-economic factors and the quality of Government and strength of the institutional framework. Given this construction, it is possible to decompose the policy advice to governments into two forms—first, the effects of changes in the structural characteristics of the economy, and second, the effects of changing institutional features (Besley et al, 2003), and there are a number of options to include different information sets within the overall benchmark.

Annex 2.95. Similarly, the Inequality-adjusted Human Development Index (IHDI) adjusts the Human Development Index (HDI) for inequality in the distribution of each dimension across the population. The IHDI accounts for inequalities in HDI dimensions by “discounting” each dimension’s average value according to its level of inequality. The IHDI equals the HDI when there is no inequality across people but falls further below the HDI as inequality rises. In this sense, the IHDI is the actual level of human development (taking into account inequality), while the HDI can be viewed as an index of the “potential” human development that could be achieved if there was no inequality. The “loss” in potential human development due to inequality is the difference between the HDI and the IHDI and is expressed as a percentage.

Annex 2.96. The PEFA framework itself points to another option. While the core framework provides data in a form that allows comparison over time (and between countries, but with a number of cautions on the comparability of the data given the methodology), the data in itself are just indicators. Key to understanding the indicators is the standardized assessment report, which systematically describes the public financial management systems, and provides the reasons for scoring.

Annex 2.97. This approach is taken a step further in the Children’s Budget exercises, where a standard budget analysis methodology is applied to the budget data collected, but the benchmarking is done through the analysis and interpretation of the resulting data in a report (see for example African Child Policy Forum 2011).

Annex 2.98. The OECD Child Wellbeing indicators are presented through a sophisticated dashboard which allows users to, among other things, organise data on scatter plot diagram, by selecting indicators for the x and y axis to show the relationship between indicators across the different countries. In our case, a similar tool could be used to plot interpretative information (such as GDP/capita; HDI) against CP expenditure, enabling an appreciation of overall correlation and where specific countries fall in relation to this.
Disaggregation potential

Annex 2.99. The potential to disaggregate the components of an indicator to provide finer grain information within a country – such as regionally, by gender, household income, or by specific beneficiary groups – would be particularly important in instances where the indicator is being used for in-country monitoring and diagnostics. As such, disaggregation capacity was central to the design of the EAPRO child protection governance indicator (which can be disaggregated by governance domain, or individual component indicator) and the SPI (which can be broken down by depth of coverage (the average size of benefits received by actual beneficiaries), breadth of coverage (the proportion of intended beneficiaries who actually receive benefits), category of SP or component programme, poor and non-poor beneficiaries, and gender). Disaggregation by age cohort – as has been the practice in some country child budget analyses – can be useful for distinguishing between children, adolescents and youth (UNICEF, forthcoming).

Annex 2.100. The logic behind this is evident; in using an indicator to diagnose weaknesses in a country’s policies and processes and identify potential reforms, the more detail provided by the indicator the more specific the resulting response can be. That is not to say that disaggregated benchmarks can prescribe detailed reform measures or offer a recipe for change, rather they should be treated as an entry point to the reform process by pointing to the principal gaps, weaknesses and distortions (UNICEF 2012). There is an additional trade-off to consider in that the more an indicator is to be disaggregated (and therefore the more informative it is), the greater the burden of data requirement (addressed below).

Annex 2.101. Related to this is the recognition that a summary aggregate score (as is found under the Human Development Index (HDI), the original SPI, the Child and Youth Well-being Index (CWI), the Worldwide Governance Indicators (WGI), and the OBI) is only necessary if a core feature of the index is cross-country comparison and ranking. If the focus is instead on in-country monitoring and diagnostics, aggregation to this level can and probably should be avoided, as it requires a weighting to be applied to composite indicators which will always leave itself open to the criticism of being arbitrary. In most of the weighted indicators reviewed, a simple average was employed (e.g. HDI, CWI and OBI) because it was considered to the most transparent, but not without criticism. The HDI’s equal weighing of health, education and income has been deemed “over-simplistic” (Klugman et al, 2013). In the case of the OBI, equal weighting has resulted in implicit bias by failing to offset the disproportionate influence of the budget documents which are scrutinised in more questions in the survey (Seifert et al, 2013). Under the WGI, component indicator are weighted according to their relative precision (or variance of the error term), but this too has been criticised for giving to much weight to data sources that make correlated errors (Arndt and Oman, 2006; Knack, 2006). In short, pursuing a framework without a summary indicator avoids problematic weighting and can be considered unnecessary for the purpose of the child protection financial benchmark anyway if it is aimed more at in-country than cross-country comparison and analysis. Indeed the movement away from an arbitrary weighting system, was a key adjustment in the revision of the SPI, for exactly such purposes (ADB, 2013). The OECD Children Wellbeing indicators, which are a series of 21 policy-focused measures of child well-being in six areas, deliberately avoid presenting an over-arching index due to the limitations in the coverage of available data and little theory to guide which aggregation method to use. Given this, it was considered that creating an over-
arching index would “distract the focus towards discussion of the aggregation method, and away from more important practical issues of improving child well-being” (OECD, 2009b).

**Flexibility for and coverage of different country contexts**

Annex 2.102. For the benchmark to be meaningfully applicable to the array of high, medium and low income countries, including emergencies settings and fragile countries as well as more stable ones, it needs to be sufficiently flexible.

Annex 2.103. The “core-plus” model adopted in the UNICEF EAPRO regional core indicators for child protection governance may provide a useful model. It proposes a core set of indicators with the expectation that these will be expanded at the country-level in response to data as they become available (UNICEF, 2012). For a child protection indicator the ‘core’ could be used across all countries, but with additional indicators or interpretive devices depending on the country context. The core can be used for cross-country analysis across country groups, if necessary, whilst the broader set can be used for in-country monitoring and diagnostics or comparison within a group of countries. One of the design questions if such a solution is used, would be the criteria against which countries are grouped (such as per capita income level; governance capacity; child protection orientation; and/or emergency status).

Annex 2.104. A disadvantage of the build-as-you-go approach at the country-level is that it complicates longitudinal analysis, as the introduction of new data sources or component indicators undermines the robustness of assessment of progress over time. This limitation is acknowledged in UNICEF EAPRO child protection governance framework (UNICEF, 2012). The WGI attempts to address a similar problem by retroactively adjusting historical data whenever changes to the set of sources are introduced (Kaufmann et al., 2007), however this is not always possible depending on data availability.

**Scope/practicality considerations**

Annex 2.105. There is an inevitable trade-off between the scope of a benchmark and the acquisition effort required. Whilst potentially being more informative, broad scope has implications for feasibility, affordability, and sustainability of needed analytical effort. Both the UNICEF EAPRO child protection governance indication and the SPI require significant input (particularly when first introduced in a country), if not for the technical rigour then for the sheer volume data to be collected, which is often not available in one central location and requires a lot of follow-up and cross-checking for accuracy. Sustaining Government ownership of such a process may prove challenging, but is essential, particularly in such cases where data are not available publically. This echoes with a UNICEF methodological guidance note on child budgeting, currently being drafted, which identifies 12 initiatives around the world to carry out annual child-focused budget analyses. Of these, only 7 were found to be still active, with an average lifespan of 9 years. Limited Government involvement (which only occurred in 2 of cases), is identified as a constraint to longevity, and is an explanatory factor for the minimal impact the recurring analysis have had on policy and budget frameworks (UNICEF, forthcoming).

Annex 2.106. There are potential lessons to be drawn from those benchmarks which have achieved substantial and sustained coverage, such as the HDI, which covered 187 countries in 2013 (UNDP, 2013), and the PEFA, which by 2010 had assessed 90% of low income, 75%
of middle income and 8% of high income countries (Lawson and Fölscher, 2011). For the HDI, this is due to this simplicity of the index’s characterization of development (as an average of achievements in health, education and income), which conveys the message that development is about much more than growth (Klugman et al, 2013). The PEFA similarly has succeeded in providing a framework for the assessment of PFM functionality which manages to be comprehensive in its coverage, but also sufficiently simple for non-technical users to understand (Lawson and Fölscher, 2011). It’s centrality to dialogue between donors and Government on the issue of budget support has further enhanced its relevance. An OECD paper on best practice in benchmarking notes that an additional benefit of engaging the central Finance Ministry in particular when imposing a benchmark, in that it creates potential for links to the budget process and other relevant decision making. Connecting the benchmark to resource allocation decisions in this manner can provide an incentive to improve performance (Helgasen, 1997). This is particularly important if benchmarking is to be used in relation to internal competition in the public sector.

Annex 2.107. In contrast, the OMBI has not achieved significant coverage, with only 16 municipalities participating (out of 444 in Ontario alone) (OMBI, 2012a). Reasons for this have not been documented, but may be associated with a hesitance on the part of municipalities to share performance data for which they could then be held accountable, or due to the associated additional reporting requirements.

Annex 2.108. One method of achieving scope whilst ensuring data collection is feasible is to rely on existing data sources where they exist. Neither the WGI nor the HDI undertake any primary data collection. To ensure the quality of these existing data sources, the HDI relies solely on well-established international institutions such as UNESCO Institute for Statistics, UN Department of Economic and Social Affairs, the World Bank, and IMF (UNDP, 2013). The WGI relies on several hundred variables obtained from 31 different data sources that capturing governance perceptions as reported by survey respondents, nongovernmental organizations, commercial business information providers, and public sector institutions. It explicitly records margins of error in each of its measures, enabling the identification of statically significant changes over time or comparisons between countries (Kauffmann et al, 2013). UNICEF child protection practitioners, as the intended user and constructor of the child protection financial benchmark at country-level, may not have the required statistical capacity for such an approach, and at the same time there is no institution collecting international data on child protection expenditures. Therefore primary data collection in country will be inevitable.

**Quality Assurance:**

Annex 2.109. Quality Assurance (QA) is critical for safeguarding the credibility of a benchmark. This point is emphasised in the 2011 PEFA evaluation which found that quality of assessments varies because the secretariat is not the “owner” of the assessments and as such, not all pass through its designated QA system (Lawson and Fölscher, 2011). Due to the reputational risks generated by (albeit isolated) cases of poor quality PEFA assessments, this is a very serious observation. The OBI also highlights the importance of a robust QA system; upon completion of the survey by local researchers, IBP staff members analyse and discuss each questionnaire with the individual researchers to ensure that all questions are answered in a manner that was internally consistent for each country, as well as consistent across countries. The answers are also cross-checked against published budget documents and reports on fiscal transparency issued by international institutions, such as the IMF, World
Bank, and the OECD. The questionnaires are then reviewed by two anonymous peer reviewers who have substantial working knowledge of the budget systems in the relevant country, and respective Governments are also invited to comment on the draft Survey results (IBP, 2012). The Ontario Municipality Benchmarking Initiative has an in-built review system wherein an overarching Financial Advisory Panel, comprised of representatives from each municipality, ensures cost data is collected in a consistent manner (OMBI, 2011). UNICEF HQ would be an obvious home for a QA system of the child protection financial benchmark, perhaps engaging Regional Bureaus to do a first phase of review and regional standardisation.

Annex 2.110. In addition to formal QA systems, the use of detailed guidelines or a handbook in addition to preconfigured spreadsheets for data entry, such as is used in the revised SPI, is reported to aide standardised collection and calculation (ADB, 2011). OMBI has developed a catalogue of tools and methodologies to facilitate standardised data collection across municipalities, including data dictionaries, a data warehouse, data collection protocols, an indirect costing methodology, data sharing and public reporting protocol, and a common measurements tool, consisting of a question bank and a user manual providing an easy-to-use method for collecting feedback from customers (OMBI, 2012b).

Annex 2 Section 3: A framework of approach for designing a Child Protection Financial Benchmark

Annex 2.111. The next step for this assignment is to design options for a benchmark. This final section of this report therefore proposes a framework to approach the design, criteria to assess different options, and next steps for the assignment.

Approach to designing a benchmark

92. The design of a UNICEF child protection financial benchmark would need to address three core questions:

- What child protection expenditure data will be included / excluded?
- What interpretive perspectives are needed as part of, or to frame, the expenditure measure(s) used in the benchmark?
- What benchmarking methodology will be used, where methodology refers to the architecture of the benchmark and its application?

Annex 2.112. These three main questions unpack into a framework of inter-related and inter-dependent central and sub-questions to address in the benchmark design. This framework is set out in the table below, identifying the criteria that will be relevant to each of the central questions.

<table>
<thead>
<tr>
<th>Annex 2 Table 5</th>
<th>Framework for designing a financial benchmark for child protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Question</td>
<td>Sub-questions</td>
</tr>
<tr>
<td>What child protection expenditure data will be included / excluded?</td>
<td>What child protection expenditures will be included?</td>
</tr>
<tr>
<td></td>
<td>What is the definition of child protection to be used?</td>
</tr>
</tbody>
</table>

Annex p36
<table>
<thead>
<tr>
<th>Question</th>
<th>In-country monitoring and diagnostics</th>
<th>Ownership and impact</th>
<th>Simplicity of approach</th>
<th>Practicality of data collection and cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Will the demarcation be defined exogenously or endogenously (see paragraph 1 Section 1.19 below for discussion)?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>What demarcation rules must be applied to identify actions and services (see paragraph Annex 2.116 below for discussion)?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Will the same rules be used across all countries; what are options for a core ‘compulsory’ framework with optional additional components? What are options for different frameworks in different contexts?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Expenditure associated with which funders will be included?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Are domestically financed / internationally financed expenditures included or both?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Of domestically financed expenditure, is only government financed expenditure included, or expenditure financed by other domestic actors (private sector, private individuals, NGOs, families)?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Will the same rules be used across all countries; what are options for a core ‘compulsory’ framework with optional additional components?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Expenditure expended by which service providers will be included?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Is expenditure by both international and domestic service providers included?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Of domestic service providers, is expenditure by government and other domestic actors (private sector, NGOs, communities) included, or only by government?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Will the same rules be used across all countries; what are options for a core ‘compulsory’ framework with optional additional components?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Expenditure financed and expended at which level will be included?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Is all expenditure included (national and sub-national)?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Will the same rules be used across all countries; what are options for a core ‘compulsory’ framework with optional additional components?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>What expenditure data will be used?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Will budget, disbursement, outturn or audited actual expenditure data be used?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>Will only aggregate data be used, or will it be disaggregated by economic classification?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
<tr>
<td>What proxy rules can be designed to assist in isolating child protection expenditure when the budget classification system does not provide enough detail to do so?</td>
<td>In-country monitoring and diagnostics</td>
<td>Ownership and impact</td>
<td>Simplicity of approach</td>
<td>Practicality of data collection and cost</td>
</tr>
</tbody>
</table>

Annex p37
<table>
<thead>
<tr>
<th>What interpretive perspectives are needed as part of or to frame the expenditure measure(s) used in the benchmark?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Will the same rules be used across all countries; what are options for a core ‘compulsory’ framework with optional additional components?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What non-financial child protection information/data will be used?</th>
</tr>
</thead>
<tbody>
<tr>
<td>What information on non-financial inputs will be used – if any – to frame an understanding of child protection expenditure?</td>
</tr>
<tr>
<td>What information on outputs, beneficiaries and outcomes – if any – will be used to frame an understanding of child protection expenditure?</td>
</tr>
<tr>
<td>How will information be used?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What other socio-economic information will be used (such as GDP, HDI, population below the poverty line, no of children with disabilities/orphans and vulnerable children etc)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>What other socio-economic information need to be provided in a benchmark – if any – and how will it be used?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What information on the budget and public financial management system will be used?</th>
</tr>
</thead>
<tbody>
<tr>
<td>What aspects of the budget / public financial management system will be assessed for a benchmark – if any – and how will these aspects be assessed?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What benchmarking methodology will be used?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Will the benchmark be applied in country-driven processes and with a country-determined timeline? Or will the benchmark we applied across countries on a benchmark-determined timeline?</td>
</tr>
<tr>
<td>Will the benchmark attempt to rank countries / localities in a country?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What aspects of the budget / public financial management system will be assessed for a benchmark – if any – and how will these aspects be assessed?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interpretability</td>
</tr>
<tr>
<td>In-country monitoring and diagnostics</td>
</tr>
<tr>
<td>Simplicity of approach</td>
</tr>
<tr>
<td>Practicality of data collection and cost</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Will the benchmark be constructed?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Will the benchmark by a unitary measure (single or composite), or a framework of measures?</td>
</tr>
<tr>
<td>Will the benchmark measure change over time, or provide a snapshot at different points?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interpretability</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-country monitoring and diagnostics</td>
</tr>
<tr>
<td>Simplicity of approach</td>
</tr>
<tr>
<td>Practicality of data collection and cost</td>
</tr>
</tbody>
</table>

Annex p38
### Explanatory remarks

**Annex 2.113. Use of the framework**: The questions posed in the framework are interrelated. For example, if the decision is to use an endogenously determined benchmark that is fully sensitive to country context, a framework of indicators is likely to be more the useful construction choice. The point of departure would be narrowly framing the objective of the indicator (for example: “assessing expenditure by the state -- domestically and externally financed -- on an exogenously determined list of key issues by country”), and how it would be applied by UNICEF (for example: country-by-country on a country-determined timetable). The team will then work on the options across the central questions posed, which are enabled by this choice, and assess the options against the agreed criteria.

**Annex 2.114. Exogenous and endogenous demarcation of child protection**: A key choice in the design of the methodology would be whether child protection would be defined or demarcated exogenously or endogenously. It is clear from the literature that there is an increasing awareness of the variation between countries -- driven by values, history and current circumstances -- in child protection systems. At the same time, the Convention on the Rights of the Child Article 19 provides an internationally agreed definitional framework for child protection, which has been interpreted further within UNICEF’s strategy instruments, orienting the focus to state parties (the key duty bearers in accordance with the CRC) and around specific risks. The choice therefore would be whether a benchmark would attempt to compare expenditure on child protection systems as they manifest within countries, given the country’s culture, history, current circumstance and the predominant risks related to violence, abuse, exploitation and neglect of children; or whether the benchmark would compare expenditure by states on specific issues, risks and services. If the latter the benchmark would define the list and require assessing expenditure against the list, or it could provide the list as a menu, with the most prominent risks being selected for costing or analysis at the country level.

**Annex 2.115. The benchmark could also strike a balance between exogenous determination of the child protection system, and endogenous determination, using a core list of issues, risks**
or services that are applied across all countries, with additional components that are country defined or centrally defined but up to the country whether to include or not.

Annex 2.116. **Demarcation rules**: Issues around demarcating child protection specific (or direct) expenditures from related (or indirect) expenditure as highlighted in the TORs remain, even if the demarcation of the child protection system is driven by a benchmark-specific list or menu of issues / services. Our discussion up front suggested that this demarcation is particularly difficult with regards to the inclusion of ‘prevention’ and ‘neglect’ in the definition of child protection. The reviewed literature suggested other options, including

- Exclude activities that are not easily quantifiable
- Include only some social transfers, but exclude others because they do not target children / are unlikely to contribute to child welfare / does not constitute a benefit flow.
- Including expenditure aimed at fulfilling the sub-set of the rights of the child under CRC article 19, as opposed to rights under other articles;
- Including all expenditure which will increase children’s safety even if this is not the explicit purpose of the expenditure;
- Including expenditure that is aimed at explicitly protection children from harm, but excluding expenditure which is aimed at their well-being even if it has the effect of protecting them from harm.
- Including expenditure on well-being/fulfilling other sub-sets of rights only where these expenditures relate to groups of children at high risk of harm.
- Exclude expenditure in sectors that if removed, would imply a redefinition of the sector.

**Criteria for assessing options**
Annex 2.117. Annex 2 Table 5 above provides the relevant criteria against which different design options in each central question dimension will be assessed:

- **Objective of the benchmark**: Will this option support the objective of the benchmark? The objective of the benchmark will frame more precisely what UNICEF wants to assess using the benchmark. Key dimensions for a point of departure are:
  - Exogenous or endogenous demarcation of the child protection services to be assessed?
  - Only government service provision or all of society?
- **In-country monitoring and diagnostics**: Will this option provide sufficient functionality for in-country monitoring and diagnostics by UNICEF Country Offices.
- **Multi-context relevance**: Can this option be applied with equal ease / is it relevant across the different contexts within which UNICEF Country Offices work?
- **Multi-context comparability**: Can this option be applied across different countries in a way that will make the results sufficiently comparable to allow UNICEF Country Offices to use the benchmark for evidence-based advocacy?
- **Interpretability**: Will this option deliver a benchmark result that will be understood easily by UNICEF’s key partners, audiences and stakeholders? Will this option contribute to the interpretation of the benchmark against country context?
- **Simplicity of approach**: Will this option deliver a benchmark that can be applied and used by UNICEF Country Offices in different contexts? Is it sufficiently simple / complex as a starting point for using financial benchmarking in the child protection sector?
• **Methodological acceptability/rigour:** Will this option deliver a benchmark that is defensible from a methodological perspective, facilitating its use for analysis, diagnosis and advocacy?

• **Ownership and impact:** Will this option facilitate ownership by UNICEF’s government partners (and donor partners in emergency contexts), ensuring that it will be used to improve child protection expenditure / outcomes in-country.

• **Practicality of data collection and cost:** Are data available for this option? What will be the likely difficulties of data collection? Will data-collection be affordable?

**Next steps**

Annex 2.118. The team will continue its data collection and review, drawing on on-going interviews with key respondents and following on pointers from the documents already reviewed. A key area of further work would be to get a better understanding of how the child protection sector works in a selection of countries, so as to be able to better judge different options against the multi-context relevance and comparability criterion. A second key area is to generate a better understanding of the links between child protection services and other socio-economic outcomes, in order to do further work on the additional interpretative information that may be needed.

Annex 2.119. Before moving to the development of options and their assessment, the team would need to engage with the UNICEF team to:

• Review the issues highlighted in this report,
• Review the proposed and agree on an indicator design framework;
• Review the proposed assessment criteria, and decide on a set of criteria, as well as the weight of different criteria;
• Generate a more precise objective of the benchmark, so as to frame the development of options and their assessment.

Annex 2.120. As soon as there is agreement on these points, and options have been developed and a preferred design identified, the benchmark will need to be piloted to assess the feasibility of data collection.
Annex 2.1: Terms of Reference

Terms of Reference
Financial Benchmark for Child Protection

Introduction
The purpose of these terms of reference is to develop the framework for a financial benchmark for child protection (CP). The purpose of financial benchmarking is to obtain: (i) a better and comparable measurement of actual expenditure on CP systems across countries; and (ii) an indicator for adequacy of public spending. While benchmarking is not limited to financial data, and a financial benchmark on its own is insufficient to assess the whole CP system, comprehending the financial resources available is critical to CP reform, and understanding the financial implications of CP reform is necessary to make a sustainable difference in development goals.

There has been a shift in the approach to CP from programming with a narrow issues and target groups lens to placing these programmes within a much broader systems strengthening approach that includes enhancing positive social change. This shift allows for capitalizing on all resources available to prevent violence against children, mobilizing all actors and components of the system to protect children. In doing so, systems thinking enables governments to more efficiently use budget resources by identifying synergies across programmes and budget lines. With ODA decreasing, and it generally being a smaller part of overall expenditure on CP, knowing the possible extent of national spending on CP is one critical component for successful reform of the CP system.

Tracking and highlighting outcomes for children as well as the expenditure used for these outcomes can be useful in creating visibility and demand for more effective action. Cross country comparisons can also be useful, by identifying ‘high’ and ‘low’ performers, in influencing national policy and highlighting good practice in the use of resources. This can be seen in indices such as the Human Development Index or in the analysis of public expenditure in specific sectors such as health or education.

A conclusion of the recent conference “A Better Way to Protect ALL Children: The theory and practice of child protection systems”\(^{17}\) highlighted the requirement for the development of financial benchmarking for CP. Developing such benchmarks through cross-country data analysis would stimulate improvement in the data on investments in CP while also better informing policy decisions. While there has been some work reviewing financial inputs to CP, to date this has been limited, and no standard methodology has been developed or applied. Collecting financial data on CP is inherently fraught with questions, particularly as public sector resources are often spread across ministries, and many protection resources are also found in the non-public sector.

There are a number of existing tools that can be drawn on to begin the benchmarking discussion. The UNICEF East Asia and the Pacific region’s tool “Measuring and Monitoring Child Protection Systems”\(^{18}\) includes indicators for public finance management (PFM) and is currently being tested in

---

\(^{17}\) The documents from this conference can be found at [http://knowledge-gateway.org/sharekluo5gnjrn31p71ra1zp7b2hnl4gj5vz27/childprotection/cpsystems/cpsconference/library](http://knowledge-gateway.org/sharekluo5gnjrn31p71ra1zp7b2hnl4gj5vz27/childprotection/cpsystems/cpsconference/library)


Annex p42
the region. The Asian Development Bank’s “The Revised Social Protection Index”\textsuperscript{19} was revised to specifically look at social assistance and social services for CP within the remit of social protection. Other indicators have been developed, such as the World Bank World Governance Indicators and the Public Expenditure and Financial Accountability PFM Performance Measurement Framework, which provides insights into overall governance and PFM, although not specific to CP. The budget classification system used in country may support this effort and there are important efforts to make better compilation and analysis of public expenditure data at national and subnational government levels. A key initiative supported by the World Bank is BOOST, which has the potential of facilitating systematic measurement, tracking, and benchmarking of the allocation and use of public financial resources in key social sector and systems including CP.

Financial Benchmarks for Child Protection

Focusing specifically on financial benchmarking and CP, there are a number of areas that need to be examined in order to define the way forward. Specifically:

- **Inclusions/Exclusions within the scope of Child Protection Expenditure**

As child protection is a sector and intersectoral, the boundaries for the benchmarking need to be clear. The general definition of CP is that it prevents and responds to the violence, abuse and exploitation of children. While the focus is on services, income transfers are at least in part also a mechanism to protect children. There is also recognition that some of the services and cash target the overall family (e.g. family counseling) or other specific family members (e.g. disability assistance), and not just the child, and yet impact on the child. Further CP is an area that is often not contained only within a particular government ministry or civil society programme, but rather crosses a range of ministries and programmes.

The breadth of child protection encompasses many child welfare services, including some in the education and health sectors. For example, teachers who are on the look-out to identify abuse or medical workers who identify violence and respond both with physical and psychological care. Taking such a wide-ranging definition of the relevant services and programmes could result in obscuring the cost of essential services to prevent and respond to violence. On the other hand, excluding such inter-sectoral work would miss many of the important responses that effectively protect children.

In general activities that should not be included are:\textsuperscript{20}

- Activities which are not amenable to quantification, such as legislative and policy reform, child rights, empowerment and consciousness raising activities.
- Public and private contributory insurance schemes such old age pensions, and (un)employment benefits.
- Social assistance and state pensions targeted to ‘priority citizens’ such as military veterans, civilian victims of war, politicians and civil servants. Also excluded would be


\textsuperscript{20} Some of the following points are adapted from the Asian Development Bank, *The Revised Social Protection Index: Methodology and Handbook*, 2011.
micro credit and microfinance as these incur loans rather than simple transfers, and cash/food for work due to the offset required in the calculation by the work.

✔ Programmes that are seen as standard programmes within other sectors, and if removed constitute a new understanding of that other sector (e.g. immunization and regular check-ups are a standard aspect of health care; delivery of schooling, preschool education, teacher training are a standard aspect of education; prevention of mother to child transmission are integral to HIV/AIDS programmes). That said, identification and treatment of abuse in health care and education facilities, support to teachers to implement positive discipline, birth registration that may occur at the same time and location as immunization, social care of children affected by HIV, and so on would be included.

✔ Programmes that fall under the general category of rural or community development along with those that concentrate on the construction of physical assets (e.g. roads) would be excluded. Those assets of social infrastructure would generally be excluded, except for those specific to CP (e.g. small group homes, offices housing social workers, day centres, etc.).

✔ Excluding tax credits, unless a specific replacement of a child assistance transfer.

What should be included? A list of child protection services was drafted in 2008\textsuperscript{21} to reflect those programmes that are critical to CP and this may be of assistance in responding to this challenge. These may provide a useful introduction to a financial benchmarking exercise in CP which may begin by identifying the services available. It should include administrative, labor and operating costs as key components of child protection. Nevertheless, there will remain a number of other areas where a discussion is specifically required to consider whether or not to include these in the calculation and benchmarking exercise.

The implications of including, or not, different areas for benchmarking is critical to understanding the meaning and significance of the end calculation. Also to be considered is what it is generally feasible to collect; while ideally the benchmark that is designed could be used for all countries, it is imperative that it can be applied in high, middle and low income countries and this will place certain restrictions on the scope of data collection.

- **Data Acquisition**

Data should be collected in agreement with the national government, and focus on obtaining information for the most significant CP programmes in terms of expenditure and beneficiaries. These ‘programmes’ may need to be ‘grouped’ or aggregated such as in cases where a general service such as ‘group homes’ is decentralized, outsourced, or is part of a service from private organisations. In countries where CP work is decentralized, there may also be a need to collect data from the relevant lower government levels.

Data sources would need to consider both ‘top-down’ and ‘bottom-up’ approaches. In the first case there may be a ministry of children’s affairs or equivalent that effectively delineates or coordinates CP actions, or an NGO specifically responsible for child protection. Within the government budget, the relevant line items available may be sufficiently disaggregated or the methodology used for budget classification may be easily interpreted to identify CP expenditures. Household surveys may also

\textsuperscript{21}See Annex with list of services defined at the UNICEF on Child Protection Systems Mapping and Capacity Assessment Consultation, 11-13 June 2008, Bucharest, Romania.
provide some information on the extent of the potential beneficiaries of the programmes, as well as those who may receive it.

A ‘bottom-up’ approach would require looking at administrative data for programmes. It would involve understanding the content of a particular programme to make an informed decision on the aspects to be considered in the financial benchmarking. This could be quite complex, particularly given the cross-cutting characteristic of CP, and require review of programme regulation documents and worker practices. There may need to be a judgment call on the extent of coverage and cost of a programme (NGO or public, provincial or national), and whether it is large enough to affect the summary results (e.g. the cost of a project may or may not have high enough expenditure to impact on the national expenditure level for child protection). Sometimes extrapolation may be required, such as assessing costs in sample provinces to extrapolate to others with a similar economy and programme size.

The focus is on the actual expenditure and delivery of CP as distinct from a focus on what is planned. Therefore it is expected that the result would be for the most recent year that expenditure date is available, acknowledging the time required to close and publish fiscal year accounts. “It should be emphasized that while much of the data work involved is not technically complex, it is rendered difficult by the multiplicity of organisations involved, the need for repeated visits to these organisations and the careful cross-checking of the data provided”

- **Numerator and Denominator**
  The numerator for the benchmark is possibly the most complex aspect. While focusing on public expenditure may be appropriate for this exercise, it is important to consider whether it is possible to include the spending of the full sector (public and private), or whether, and how, the extent of non-public sources (such as through a mapping of use of non-public service providers) could be assessed to inform the end result. Another question would be on whether to focus on only national or also sub-national expenditure, noting that sometimes CP services are decentralized and tailored to local requirements.

  With regards to the denominator for the benchmark (GNI, GDP, government budget), the principal limitation is that this does not indicate the impact on the child and its family. While some of this information may come from other indicators, such as those found in UNICEF’s Strategic Plan (2014-2017), and may evolve as further performance benchmarks are identified, options that can indicate the extent of funding provided to children and their care givers for their protection may also be explored. One option may be to look at the funding provided by child population and translating the figures into purchase power parity to allow for further cross country comparisons.

- **Interpretation**
  Considering only finances used for CP does not allow for a full understanding of the impact on child protection. Fiscal resources may be used in ways that are deemed more negative than positive in terms of the care and protection of children, for example for detention rather than diversion in the case of children in conflict with the law. It may however provide a basis for international comparison. The value may also reflect the country’s priorities, its wealth and the demand for CP services. In itself though it is important to clearly understand what it says, and potentially understand the extent of


Annex p45
cross-country comparison that it may provide, e.g. similar economic profiles, population compositions and political contexts (including conflicts and emergencies) etc.

- **Disaggregation**
  The level of disaggregation desired or possible is dependent on the numerator. If the number of beneficiaries is also included, it may be possible to assess the distribution of expenditure by sex, age, location, level of poverty, category, etc., although this may be more limited if focusing solely on finances. It would also be dependent on the possibility of tracking expenditures, for example the extent by which national expenditure is spent at the local level. That said, in countries with decentralized expenditure systems, it may be possible to look at distinctions between rural and urban, public and private, administration and service delivery. Currently the focus is on the overall national spending, considering that many CP programmes are a relatively small proportion of the overall government expenditure.

**Objective of the Assignment**

The purpose of this assignment is to lay the groundwork for the development of financial benchmarks for child protection. The aim is to review existing materials and develop a proposal for a financial benchmark for child protection. Working with a reference group it would lay out a potential roadmap for next steps such as the development of a manual, and testing in countries.

**Duties and Responsibilities**

- To conduct a literature review of existing financial benchmarking methodologies and uses, focusing both in the private and public sector. While the social sector is the main area of interest, other examples may also be drawn upon. The output should provide an indication of good practice that can inform the options for financial benchmarking for child protection.

- Discuss potential options for CP financial benchmarking. This would include a discussion on: (i) the scope of CP expenditure to be included; (ii) data acquisition; (iii) numerator and denominator; (iv) interpretation; and (v) disaggregation. Within each, the paper should discuss the risks and benefits, including regarding the potential of each option for allowing identification of synergies across budget lines as well as solutions to existing inefficiencies of a CP system

- Propose three options for financial benchmarking of child protection that can be applied in all contexts. This would include the scope and the components to be included. It would also articulate briefly the advantages and obstacles for implementing the methodology that would be required.

- After reviewing the options, develop a roadmap of next steps that would result in a tested manual for determining the financial expenditure on CP in a country and further financial benchmarking for child protection.

**Deliverables**

- Literature review
- Characteristics of CP financial benchmarking
- Three options for financial benchmarking
- Roadmap of next steps

Annex p46
CONTRACT ARRANGEMENTS

A desk assignment, it is expected to last approximately 30 days over 60 days, with the end date being before 30 September 2014.

Payment schedule will be as follows:

<table>
<thead>
<tr>
<th>Task</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature review and characteristics of CP financial benchmarking</td>
<td>25%</td>
</tr>
<tr>
<td>3 options for CP financial benchmarking</td>
<td>25%</td>
</tr>
<tr>
<td>Roadmap of next steps</td>
<td>25%</td>
</tr>
<tr>
<td>Final report</td>
<td>25%</td>
</tr>
</tbody>
</table>

LIAISON AND REPORTING

The selected candidate will work under the supervision of the Senior Advisor, Social Welfare and Justice Systems, Programming Division, and in Consultation with the Chief of Public Finance and Local Governance, Division of Policy and Strategy. Outputs will be submitted in English.

A reference group will be established by UNICEF to provide feedback on each draft output such that any adjustments can be made for the final product. It is also expected that the reference group will provide clear directions on the suggested options so that the roadmap is tailored to a selected CP financial benchmarking option.

- Approval of the final product shall be provided by UNICEF.
- The contract will be terminated with immediate effect if the performance of the consultant is not satisfactory.
- The consultancy does not guarantee any future work with UNICEF.

CONSULTANT QUALIFICATIONS

- Graduate degree in economics, public finance accountancy or audit, public policy or equivalent
- A minimum of 10 years professional experience working in public finance management, particularly relating to child or social protection
- Proven research skills
- Excellent English communication skills, fluency in another UN language an asset
- Fluency in written and spoken English, and excellent report writing skills
- IT skills (ie Word, Excel or equivalent)
- Experience working in international environments

HOW TO APPLY

Qualified candidates are requested to submit a cover letter, outline methodology for the assignment, and an example of their published work. In addition provide your CV and P11 (the
P11 template can be downloaded from [http://www.unicef.org/about/employ/index_53129.html](http://www.unicef.org/about/employ/index_53129.html) to pdconsultants@unicef.org with subject line “Child Protection Benchmarking Consultant” by **26 March 2014**. Please indicate your availability, to undertake the terms of reference above. Applications without a daily fee rate or methodology will not be considered.

---

**Annex A**

Summary list of services and sectors it may connect with, adapted from the Bucharest workshop

<table>
<thead>
<tr>
<th>Prevention</th>
<th>Social welfare Sector</th>
<th>Education Sector</th>
<th>Health Sector</th>
<th>Justice &amp; Security Sectors</th>
<th>Community, NGOs, Private Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitive health, police, judicial, social work case management, shelter services for child victims of violence or exploitation</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Verification &amp; investigation of cases</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Psycho-social support/mental health services</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Identification, reporting, verification of suspected abuse/exploit</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Tracing, reunification, reintegration</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diversion &amp; alternatives to custody</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alternative care (foster, residential, emergency shelter, adoption)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best interests determination &amp; gate-keeping procedures</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complaints mechanisms—care, detention, schools, health facilities</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Individual family support—mediation, assist with entitlements, service access, respite entitlement, legal aid</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Identification of at-risk children &amp; families</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Public education &amp; community mobilization for social change</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Life skills, youth civic engagement</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Background checks &amp; codes of conduct for those working with children</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Daycare</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth registration</td>
<td>X or civil affairs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income supplements</td>
<td>X and others</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Annex 2.2: UNICEF CP Financial Benchmarking -- Review of Comparative Benchmarks

REGIONAL CORE INDICATORS FOR MEASURING AND MONITORING GOVERNANCE OF CHILD PROTECTION SYSTEMS – UNICEF East Asia and Pacific

Summary

Presents a framework of core indicators to support monitoring and assessment of the enabling environment or ‘governance’ of national child protection systems.

It is grounded in the recognition that the translation of inputs (material and financial) into CP outcomes, depends on the CP system’s governance: that is, the network of institutional arrangements (including government) and processes that determine decision making and the process by which those decisions are/are not implemented. Governance with respect to CP includes critical processes related to the management of financial, human and information resources, legislative and policy frameworks, as well as a dynamic cultural context which shapes social expectations with regard to protection of children.

Primary Purpose

The primary purpose of this indicator framework is to enable country stakeholders, led by Government, to assess the impact of policies and programmes on child protection systems over time, thereby developing an evidence base for the strengthening of national child protection systems.

Being standardised across the region, the indicators can also be used at the regional level for comparing progress between countries, analysing comparative strengths and weaknesses of existing systems, and jointly using emerging information to improve child protection systems. However ranking is deemed not possible (see below).

The PFM indicators are envisaged to also:

- Contribute to wider reforms in PFM, and thus help child protection actors engage in wider partnerships within the government and with other development agencies such as the World Bank and the IMF, potentially strengthening leverages for child protection reforms;
- Facilitate the review of public financing in child protection with the view of eventual development of sector-wide costing of child protection programmes;
- Contribute to reforms related to the division of responsibilities for child protection across levels of government by highlighting financial aspects of related arrangements.

Construction

The framework contains a menu of 37 “core Indicators” of Child Protection System Governance, which are spread across seven major governance domains. Each indicator has a related benchmark and series of Standard (ranked A-D).

Domains:

1. Child protection policy framework
2. PFM
3. Human resources child protection
4. Statistics and information
5. CP systems coverage
6. Surveillance, gate keeping, referral & quality assurance
7. Public attitudes and values

Seven indicators under domain 2 (PFM) are:
2.1 **Credibility of budget allocations across Child protection-related sectors**: Benchmark: Composition of expenditure out-turns compared to original approved budgets on Child protection-related programmes (administrative and functional) over the last three years. Standards: A. Variance in expenditure composition less than 5 per cent in any of the last three years; B. Variance in expenditure composition by 5 per cent in one year over three years; C. Variance in expenditure composition by 10 per cent in no more than one out of the last three years; D. Variance in expenditure composition by 10 per cent in two out of the last three years. Measurement against this indicator requires an empirical assessment of expenditure out-turns against the original budget at a sub-aggregate level for related sectors. Variance is calculated as the weighted average deviation between actual and originally budgeted expenditure calculated as a percentage of budgeted expenditure on the basis of administrative or functional classification, using the absolute value of deviation.

2.2 **Robustness of budget expenditure classification** Benchmark: Use of administrative, economic and functional classification standards.

2.3 **Transparency of intergovernmental fiscal relations** for child protection. Benchmark: Transparency and objectivity in the horizontal allocation of transfers linked to expenditures on child protection among sub-national governments.

2.4 **Conduciveness of public financing framework** to development of balanced and child-focused menu of Child protection services. Benchmark: Neutrality of financing framework with regard to types of Child protection services and an absence of financial incentives that have detrimental effects on children, for example, capitation payments that provide incentives to place children in residential care.

2.5 **Realistic intergovernmental financing** of Child protection programmes Benchmark: Accuracy of reimbursement of financial costs imposed on sub-national budgets by the central government within decentralized spending programmes related to child protection.

2.6 **Multiyear perspective in fiscal planning**, expenditure policy and budgeting. Benchmark: Preparation of multiyear fiscal forecasts and functional allocations.


The above draw heavily on the PEFA framework.

**Scope of expenditures considered**

Focusing on child protection systems at the national level, but some indicators have been designed specifically to assess and measure the effectiveness of decentralization arrangements salient for child protection outcomes. In principle, the overall Indicator Framework is also open for adaptation for possible application to sub-national governance systems (although no adapted template is yet available).

The scope of expenditures considered differs by indicator: Given that CP expenditures are fragmented across several sectors and are not readily reported, the benchmark for indicator 2.1 covers all sectors with relevance to child protection, with a view to measuring how stable allocations across these segments of public spending are. Other indicators, such as 2.3 and 2.5 covers all transfers that are linked to expenditure on CP. 2.6 on multi-year forecast, is focused on overall expenditure totals for all sectors of spending which are usually prepared in an aggregated medium-term expenditure framework, as what matters for the sustainability of child protection expenditure in terms of a multiyear perspective is the political willingness and technical capacity of the government to engage in medium-term budgeting, with a particular focus on forecasting resulting public deficits and debts, and preparing to fund them in a responsible way.

**Spheres of expenditure**

In terms of expenditures, the focus in on public funds, but other aspects consider CSOs and private sector as Governance of CP encompasses the state, but it transcends organs of the state by including the private sector and civil society organizations.

**Stock / Flow**

Some of the indicators are related to a stock measure in a single year, others (such as 2.1) over multi-years.
A framework of core indicators for measuring and monitoring national child protection systems which can be disaggregated by domain or indicator.

Under each criteria are four benchmark, which receive a score between 0 and 1 to the extent that they have been achieved. These are then aggregated (summed) for an overall score that is presented in a colour coded results matrix. This is accompanied by a report, detailing key strengths and weaknesses, and progress and challenges against each domain.

Data acquisition process

The indicators had not, at the time of the report, been put against the practical realities of actual data collection and of building dialogue with local stakeholders and the government on the possible role of this tool.

Government engagement is critical for collecting the necessary information for the indicator framework as it spans several public administration fields and which must furthermore be broken down to be made relevant to CP.

Indicator framework is “as much about the process as it is about the product” requiring ongoing refinement and expansion beyond the core set proposed in response to data as they become available and to related legal and policy developments.

Recognition that other international actors have stronger expertise in the area of governance – e.g. World Bank and UNDP – and UNICEF may have to work in partnership with them.

Strengths

Framework is flexible so as to capture a significant amount of country-specific qualitative perspectives within the standardized scores. This allows users of the Indicator Framework to combine opportunities of broad regional comparisons with generating assessments useful for more in-depth policy work at a country level.

Limitations

Regarding the indicator framework as whole:

- **Large size and scope** has implications for feasibility, affordability, and sustainability of needed analytical effort. The framework requires a lot of technical rigour particularly in view of the coverage of several sectors in public administration (such as public financial management, human resources management, decentralization arrangements, statistics and information management) and, at the same time, the need to break these data down by categories to make them relevant for sectorwide analysis specific to child protection. This demands considerable knowledge of related concepts and of local institutional contexts to arrive at meaningful assessments. In practical terms, the sheer number of new themes and concepts that have to be analysed implies a significant effort is initially required.

- **Reliance on qualitative analysis**, which whilst helping to capture specifics of the context, may also create opportunities for manipulation with ratings and with the depth of analysis undertaken to arrive at standardized scores, which is difficult to verify.

- **Does not provide a very robust way of assessing progress in CP systems over time** as the methodologies and information sources of some indicators change from year to year either due to improvements in the construction of indicators or the (un)availability of data sources.

- While standardized, this instrument is not **sufficient to make far-reaching conclusions on the relative rank of individual countries**. Standardized scores provided in the Indicator Framework are designed to “help countries to broadly assess their progress compared to regional averages to stimulate policy dialogue and reforms, rather than compete for higher ranks”.

- **Cannot prescribe reform measures or guide practical steps for action**: it is a broad instrument aimed at initial diagnostic assessment; it should be treated as an entry point to the reform process, rather than a method for formulation of quick and concrete recipes for change. However, it can point to the principal gaps, weaknesses and distortions which require attention and change.

Annex p51
- The indicators focus deliberately on the processes relevant to child protection, and not on the actual results such as prevalence rates of maltreatment. While this is a deliberate and defining feature of the Framework, its lack of coverage of actual outcomes may be received with caution.

- Although not yet tested in country, expectation that (as with any other framework for the measurement of policy progress), Child Protection Indicators would be difficult to implement and sustain over time. Ensuring results are utilized in local policies would also be problematic if they are not ‘owned’

**Regarding the PFM indicators:**

- they help consolidate basic information whether core systems are in place to ensure strategic, realistic and efficient use of funds, but this is not sufficient to make an ultimate assessment of PFM systems.

- Corruption is deemed too complex to include in the framework, whilst recognising that it is closely linked to the quality of governance systems.

- The assessment of national PFM systems is linked to an analysis of budget data which are not always readily available. In particular, child protection requires analyses which are ‘sector-wide’ and stretch across traditional boundaries of existing ministries and sectors (such as healthcare and education).

**Sources**


---

**REVISED SOCIAL PROTECTION INDEX (SPI) – Asian Development Bank**

**Summary**

The revised SPE, adopted in 2010 is designed to measure national social protection programmes and their broad impact on the poor and vulnerable in Asia and the Pacific, in a standardised way.

**Primary Purpose**

The SPI is not designed for ranking the relative performance of countries, although it is internationally comparable. Rather, it is designed for practical use to assess the nature of social protection programs at country level and to identify these programs’ broad impact on the poor and vulnerable.

**Construction**

**Numerator:** Total social protection expenditures per total reference population

Equivalent to: total expenditures/total beneficiaries (depth of CP coverage) × total beneficiaries/total reference population (breadth of SP)

**Denominator:** regional poverty-line expenditures (25% GDP/capita)

This normalises the SPI. It is an average of national poverty lines across the region expressed as percentage (25%) of GDP/capita. It denotes the threshold of total expenditures that each person needs to exceed to be considered non-poor. It eliminates the problem that expenditures are expressed in national currencies since national currencies drop out of the two combined ratios. It is an alternative to just using GDP/capita which would make the index small and would not carry much normative significance since no one would plausibly argue that SP expenditures should approximate average income per person. The choice of a uniform percentage of GDP per capita mitigates the problem that some countries set their national poverty lines very low while others set them very high.
Given this construction, the SPI can be interpreted directly as SP expenditures per potential beneficiary as a percentage of per capita poverty line expenditures.

For example, if the SPI were 0.100 in country X, this index number would mean that total social protection expenditures (per intended beneficiary) represent 10% of poverty-line expenditures. The higher this index number, the better a country’s performance.

The overall SPI is a weighted sum of the SPIs for social insurance, social assistance, and labor market programs. The weights are the relative sizes of the groups of potential beneficiaries of each of these three major programs. However, at the same time, unweighted SPIs for the three major programs of social protection (simple ratios of expenditures to intended beneficiaries) are presented as this reveals the scale of program-level spending relative to the intended beneficiaries—indeed independently of the relative size of the three groups of beneficiaries.

### Scope of expenditures considered

**SP services** included are (i) social insurance, that is, the categories of old-age insurance, programs for the disabled, and health expenditures on insurance and pensions; (ii) social assistance, that is, the categories of non-contributory health insurance, conditional cash transfers, child protection, and unconditional cash transfers; and (iii) labor market programs.

The emphasis is on obtaining information for the **largest SP programs** in terms of expenditure, number of beneficiaries, gender, and poverty. Smaller programmes can be excluded given their minimal influence on summary results.

**SP activities not amenable to quantification are excluded**, such as legislation on labor standards, women and children’s rights, as well as empowerment and consciousness-raising initiatives.

**Programs that are generally seen as falling within the health or education sectors are excluded**, such as health care (including HIV/AIDS and reproductive health), immunization, nutrition, preschool education, and general vocational and technical education.

**Some effort to collect decentralised data** is included in the handbook, but recommended only in countries that are sufficiently decentralised to merit the additional effort. In countries where administration of some SP programmes are decentralised, the first port of call is the central government which often collects adequate information. The consultant should visit 2/3 local administrations to collect information on the types of programmes administered locally to judge whether these programmes are large enough to affect the summary results. If yes additional visits and extrapolation need to be made to produce national estimates of expenditures and number of beneficiaries for these programs. Based on the previous SPI exercise, it is unlikely that SP programs financed by provincial governments will be significant except in the largest countries, such as the People’s Republic of China, India, and Indonesia.

### Spheres of expenditure

Data for the revised SPI are confined to **government or government-supported programs**, but include Government and international development agency funding to those programmes.

This is a departure from the original SPI which included projects implemented independent to Government. The basic motivation for this change in approach is that the current project assumes that governments should take ownership of their social protection systems and be held accountable ultimately for the success or failure of those systems.

### Stock / Flow

Data for the basic statistics are only collected for the reference year. Data on SP programs should be collected for the three most recent calendar or fiscal years.

### Presentation

The revised SPI is a **unitary summary indicator**. Note: this is a departure from the original SPI which was a summary composite index.
The index can be **disaggregated** or decomposed in different ways, such as by depth and breadth of coverage (depth means the average size of benefits received by actual beneficiaries, and breadth means the proportion of intended beneficiaries who actually receive benefits); by category of SP (including social insurance, social assistance, and labour market) or component programmes; by its impact on the poor and the nonpoor, as well as into its impact on women and men.

**Data acquisition process**

Data is collected by SPI country consultants. There is usually one consultant per country, who should be assigned a Government person or agency to be the prime contact point. The contact point should be able to provide direct access to data on most SP programmes and introductions to other agencies. International finance institutions may also provide useful contacts and introductions.

**Strengths**

- What each country consultant will need to do to calculate the revised SPI is not technically complex, but may be time consuming.
- Preconfigured spreadsheets and detailed guidelines aides standardised collection and calculation.
- The indicator has a directly meaningful interpretation: SP expenditures per potential beneficiary as a percentage of per capita poverty line expenditures.

**Limitations**

- Data needed are unlikely to be available from centralized or easily accessible sources, and obtaining them requires persistence in following up, clarifying, and double checking. Unlike education or health data, information on social protection programs is not centralized in one ministry or government department. This is an indication of the general lack of strategic focus for social protection. Invariably, no one government agency is given the central authority to coordinate social protection efforts. This lack of centralization of efforts meant that national consultants had to spend a great deal of time with various government departments to gather comprehensive data. Such efforts could thus involve coordinating with a wide range of government institutions, ranging across ministries of labour, social welfare ministries, and ministries of education and health, among others.
- Given the above time requirement, during the 2011 data collections process, **funding was inadequate to finance national consultants to engage directly with subnational units**, such as state or local governments, some of which might be directly implementing their own social protection programs. In large countries, such as the People’s Republic of China, India, and Indonesia, this is likely to lead to a greater underestimation of the extent of social protection programs than in smaller countries, where independent subnational programs are less likely to exist.
- There are **significant exclusions** in terms of SP programmes not considered.
- It is recommended that some SP programmes – specifically for example CP – be grouped into a single expenditure entry (as long as they are implemented by the same agency, targeted at the same group with similar eligibility criteria, and small).
- One of the most difficult aspects of computing the SPI is the identification of the appropriate categories of “intended beneficiaries” for each of the three major social protection programs and for each of their major subcomponents. Since many forms of social insurance, such as pensions or health insurance, are contributory, it was decided to use the employed as the most appropriate potential beneficiaries of some of the main forms of social insurance (particularly health insurance and unemployment insurance). For pensions, the elderly (those aged 60 years or older) were designated the potential beneficiaries. If, however, health insurance is universal, then the SPI project used the entire population as potential beneficiaries. For social assistance, the most difficult group to identify, in practical terms, was those affected by disasters.

**Sources**


Asian Development Bank, 2013. The Social Protection Index: Assessing Results for Asia and the Pacific

Annex p54
**SOCIAL PROTECTION INDEX FOR ASIA (SPI)**

**Summary**

The SPI provides a summary measurement tool of the extent to which Asian and Pacific countries provide welfare, labour market, social security, health insurance, microcredit, child protection, and targeted education and health support programs to their citizens, and especially those living below the poverty line.

**Primary Purpose**

The SPI was designed to be a starting point for diagnostic evaluations of national social protection programs to enable the monitoring of changes in social protection provision over time, and to facilitate international comparisons of social protection provision.

The SPI and its summary indicators provide the starting point for the appraisal of existing SP provision, and help policy-makers identify key issues concerning the level and effectiveness of a country’s SP programmes. Once this initial stage has been achieved, attention then needs to be focused on the information on the coverage of the priority target groups (the subcomponents of the coverage indicator) and the targeting of individual SP programmes. This more detailed analysis can identify where there are major problems in a national SP system, what are the reasons for this, and how the situation could be improved. This approach to using the SPI and its constituent information is considered analogous to the use of the HDI.

Furthermore, the authors believe that the SPI and the information used in its computation can become an important tool in comparing and examining countries’ SP activities by: enabling international comparisons of SP activities to be made, as well as the monitoring of these activities by individual countries over time; providing a database which can be used to define SP targets both internationally and nationally; providing the basis for more rigorous diagnostic evaluations of countries’ SP provision; raising awareness of the importance of social protection through the very process of obtaining the information needed to calculate the SPIs and the subsequent dissemination of the results.

**Construction**

The SPI, conceptually based on the human development index, was derived from four social protection summary indicators (SPSIs), namely, social protection expenditure (SPEXP), social protection coverage (SPCOV), social protection distribution/poverty targeting (SPDIST), and social protection impact on incomes of the poor (SPIMP).

<table>
<thead>
<tr>
<th>SPSI</th>
<th>Numerator</th>
<th>Denominator</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPEXP</td>
<td>Total expenditure on all social protection programs</td>
<td>Gross domestic product</td>
</tr>
<tr>
<td>SPCOV</td>
<td>Beneficiaries of social protection programs targeted at key target groups</td>
<td>Reference populations for key target groups</td>
</tr>
<tr>
<td>SPDIST</td>
<td>Number of social protection beneficiaries who are poor</td>
<td>Poor population</td>
</tr>
<tr>
<td>SPIMP</td>
<td>Average social protection expenditure for each poor person</td>
<td>Poverty line</td>
</tr>
</tbody>
</table>

It is modelled closely on the HDI under the expectation that this would “increase the likelihood of the SPI becoming widely accepted”.

The indicators are then scaled and then aggregated using an average due to equal weighting assigned to each (as with the HDI)

**Scope of expenditures considered**

Five kinds of SP activities are included:

- Labour Market Programs
- Social Insurance Programs

Annex p55
- Social Assistance and Welfare Programs
- Micro and Area-Based Schemes (community-based)
- Child Protection; including:
  - Early child development activities—e.g., basic nutrition, preventative health, and educational programs
  - Educational assistance (e.g., school-feeding, scholarships, fee waivers)
  - Health assistance (e.g., reduced fees for vulnerable groups)
  - Street children initiatives
  - Child rights and advocacy/awareness programs against child abuse, child labour etc,
  - Youth programs to reduce health risks (especially HIV/AIDS and drugs) and anti-social behaviour.
  - Family allowances (e.g., in-kind or cash transfers to assist families with young children to meet part of their basic needs).

**Non-targeted programmes that are generally seen as falling within the health, education and rural development programmes are excluded.** It was noted that this “facilitates the acceptability of the results by national governments and international financing institutions by excluding programs and activities that are normally seen as falling within the ambit of other sectors”

SPSI can be applied to subnational levels and have been used to show disparities in rural and urban areas.

**Spheres of expenditure**

Government, private sector and nongovernment organization expenditures are included in the calculation of SPEXP. i.e. the original SPSI included programmes supported by external development agencies that were implemented largely independently of governments. The same criteria have been applied to programs administered by nongovernment organizations.

**Stock / Flow**

Indicators are for specific years

**Presentation**

SPI is a final summary score, which can be broken down into its composite indicators. SPEXP and SPCOV and be further disaggregated to their subcomponents.

**Data acquisition process**

Data compilation undertaken by local consultants. A 3-year updating cycle is envisaged.

**Strengths**

The scaling up of the TA 6120 methodology to over 20 additional countries in Asia has confirmed the formulation of the expenditure (SPEXP), poverty targeting (SPDIST), and impact on income indicators (SPIMP) to be “robust” according to the ADB report.

Facilitates the acceptability of the results by national governments and international financing institutions by excluding programs and activities that are normally seen as falling within the ambit of other sectors

**Limitations**

The data requirements for calculating the SPI are substantial, involving a mixture of primary and secondary data sources, some of which will be published and be publicly available, but most of which will only be available via unpublished reports or the reports of the government departments, NGOs and private organisations involved in SP activities.

The formulation of the coverage indicator (SPCOV) is problematic due to the lack of a one-to-one correspondence between program beneficiaries and reference populations for several of the designated target groups such as the unemployed, the poor, and children with special needs due to the existence of
beneficiaries not forming part of the reference population and the consequent potential for the number of beneficiaries to exceed the reference population.

Significant exclusions by not considering expenditures in other sectors

The percentage of the population receiving SP is a separate indicator from the indicator registering the size of their benefits, but as, for instance, a large share of the reference population could receive very small benefits or a small share of the reference population could receive very large benefits, neither a high degree of coverage nor a large size of received benefits would necessarily be desirable on its own. This was one rationale for subsequent revision of SPI.

(SPEXP) compares the magnitude of SP benefits to the size of GDP instead of the number of beneficiaries. Hence it gives only an indirect indication of the impact of the expenditures on human well-being. This was one rationale for subsequent revision of SPI.

Another concern was that the abstract index form of the original SPI was difficult to interpret.

Any weighting scheme—even such a simple one as equal weighting of components—always leaves itself open to the criticism that it is essentially arbitrary.

A particularly contentious issue in the application of the original SPI was its ranking of countries’ performance in social protection. While such an approach might succeed initially in capturing people’s attention, it is likely to be less successful in encouraging governments to make practical use of the SPI.

Sources:

Asian Development Bank, 2013. The Social Protection Index: Assessing Results for Asia and the Pacific

HUMAN DEVELOPMENT INDEX (HDI)

Summary

The Human Development Index (HDI) is a summary measure of key dimensions of human development. It measures the average achievements in a country in three basic dimensions of human development: a long and healthy life, access to knowledge and a decent standard of living. The HDI is the geometric mean of normalized indices from each of these three dimensions.

Primary Purpose

The HDI was part of an effort to shift attention away from an overly large focus on economic growth as the objective of development policies. In this sense, one of the main purposes of the HDI is to play an advocacy role, seeking to re-establish balance in priorities to counteract a focus distorted by a growth-centric approach. Its primary purpose is cross-country comparison derived from the rankings. The HDI can also be used to question national policy choices, asking how two countries with the same level of GNI per capita can end up with such different human development outcomes.

Construction

The HDI measures the average achievements in a country in three basic dimensions of human development:

- a long and healthy life, (indicator: life expectancy at birth)
- access to knowledge (indicators: mean years of schooling at age 25; expected years of schooling for children entering school age)
- decent standard of living (indicator: GNI per capita, PPP$)

Minimum and maximum values (goalposts) are set in order to transform the indicators into indices between 0 and 1. The maximums are the highest observed values in the time series (1980–2012). The minimum values

Annex p57
can be appropriately conceived of as subsistence values. The minimum values are set at 20 years for life expectancy, at 0 years for both education variables and at $100 for per capita gross national income (GNI).

The dimension indices are calculated as follows:

**Nominator:** actual value – minimum value

**Denominator:** maximum value – minimum value

The HDI is the geometric mean (cube root) of the three dimension indices: life, education and income

The Inequality-adjusted Human Development Index (IHDI) adjusts the Human Development Index (HDI) for inequality in the distribution of each dimension across the population. The IHDI accounts for inequalities in HDI dimensions by “discounting” each dimension’s average value according to its level of inequality. The IHDI equals the HDI when there is no inequality across people but falls further below the HDI as inequality rises. In this sense, the IHDI is the actual level of human development (taking into account inequality), while the HDI can be viewed as an index of the “potential” human development that could be achieved if there was no inequality. The “loss” in potential human development due to inequality is the difference between the HDI and the IHDI and is expressed as a percentage.

HDI indicators can be adapted at the country level provided they meet other aspects of statistical quality.

### Scope of expenditures considered

<table>
<thead>
<tr>
<th>N/A</th>
</tr>
</thead>
</table>

### Spheres of expenditure

<table>
<thead>
<tr>
<th>N/A</th>
</tr>
</thead>
</table>

### Stock / Flow

HDI is up of both stock and flow variables. This is reportedly a source of frustration for many governments

### Presentation

HDI is a Summary measure, between 0 and 1 which gives way to a ranking of countries. The countries are grouped into categories Very High, High, Medium, Low, by quartile of the HDI distribution.

The HDI can be disaggregated across population sub-groups defined relative to income, geographical or administrative regions, urban/rural residence, gender and ethnicity

It can also be disaggregated at sub-national level to compare levels and disparities among different subpopulations within a country, provided that appropriate data at the level of disaggregation are available; or can be estimated using sound statistical methodology.

### Data acquisition process

The data is collected by the Human Development Report office from existing sources. Life expectancy at birth is provided by the UN Department of Economic and Social Affairs; mean years of schooling are based on UNESCO’s Institute for Statistics (UIS) educational attainment data and Barro and Lee methodology; expected years of schooling are provided by UIS; and GNI per capita by the World Bank and the International Monetary Fund. For a few countries, mean years of schooling are estimated from nationally representative household surveys, and for few countries GNI was obtained from the UN SNA Main Aggregates database. The Human Development Report Office does not collect data directly from countries.

### Strengths

- The simplicity of the index’s characterization of development (as an average of achievements in health, education and income), linked to the basic message that development is about much more than growth, has contributed to its popularity.

### Limitations

- HDI’s very simplicity prompted critiques from the start, with some contending that it was too simplistic. Its choice of indicators is severely restricted by data availability. Many data gaps still exist in even some very basic areas of human development indicators.
The HDI is not a comprehensive measure of wellbeing: The policy of the Human Development Report Office has always been to construct additional complementary composite indices for covering some of the “missing” dimensions in the HDI. Gender disparity, inequality and human deprivation are measured by other indices (see Gender Inequality Index, Multidimensional Poverty Index and Inequality-adjusted HDI). Participation and other aspects of well-being are measured using a range of objective and subjective indicators and are discussed in the Report.

It has also been noted that the reported indicators do not truly capture variations across countries in the relevant dimensions. Life expectancy is a measure of longevity but tells us nothing about the health of people during the time that they are alive.

There are constant marginal returns to improvements in each dimension, and therefore the marginal rate of substitution between dimensional achievements is also a constant. This would seem to run counter to the intuition that, the worse the deprivation in a particular dimension, the more urgent the efforts to improve achievements in that dimension should be regarded.

The choice of equal weights has been criticized.

The main disadvantage is that the IHDI is not association sensitive, so it does not capture overlapping inequalities. To make the measure association sensitive, all the data for each individual must be available from a single survey source, which is not currently possible for a large number of countries.

Sources:

CHILD AND YOUTH WELL-BEING INDEX (CWI)

Summary
Each year, the Foundation for Child Development and the Child and Youth Well-Being Index Project at Duke University issue a comprehensive measure of how children are faring in the United States. The resultant National Child and Youth Well-Being Index (CWI) is based on a composite of 28 Key Indicators of Well-Being, grouped into seven Quality of Life/Well-Being Domains. These Domains are: Family Economic Well-Being, Safe/Risky Behaviour, Social Relationships, Emotional/Spiritual Well-Being, Community Engagement, Educational Attainment, and Health.

Primary Purpose
In-country monitoring: The objective of the CWI is to give a view of changes over time in the overall well-being of children and youth in the United States. The emphasis is on the direction of change in overall well-being, as compared to the 1975 base year of the indicators, rather than the value itself.

Construction
Annual time series data (from vital statistics and sample surveys) is collected against 28 Indicators which are spread over 7 Quality of Life/Well-Being Domains. These Domains are: Family Economic Well-Being, Safe/Risky Behaviour, Social Relationships, Emotional/Spiritual Well-Being, Community Engagement, Educational Attainment, and Health. Each domain has between 2 and 6 associated indicators.

Where data are not observable, they are projected by use of statistical time series models.

The 28 indexed Key Indicator time series are grouped into the seven Domains of Well-Being by equal weighting to compute the Domain-Specific Index values for each year. The composite Index is an equally-weighted average of the seven Quality of Life/Well-Being Domains.

Scope of expenditures considered
N/A

Spheres of expenditure
N/A

Stock / Flow
Annex p59
Stock: all indicators are computed annually

**Presentation**

The CWI is a *composite summary index*. The value of the CWI in the base year 1975 set equal to 100 so that a value less/more than 100 represents a deterioration/improvement in overall child and youth well-being, as compared to 1975 base-year values.

**Data acquisition process**

The index is computed and updated annually in a collaborative effort between the Foundation for Child Development and the Child and Youth Well-Being Index Project at Duke University, using existing sources of vital statistics and surveys.

**Strengths**

- CWI has had significant impact because it represents one of the first significant efforts to summarize child well-being in one single number, for the United States, using longitudinal data, which allows tracking child well-being through time and identifying trends.

**Limitations**

- The usage of aggregated data only allows to describe the proportion of children with a particular outcome, as opposed to microdata which allows to determine whether an individual child has one or more particular outcomes
- Construction of this index has different origins; it consists of an array of indicators compiled from different surveys which means that the sample is not stable throughout that set of indicators.
- Children’s own views about their well-being are not properly considered, more often the unit of analysis of surveys is the family.
- Not all of the indicators are robust measures of the outcomes in question. Some domains rely on two indicators which don’t comprehensively reflect the outcome.
- No interactions between dimensions are considered

**Sources:**

Fernandes 2010, A review essay on child well-being measurement: uncovering the paths for future research

Foundation for Child Development, 2013, NATIONAL Child and Youth Well-Being Index (CWI)

---

**PUBLIC EXPENDITURE AND FINANCIAL ACCOUNTABILITY (PEFA)**

**Summary**

PEFA measures the operational performance of the key elements of the PFM systems, processes and institutions of a country central government, legislature and external audit, against six *core dimensions of an open and orderly PFM system*

1. Credibility of the budget - The budget is realistic and is implemented as intended
2. Comprehensiveness and transparency - The budget and the fiscal risk oversight are comprehensive and fiscal and budget information is accessible to the public.
3. Policy-based budgeting - The budget is prepared with due regard to government policy.
4. Predictability and control in budget execution - The budget is implemented in an orderly and predictable manner and there are arrangements for the exercise of control and stewardship in the use of public funds.
5. Accounting, recording and reporting – Adequate records and information are produced, maintained and disseminated to meet decision-making control, management and reporting purposes.
6. External scrutiny and audit - Arrangements for scrutiny of public finances and follow up by executive are operating.

**Primary Purpose**

The primary purpose of the PEFA is *in-country monitoring*: A PEFA assessment sets the stage for preparing measures to strengthen PFM system performance (perhaps formalised as a revised PFM reform strategy) and

Annex p60
subsequently for measuring the performance impact of its implementation by providing information on the extent to which country PFM performance is improving or not. In addition, the PFM-PR recognizes the efforts made by government to reform its PFM system by describing recent and on-going reform measures, which may not have yet impacted PFM performance.

While the PEFA framework was not originally intended for the purposes of cross-country comparisons, it is used for this purpose, often by comparing scores for each indicator and using the narratives in the PEFA reports (PFM-PRs) to draw conclusions on reasons for differences in scores. Where this includes more that a few countries, researchers have sought an aggregation of the results of a country assessment. There is no PEFA-methodology for this, but typically it involves both the conversion of ordinal indicator ratings to numerical value and allocation of weights to the individual indicators. Where country results are clustered, aggregation also requires assignment of weights to each country.

**Construction**

**High-level performance indicators:** The selected 28 indicators for the country’s PFM system are structured into three categories:

a. PFM system out-turns: these capture the immediate results of the PFM system in terms of actual expenditures and revenues by comparing them to the original approved budget, as well as level of and changes in expenditure arrears.

b. Cross-cutting features of the PFM system: these capture the comprehensiveness and transparency of the PFM system across the whole of the budget cycle.

c. Budget cycle: these capture the performance of the key systems, processes and institutions within the budget cycle of the central government.

In addition to the indicators of country PFM performance, this framework also includes

d. Donor practices: these capture elements of donor practices which impact the performance of country PFM system.

Each indicator seeks to measure performance of a key PFM element against a four point ordinal scale from A to D. Guidance has been developed on what performance would meet each score, for each of the indicators.

Most of the indicators are composite and the overall score for an indicator is then based on the assessments for the individual dimensions of the indicator. Intermediate scores (denoted by a ‘+’ sign) available for composite indicators that include two or more sub-dimensions. Combining the scores for dimensions into the overall score for the indicator is done by one of two methods Method 1 (M1) for some indicators (single dimension/weakest link); and Method 2 (M2) for other indicators (simple average).

**PFM Performance Report:**

The PFM–PR is a concise document, which has the following structure and content:

- A summary assessment uses the indicator-led analysis to provide an integrated assessment of the country’s PFM system against the six core dimensions of PFM performance and a statement of the likely impact of those weaknesses on the three levels of budgetary outcomes.
- An introductory section presents the context and the process of preparing the report and specifies the share of public expenditures captured by the report.
- A section presents country-related information which is necessary to understand the indicator-led and overall assessment of PFM performance.
- The main body of the report assesses the current performance of PFM systems, processes and institutions based on the indicators, and describes the recent and on-going reform measures implemented by government.
- A section on government reform process briefly summarizes recent and ongoing reform measures implemented by government and assesses the institutional factors that are likely to impact reform planning and implementation in the future.

**Scope of expenditures considered**

The focus of the PFM performance indicator set is the public financial management at central government level, including the related institutions of oversight.

Additional guidelines are available for application at sub-national levels. Minor changes to the wording of some rating criteria allow it to be applied to the diverse structures of sub-national governments. If the
purpose of the assessment is to gain an opinion of the overall state of PFM at the SNG level, a sample of SNGs are usually selected. The adapted template is based on in-depth analysis of the diverse possible models of decentralization, including distinctions between: Fiscal, Administrative and political decentralization; Federal and unitary states; and Francophone and Anglophone countries.

If two or more levels of government have been assessed e.g. central and state, it is not considered appropriate to aggregate the results. In general, aggregation of scores even at the same level is not recommended as key differences will be lost. Also, differences across SNG create methodological difficulties for the aggregation of results. Indeed some care must be taken in ranking or even direct comparison as they may become highly politicized. However, some governments may wish to rank their own SNGs which could be done using the statistical mode or by using the median value.

### Spheres of expenditure

#### Public

#### Stock / Flow

Diffs by indicator: most are an annual figure, but indicators PI-1, PI-2, PI-3 and D-1 require data for three years as a basis for the assessment. The data should cover the most recent completed fiscal year for which data is available and the two immediately preceding years. The assessment is based on the performance in two out of those three years i.e. allowance is made for one year to be abnormal (and not contributing to the score) due to unusual circumstances such as external shocks.

### Presentation

Summary scores (A-D) across a framework of 28+3 indicators, presented next to previous PEFA score (where available).

### Data acquisition process

The assessment may be undertaken by means of one of the following three assessment models:

i) a self-assessment undertaken by the government (with arrangements for external validation);  
ii) a joint assessment i.e. government working with other stakeholders – e.g. development partners, domestically-based academia and civil society organisations;  
iii) an external assessment led by a non-government stakeholder, with technical and logistical support provided by government, (may be preferable if the government is facing capacity and time constraints).

An oversight team (OT) made up of Ministry of Finance, other government agencies involved and non-government stakeholders, such as the Auditor General’s Office, Parliament, and development partners play the governance role in the assessment process. The OT may delegate an individual, here called the Assessment Manager (AM) to manage the assessment process, starting off with the selection of the Assessment Team (AT).

The PEFA’s periodicity should be a minimum of 3 years and maximum of 5 years

### Strengths

- Coverage is impressive, 90 per cent of low income, 75 per cent of middle income and 8 per cent of high income countries had been assessed, were in the process of assessment or were going to be assessed by October 2010.  
- The PEFA Programme has succeeded in creating a credible framework for the assessment of PFM functionality, which manages to be comprehensive in its coverage and yet sufficiently simple for the non-technical user to understand.  
- The framework is comparable over time and, subject to certain caveats (see below), also comparable across countries.  
- In recent years the PEFA has begun to display evidence of global impact. The 2011 evaluation found that, within the sample of countries assessed, progress is being made in the use of the PEFA framework to enhance country ownership of reforms, and to improve alignment of donor support to PFM reform.

### Limitations
• Aggregation is potentially hazardous: Comparing progress across time and space requires the conversion of ordinal into numerical scores or using a frequency distribution. Issues associated with this include whether it is valid to assume that quality differences between each grade are the same and that changes in scores mean the same for all initial scores, how to handle absent scores.

• Other issues arise regarding cross-country comparison due to not comparing like with like: The scope (central government only, SNG only, combined) and year of the assessment may not be the same for the countries being compared; definitions of PFM terms may differ across countries, the reasons for the same scores across countries for each indicator may vary considerably, as most indicators have at least two and sometimes four dimensions, the quality of the assessments may vary, there may be differing proportions between countries of public expenditure financed directly by the government and financed directly by donors, and the case for assigning weights will vary between countries.

• Quality of assessments varies: the secretariat is not the “owner” of the assessments and not all pass through the secretariats QA system. Due to the reputational risks generated even by isolated cases of poor quality PEFA assessments, this is a very serious observation.

• There is a lack of awareness of the scope and potential of the PEFA assessment framework amongst governments, and its potential as a benchmarking tool around which to organise dialogue with donors, or with sub-national governments, or with civil society, is not everywhere fully appreciated, in part because it is perceived by some stakeholders as a tool of fiduciary risk assessment, rather than a more developmental, diagnostic tool.

• The Governance structure includes relatively small number of PEFA Partners which allows for efficiency but it only permits a limited number of stakeholder voices, excluding many key development agencies, professional bodies, non-OECD governments and NGOs.

• 2011 Impact assessment found that despite the PEFA Framework being used as a common pool of information, there was evidence of duplicative broad based diagnostic work. Government officials also noted that DPs are conducting a considerable amount of uncoordinated PFM assessments to inform their own operational decision making that places an additional burden on government time.

• The Study provided little evidence of any impact of PEFA assessments beyond central finance/planning ministries (e.g. line ministries, SNGs, SAIs and Parliament). This preliminary finding, uncorroborated by a more comprehensive in-country mission also suggests that more effort needs to be made to address issues of communicating results, public disclosure and improving dissemination workshops which incorporate the media and NGOs.

Sources:
PEFA, 2009, Issues in Comparison and Aggregation of PEFA Assessment Results Over Time and Across Countries

De Renzio, 2009, Taking Stock: What do PEFA Assessments tell us about PFM systems across countries?

PEFA 2011, PEFA - PFM Performance Measurement Framework


Mackie and Caprio, 2011, Assessing the Impact of the PEFA Framework: A Study for the PEFA Steering Committee

PEFA, 2012, Good Practices in Applying the PFM Performance Measurement Framework


PEFA, 2013. Supplementary Guidelines for the application of the PEFA Framework to SNGs

WORLDWIDE GOVERNANCE INDICATORS (WGI) – World Bank

Summary

World Governance Indicators (WGI) are six composite indicators measuring broad aspects of governance: Voice and Accountability, Political Stability and Absence of Violence/Terrorism, Government Effectiveness,
Regulatory Quality, Rule of Law, and Control of Corruption. The aggregate indicators are based on several hundred individual underlying variables, taken from a wide variety of existing data sources. The data reflect the views on governance of survey respondents and public, private, and NGO sector experts worldwide.

Primary Purpose

The individual indicators can be used to make comparisons of countries over time, as underlying sources use reasonably comparable methodologies from one year to the next. Confidence intervals allow an assessment of whether the change is statistically significant.

They also can be used to compare the scores of different countries on each of the individual indicators, recognizing however that these types of comparisons too are subject to margins of error, however if confidence intervals do not overlap, the comparison can be assumed to be significant.

Construction

The six composite indicators are based on several hundred variables obtained from 31 different data sources, capturing governance perceptions as reported by survey respondents, nongovernmental organizations, commercial business information providers, and public sector institutions.

For each of these six clusters an Unobserved Components Model is used to (i) standardize the data from these very diverse sources into comparable units, (ii) construct an aggregate indicator of governance as a weighted average of the underlying source variables, and (iii) construct margins of error that reflect the unavoidable imprecision.

All the individual variables have been rescaled to run from zero to one, with higher values indicating better outcomes.

Scope of expenditures considered

Defines governance as “the traditions and institutions by which authority in a country is exercised. This includes (a) the process by which governments are selected, monitored and replaced; (b) the capacity of the government to effectively formulate and implement sound policies; and (c) the respect of citizens and the state for the institutions that govern economic and social interactions among them.” Two measures of governance corresponding to each of these three areas, resulting in a total of six dimensions of governance:

- Voice and Accountability,
- Political Stability and Absence of Violence/Terrorism,
- government Effectiveness,
- Regulatory Quality,
- Rule of Law,
- Control of Corruption.

Spheres of expenditure

N/A

Stock / Flow

Annual (almost all data sources available annually, in a few cases data sources are updated only once every two or three years, in which cases WGI uses data lagged by one or two years to construct an estimate)

Presentation

Six composite indicators, reported in in two ways: in the standard normal units of the governance indicator, ranging from around -2.5 to 2.5, and in percentile rank terms ranging from 0 (lowest) to 100 (highest) among all countries worldwide.

Data acquisition process

Data obtained from 31 different data sources, capturing governance perceptions as reported by survey respondents, nongovernmental organizations, commercial business information providers, and public sector institutions. Data sources include surveys of firms and households, as well as the subjective assessments of a variety of commercial business information providers, non-governmental organizations, and a number of multilateral organizations and other public-sector bodies.
Strengths

- Explicitly recognizes imprecision, and reports the margins of error enabling identification of statistically significant changes over time or comparisons between countries.
- Very comprehensive, bringing together and averaging a large number of data sources.

Limitations

- Relies exclusively on perceptions-based governance data sources leaving it open to various perception biases (but it is noted, perceptions data have particular value in the measurement of governance).
- Small changes from year to year in the set of sources on which the WGI scores are based undermine the robustness of WGI to track progress over time; however wherever possible the team make changes consistently for all years in the historical data as well.
- Differences in availability of some indicators in some countries undermines the robustness of comparison between countries: For example, the WGI “Control of Corruption” for Eastern Europe and Central Asia has 23 different combinations of sources, but only four pair of countries ratings are based on a common set of sources.

Sources:


OPEN BUDGET INDEX (OBI)

Summary

The Open Budget Survey measures the state of budget transparency, participation, and oversight in countries around the world. It consists of 125 questions and is completed by independent researchers in the countries assessed. Ninety-five of the questions deal directly with the public availability and comprehensiveness of the eight key budget documents that governments should publish at various points of the budget cycle. Answers to these are given numerical value and aggregated to form the Open Budget Index. (The remaining 30 questions relate to opportunities for public participation in the budget process, and to the roles played by legislatures and supreme audit institutions in budget formulation and oversight.) As such, the OBI provides cross-country measure of public access to information on the receipt and expenditure of public funds.

Primary Purpose

The OBI provides cross-country measure of public access to information on the receipt and expenditure of public funds.

It is designed to provide a better understanding of the current state of budget transparency and accountability, as well as how these have changed over time.

The OBI is being used by governments, civil society groups, and international donor agencies to both measure progress toward more transparent and responsive budgeting and to encourage civil society and citizens to participate in budget processes. The OBS tool has been used by organizations in various countries to support their advocacy efforts and in particular to ask for improvements in budget transparency.

Some civil society and international development organizations producing governance indicators have included elements of the Open Budget Survey in their own indices and studies. Donor agencies, such as the European Commission and the U.K. Department for International Development, have drawn extensively on the Open Budget Survey recommendations in developing their new guidelines for providing direct budget support.

Construction

Annex p65
The OBI is constructed from the responses to the 95 Open Budget Survey’s questions on the amount of budget information that is made publicly available in eight key budget documents.

Most of the questions require researchers to choose from five responses. Responses “a” or “b” describe a situation or condition that represents good practice regarding the type of budget information (or budget practice) that the question assesses, with “a” indicating that the standard is fully met. Response “c” corresponds to minimal efforts to attain the relevant standard, while a “d” indicates that the standard is not met at all. An “e” response indicates that the standard is not applicable, for example when an OECD country is asked about the foreign aid it receives.

Some questions, however, only have three possible responses: “a” (standard met), “b” (standard not met), or “c” (not applicable).

Researchers are required to provide adequate evidence for each of their responses, and to supplement their answers with comments, clarifications, and links to relevant documentation.

Once completed, the questionnaire responses are quantified. For the questions with five response options, an “a” receives a numeric score of 100, a “b” receives 67, “c” answers get 33, and “d” responses are scored 0. Questions receiving an “e” are not counted as part of the country’s aggregated scores. For the questions with three response options, the scores are 100 for an “a” response, 0 for a “b,” and “c” responses are not included in the aggregated score.

Upon completion, IBP staff members analyzed and discussed each questionnaire with the individual researchers over a three- to six-month period. The IBP analysis sought to ensure that all questions were answered in a manner that was internally consistent for each country, as well as consistent across countries. The answers were also cross-checked against published budget documents and reports on fiscal transparency issued by international institutions, such as the IMF, World Bank, and the OECD. The questionnaires were then reviewed by two anonymous peer reviewers who have substantial working knowledge of the budget systems in the relevant country. The IBP also invited the governments of 95 of the countries surveyed to comment on the draft Survey results.

Though each of the eight key budget documents assessed may have a different number of questions related to it, the overall OBI score is a simple average of all 95 questions. In calculating the OBI scores, no method of explicit weighting was used to offset the disproportionate influence of documents that have more questions about them on the Survey.

**Scope of expenditures considered**

| N/A |

**Spheres of expenditure**

| Public |

**Stock / Flow**

| Annual/multiyear average |

**Presentation**

| OBI is a summary indicator: an average of the response scores against survey questions. Countries are ranked in regards to their score. The countries are also catagorised by their score: Extensive Information (OBI Scores 81-100), Significant (OBI Scores 61-80), Some (OBI Scores 41-60), Minimal (OBI Scores 21-40), Scant or No Information (OBI Scores 0-20) |

**Data acquisition process**

| In-country researchers complete the questionnaires through a combination of desk-based research and interviews with relevant officials, while being supervised by the IBP staff, over a clearly defined research period to guarantee comparability of data. Almost all of the researchers responsible for completing the Open Budget Questionnaire belong to either academic institutions or civil society organizations. The results were thoroughly vetted by two anonymous peer reviewers and by the IBP. |
The survey is repeated biannually.

**Strengths**

- Avoiding the de-jure emphasis of other PFM assessments, the majority of the OBI Survey questions assess what occurs in practice, rather than what is required by law.
- OBI allows for comparison across time for a large number of countries.
- OBI is based on objective assessments rather than perceptions, making it a useful diagnostic tool. Researchers must substantiate their responses to each question with existing documentation that is publicly accessible and cross-checked. In order to justify responses given, researchers must provide a clear citation of the page, chapter, or section of the appropriate budget document that includes the relevant information. Peer reviewers as well as IBP staff then check the researchers’ assessment using the same sources as the researchers.
- The research to inform the Survey is conducted by independent local experts. Where possible, to maintain continuity and understanding of the research process and interpretations, the same researchers (or institutions) as in previous rounds are employed.
- Having IBP staff as the ultimate referee for the answers ensures consistency in the interpretation of answers or the eligibility of documents taken into consideration. All comments are published so that in the questionnaire it is possible to follow the dialogue that took place between researchers, IBP referees, government, and peer reviewers.

**Limitations**

- OBI does not assess the quality or credibility of the information published in the budget (unlike the PEFA or the IMF fiscal ROSCs mentioned above). i.e., whether the information presented in government budget documents is an accurate reflection of state finances. This means that countries which score highly on the OBI may be presenting their publics with faulty information.
- Potential endogeneity: given that the OBI is used not only as a diagnostic but as an advocacy tool, substantial improvements in budget transparency may in fact reflect OBI researchers’ successful advocacy campaigns.
- Though using a simple average is clear, it implicitly considers certain budget documents as more important than others depending on how many questions are dedicated to them.
- The Survey is not an all-encompassing tool and, notwithstanding modifications made during each round, there will always be issues that are left out and that should be taken into account when assessing budget transparency in a given country. Nonetheless, the IBP holds that the OBS establishes the lowest possible benchmark that can be used to measure and compare countries on budget transparency.
- OBI does not have global coverage and the time-series data has limitations (four data points per country at the most)

**Sources:**


Seifert et al, 2013, The Open Budget Index (OBI) as a Comparative Statistical Tool

---

**ONTARIO MUNICIPAL BENCHMARKING INITIATIVE (OMBI)**

**Summary**

OMBI is a collaboration among 16 municipalities working together measure performance and share best practice regarding municipal service. It sets out to provide credible municipal data that measures and compares how efficiently and effectively services are delivered.

**Primary Purpose**

OMBI provides year over year internal measurement of progress and also comparisons with other municipalities

Stated reasons for measuring include:

- Sound business practice and expectation in today’s environment

Annex p67
• Overall acceptance by Government
• Creates a culture of continuous improvement
• Measurement is the ‘jump-off’ point leading to discussions about underlying practices or policies in achieving certain results
• Citizen expectations for more effective and efficient programs and services
• On-going fiscal challenges to do more with less
• Need to align service requirements with budget priorities
• Contributes to the public trust and confidence in municipal government with its progressive benchmarking framework and forums for the exchange of ideas.

Data collection is part of a broader benchmarking exercise intended to improve performance by way of the following process:

1. Select Programs for Benchmarking
2. Develop performance measures
3. Collect and analyze the data
4. Establish the ‘Zone of Performance’
5. Assess/Recommend best/better practices
6. Develop emulation strategies
7. Evaluate benchmarking process

### Construction

Framework of (‘more than 850’) individual indicators under the 37 service areas. The benchmarking framework includes four types of measures:

1. Community impact measures: The effect programs and services have on our communities
2. Service level measures: The number, type or level of service delivered to residents in municipalities
3. Customer Service Measures: Measure the quality of service and degree of customer satisfaction with the service received, relative to service standards and customer needs and expectations.
4. Efficiency measures: How municipalities use their resources, often expressed as a cost per unit of service or the volume of output per staff member.

The initiative is led by Chief Administrative Officers and City Managers who form the OMBI Board. Their role is to represent the interest of their respective municipality and determine practices and processes that drive the initiative. The Municipal Lead also serves as a conduit within their municipality to coordinate the annual OMBI Data Call, support internal experts and is responsible for facilitating various Expert Panels.

Expert Panels collect data across the municipal services, and each panel consists of staff with expertise in the particular service area. They meet throughout the year to develop their respective measure decks, refine technical definitions, and also to learn, network and exchange information. The Financial Advisory Panel (FAP), comprised of representatives from each municipality, ensures cost data is collected in a consistent manner.

In order to support the overall benchmarking model and the implementation of the performance measurement framework, OMBI has developed a number of practices and processes including:

• Data Dictionaries
• Data Warehouse
• Data Collection Protocols
• Indirect Costing Methodology
• Data Sharing and Public Reporting Protocol
• Common Measurements Tool, consisting of a question bank and a user manual providing an easy-to-use method for collecting feedback from customers.

### Scope of expenditures considered

37 Municipal services included:

**Direct Services:** Building, By-law, Child Care, Culture, EMS, Fire, Hostels, Libraries, Licensing, Long Term Care, Parking, Parks, POA (Court Services), Planning, Police, Public Health, Roads, Social Assistance, Social Housing, Sports & Recreation, Transit, Waste Management, Wastewater, Water

Annex p68
**Support services:** Accounts Payable, Clerks, Customer Service (new), Facilities, Fleet, General Revenue, Human Resources, Information Services, Investment, Legal Services, Payroll, Purchasing, Taxation

Municipal level services only considered.

<table>
<thead>
<tr>
<th>Spheres of expenditure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stock / Flow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stock</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Framework of indicators, no composite indicators identified.</td>
</tr>
</tbody>
</table>

Selected indicator results are presented in annual reports, in graphs are designed to show how participating municipalities compare with each other on selected service measures and against the median line.

Influencing factors, such as municipality population, geographic size, organizational form, government type, etc., as they relate to the results are listed and “should be considered when reviewing the results”.

<table>
<thead>
<tr>
<th>Data acquisition process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seemingly self-measurement by the municipalities. No further detail identified.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strengths</th>
</tr>
</thead>
<tbody>
<tr>
<td>According to OMBI (no independent evaluation identified)</td>
</tr>
</tbody>
</table>

- Expert panel meetings provide a forum for the exchange of ideas and posing of business questions facing the program
- External peer review to complement each municipality’s examination of their own internal performance
- Strengthens municipalities accountability and improves the level of transparency in how services are provided and reported
- Enables Councils and Staff to make informed decisions and/or set policy based on service quality, quantity and cost

<table>
<thead>
<tr>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognizing the uniqueness of each municipality, e.g. population, geographic size, organizational form, government type, etc., results are influenced to varying degrees by a number of factors.</td>
</tr>
</tbody>
</table>

- Datasets are not entirely comparable across municipalities because of differences in services delivered.
- Full datasets not available: only selected indicators are presented in the report
- Coverage is poor, with only 16 municipalities participating (out of 444 in Ontario alone).

<table>
<thead>
<tr>
<th>Sources:</th>
</tr>
</thead>
<tbody>
<tr>
<td>OMBI 2012a, Performance Management Report</td>
</tr>
<tr>
<td>OMBI 2011, Performance Management Report</td>
</tr>
<tr>
<td>OMBI, date unknown. Overview: Ontario Municipal Benchmarking Initiative</td>
</tr>
<tr>
<td>OMBI, 2012b, Developing a Survey Instrument for Local Governments: OMBI CMT</td>
</tr>
</tbody>
</table>

**BENCHMARKING GOVERNMENT PROVISION OF SOCIAL SAFETY NETS, World Bank**

**Summary**

In an effort to address the question of how much governments should spend on social safety nets, this paper presents an analysis of 63 countries spending patterns from 1972-1997 using a comparative benchmarking methodology, based three different rankings; a ranking based on

Annex p69
unconditional expenditures on safety nets; a ranking conditioning on structural features of the economy; and a ranking also conditioning on the quality of institutions within the country.

**Primary Purpose**

The intended use is for cross-country comparison, particularly between countries of similar income levels, to generate a picture as to how states are performing relative to international expenditure norms. As such, it may be useful to policymakers in determining the appropriate level of overall spending, with the end intention to augment social safety expenditure where necessary: “Telling policymakers how they perform relative to their economic neighbours may prove to be an effective means of providing the incentives to increase safety net expenditures by introducing a form of yardstick competition across economies.”

**Construction**

Countries are benchmarked through three distinct rankings:

**R1 Time-averaged unconditional ranking of safety net spending:**

A ranking of the share of gross domestic product (GDP) devoted to each type of safety net. According to R1, the country with the highest ranking is that country which spends the most on safety net expenditures as a share of its GDP. Safety net expenditures are expressed as a share of GDP. Two alternative normalizations are considered: (i) expenditures as a share of total government expenditures, and (ii) expenditures as a share of total government tax revenues, which capture the notion of a government budget constraint. On the whole the rankings to be largely unchanged by any of these normalizations, although certain oil-dependent economies perform better when we take expenditures as a share of total tax revenues.

There is enough variation in this data to suggest that policymakers cannot all be optimizing.

**R2 Ranking of social spending conditional on structural factors.**

To account for the variation under R1 through adjusting certain structural characteristics of the economy, such as the level and distribution of income, the level of urbanization, etc.

Structural factors included include: (i) the (log of) per capita income of the country which proxies for the government budget constraint; (ii) the fraction of the population of working age (between 15 and 64), which captures the level of dependency in a country—in most countries, children and the elderly are particularly susceptible to poverty; (iii) the fraction of the population residing in urban regions, which again can capture an element of neediness in the population, given that the poor most often reside in rural locations; (iv) a measure of shocks to income per capita, i.e., how far a country is from its long-run sustainable income level in any given year, because such cyclical components could also be correlated with expenditures. Having controlled for these structural features, there remains an unexplained component of the level of expenditures.

**R3 Ranking also taking into account quality of government.**

The third ranking takes into account the institutional framework within a country, which influences the levels of safety net spending and the effectiveness of a given level of expenditures. The set of institutional features taken into account are the (i) the level of repudiation of government contracts; (ii) the threat of expropriation of assets by government; (iii) the level of corruption in civil society; (iv) the effectiveness of the rule of law; and (v) the amount of government bureaucracy. In order to form a cross-country ranking taking into account these institutional features, they take the amount of unexplained expenditures from R2, and regress these on these various quality-of-government indicators. They rank these unexplained components in R3, which now takes account of structural and institutional factors. A country with...
a positive unexplained component is interpreted to be spending more than we would expect given the quality of its institutions and structural characteristics of the economy.

Given this construction, it is possible to decompose the policy advice to governments into two forms—first, the effects of changes in the structural characteristics of the economy, and second, the effects of changing institutional features.

The framework is disaggregated by the three rankings and also by the category of safety net (social security and welfare, transfers to organizations and households, health and education).

**Scope of expenditures considered**

All expenditures on safety nets, where safety nets are defined as those public interventions which are designed to serve two key functions: (i) to play a redistributive role transferring resources toward the poorer members of society to bring them out of poverty, and (ii) to provide greater opportunities for individuals to mitigate risks from unforeseen contingencies.

**Spheres of expenditure**

Public

**Stock / Flow**

Flow: cross-country panel data set over the period 1972–97 with data averaged over five-year periods

**Presentation**

Framework of three ranking scores

**Data acquisition process**

By authors

**Strengths**

- The authors argue that benchmarks of this type are useful as they can generate incentives to improve performance to conform with international and regional norms. Pressure to do this can come both from the international community and from the domestic dialogue which is generated by the publication of such information. This has often been the case in such areas as labor and environmental standards, and it seem sensible that this process of “yardstick competition” be extended to cover the case of social protection.
- The approach is a deliberate move away from the “optimizing approach” that seeks to answer how much governments ought to spend, as and such avoids having to specify any welfare criterion which would require unfeasible amounts of data. It posits that what countries are actually observed to be doing on average is a good indication of feasibility constraints and provides a natural focal point for what countries ought to be able to do.
- It avoids many of the disadvantages of incidence studies, including the fact that such studies take the benefits that accrue to an individual to be well proxied by the average cost of provision to that individual can also be called into question. Benefit incidence studies also implicitly assume that there is a uniform cost of service provision over all households, or that the public good is homogeneous. To the extent that this is not the case, they may lead to incorrect inferences regarding the distribution of benefits of social expenditures.

**Limitations**

- The value of this benchmarking exercise is contingent on data being reflective of government activities in the safety net area and being comparable across countries. On both these counts, available data sets are somewhat lacking.
• The currently available cross-country data series do not allow authors to control for all the structural features that we believe would drive safety net spending. This may lead to concerns about potential sources of bias affecting the estimates. However the use of rankings, rather than a literal interpretation of the parameter estimates mitigates this to a degree.

Sources:
Timothy Besley, Robin Burgess, and Imran Rasul, 2003; Benchmarking Government Provision of Social Safety Nets

CHILD BUDGETING, Save the Children

Summary

Child centred budget analysis is a tool for increasing good governance in support of the implementation of the UN Convention on the Rights of the Child.

It recognises that Children’s issues most often receive less attention and are given low priority in the state budget, and even when adequate money is allocated to a children’s programme, its effectiveness is limited as it ends up delivering less than what it was supposed to. Analysing budgets gives them the means to create effective advocacy strategies for seeking more commitment from the state, holding it accountable, and ensuring more effectiveness and transparency in the expenditures.

For this purpose, Save the children developed a toolkit training CSOs on child budget analysis. It entails building the knowledge and capacities of civil society organisations, including children’s organisations, to understand the national budgetary process and undertake State budget analysis.

Primary Purpose

In-country advocacy through diagnosis of deficiencies: this toolkit is designed to train individuals and civil society on how to conduct a child centered budget analysis to support their advocacy work and to hold state accountable for the fulfilment of child rights.

Construction

The Save the Children toolkit identifies six steps of the child budget analysis methodology:

1. Assemble a team: Given the time-consuming and detailed nature of the work, this team needs to be dedicated and have the required skills and analytical abilities. An advisory committee of experts, such as retired or/serving government officials and an economist, along with child rights practitioners, is also recommended.
2. Collect budget documentation, which corresponds to the whole budget cycle from budget proposals, to laws, to draft and approved budgets, mid and end of year reports, audit reports etc.
4. From the budget documents of each ministry and department carefully mark out the line items ones that are focussed on children. See below of inclusions and exclusions; two key groups are considered:
   o Schemes and programmes and sectors that are specifically aimed at children (0-18 years).
   o Schemes and programmes that fit within the sectoral categories identified above but do not specify the exact age group aimed at. In effect, this means that the financial allocation for programmes related to mother-and-child health, the rehabilitation of girls and women in distress, and those for the development of sports and youth are a part of the BFC analysis.
5. Once the programmes have been identified they have to be clustered into the sectoral divisions under which they fall
6. Data entry: to study how much was allocated (in planned and revised budgets), how much was spent and what is the trend. Once entered, totals across sectors and overall are calculated. Subsequent calculations include share of the child budget in the total budget as well as the shares of the different sectors, share of the child budget in the social sector.
7. Data analysis and presentation.

Annex p72
As part of the analysis, research should try to ascertain whether government is meeting its budget input obligations. The ‘how much questions’ (How much is allocated to the programmes? What is the rate of the real increases in these allocations? Are programmes aimed at realising the right being phased out without substitutes being put in place?) and the ‘allocative efficiency’ questions (How much geographical inequity is there in the allocations? Do they discriminate against any children? Do they target the poorest of the poor? Are they prioritising basic services? Etc). Researchers should also seek to answer whether government is meeting its budget output obligations (Is output increasing, costs falling and quality improving? Is there discrimination in access to services? What steps is government taking to remove remaining access and discrimination hurdles?)

**Scope of expenditures considered**

**Defining expenditures:**

The Save the children methodology considers public expenditures, focusing exclusively on targeted allocations made specifically and deliberately to improve children’s lives, either directly as in setting up a school or an immunisation centre, or indirectly as in free health checkup for pregnant women.

It is noted that the CRC lists the four basic rights of children as the rights to survival, development, protection and participation. However experience from India suggested that this typology was difficult to follow as expenditures are distributed as per sectoral ministries rather than as rights groups; given that the purpose of the BfC exercise is advocacy and the targets are the ministries and departments it is critical that the study is organised in a manner that our policy makers can identify with and understand.

Given this, the following sectors are to be considered for analysis:

- **Education:** Includes elementary and secondary education programmes
- **Development:** Includes programmes and schemes for early childhood care and education; and other general schemes and programmes that do not fit within any particular sector category, but are aimed at the overall development of children
- **Health:** Includes programmes and schemes related to the health care needs of children
- **Protection:** Includes interventions that are aimed at children at work, children in need of adoption, children in prostitution, children who are physically or mentally challenged, street children, children who are neglected or treated as juvenile offenders

Can be extended to sub-national Government, but it notes, the divisions of programmes and schemes at the provincial levels may not be the same at the national level, and so the identification of relevant expenditures will need to be done on a case by case basis.

**Spheres of expenditure**

**Public**

**Stock / Flow**

Flow: a timeframe of a minimum of three years is recommended in the Save toolkit, to allow for trend analysis.

**Presentation**

Under the Save methodology, the endpoint of the analysis is a final total expenditure against total plan, which can be disaggregated by sector. A more detailed budget analysis could go deeper into the issues of how the budget is spent, disaggregating its impact on various groups of children based on gender, age, ability, background, and so on.

**Data acquisition process**

Save toolkit is to be used by CSOs and NGOs. Note: key in team spec is not specialist budget skills but availability to commit to a significant time investment "Don’t insist on skilled staff: It is common for groups to have some difficulty in finding the right people to work on budget issues. There is a feeling of budget work being elitist and needing persons with specific skill sets. However, it is important to remember that interested individuals with strong analytical and communication skills would do well. More important is dedication and a fulltime engagement in budget work because developing budget expertise would not be difficult if one is fully immersed in budget issues.”

Annex p73
Strengths

- In countries that have experimented with budget analysis for children, some concrete results were seen: In South Africa, the combination of CBU-IDASA’s solid information and the mobilising power of Alliance for Children’s Entitlement to Social Security (ACCESS), led to the government increasing the total resources for children in the 2003/04 budget in real terms, as also the age until which a child could benefit from the Child Support Grant up to 14. In India, the government not only accepted and acknowledged the concept and methodology, but also used it itself. There are many more examples in the education sector particularly where budget analysis has been used as an effective advocacy tool.
- Straightforward, does not require specialist expertise, rather designed for CSOs who may not have PFM skills.

Limitations

- Excludes spending directed towards the larger society or family unit as a whole, even when that has a clear impact on children, be it a road, a hospital or an airport. It noted that “it would be excessively complicated, and even impossible, to disaggregate the amount going to children from such programmes.”
- Defining a child. This is particularly difficult in countries such as India where different laws define children differently. In Bangladesh, for example, the child policy defines a child as persons up to 16 years. Yet, the country is a signatory to the UNCRC which defines children as all persons up to the age of 18 years.
- Accessing government data remains difficult and unpacking budget data can prove tedious and confusing.
- The selection of programmes and schemes for children has an element of subjectivity.
- Although attention is paid to both input and output obligations of Government, there is little linkage between them – that is, questions of effectiveness and efficiency are largely overlooked.

Sources:

Save the Children, 2010, Budget for Children Analysis

Annex p74
CHILD BUDGETING, African Child Policy Forum

Summary

The primary purpose of budget analysis from a child rights perspective is to examine if governments are utilising the maximum of their available resources to meet their commitments to children. The approach provides a framework for assessing budget expenditure in terms of sufficiency, trends over the years, equity, and efficiency in giving children a good start in life and in progressively realising their wellbeing. From this analysis a composite performance index is constructed to compare and categorise the performance of African countries in terms of their efforts to allocate the maximum amount of available financial resources in sectors that benefit children.

Primary Purpose

The analysis primarily focuses on in country diagnosis of deficiencies to support advocacy efforts, but the index can be used for cross-country comparison.

Construction

Budget analysis from these perspectives looks into seven basic issues:

1. Availability of resources: tax and non-tax revenues, grants and other funds from external sources are used as a proxy for resource envelope. Attempt was also made to look at the economic outlook to examine the prospect for future revenue generation. In the African context, population growth is another important consideration.

2. Adequacy of budgets for children: Assessment of sufficiency of budgets for children is usually carried out by comparing the amount spent in each of the sectors targeting children against international or regional targets related to the sectors. For instance, health spending is compared with the Abuja commitment. In this case, the budget allocations are evaluated against child wellbeing outcomes and per capita cost required to achieve desired outcomes.

3. Priority given to sectors/programmes benefiting children: By comparing the proportions of the national budgets that go to programmes or sectors benefiting children with the other sectors, one can easily see where children stand in governments’ resource prioritisation. Such analyses, among others, help to make inference on whether governments are keeping their promises and committed to children.

4. Progressive increase in allocation: Not all rights of children can be realised in a given period of time and, therefore, governments are obliged to realise them progressively over time. Budgets for children need to reflect this intention to be considered as ‘progressive’. Monitoring the progressive realisation of child rights through budget analysis entails a multi-year comparison of budgets for children and examining the trend over time.

5. Equity: budgets need to be reviewed to examine if resources are allocated fairly to all children irrespective of their gender and socio-economic background. The challenges with such kinds of analyses relate to their requirement for detail data on each programme and projects that target these groups of children. Most budget experts in Africa do not compile budget data by programmes or projects. Even if they do, the information is not made available for public use. This situation makes it difficult to carry out the analysis, and alternatively, attempt is made to use and compare outcome related proxy indicators to assess budgets’ sensitivity to the needs of vulnerable children and make inference about equity.

6. Efficiency and Effectiveness of the budgets in terms of bringing about concrete child wellbeing outcomes: One of the problems in assessing budget efficiency and effectiveness relates to the intensity of data and information requirements. These parameters demand detail administrative, systemic, process and capacity related data and information. In the absence of data and information on budget monitoring and evaluation, attempt is made to assess budget efficiency and effectiveness by comparing budget expenditures against objective child wellbeing outcomes to determine, at least in a crude form, whether budgets for children are well targeted and efficiently utilised to result in concrete outcomes reflected on children themselves.

Annex p75
A composite measure of performance – the Performance Index for Budgeting for Children - is constructed using common indicators on budget expenditures on sectors and programmes targeting children. The indicators used are as follows:

- Expenditure on health as a percentage of total government expenditure
- Total public expenditure on education as a percentage of GDP
- Percentage of the budget for routine EPI vaccines financed by government
- Military expenditure as a percentage of GDP
- Percentage change in governments’ expenditure on health, 2004-2008. These indicators are first converted into standardised performance score values, then aggregated to yield the Performance Index for Budgeting for Children.

Serving as a policy instrument, the index is used to compare the performance of African governments on their commitment to maximising the use of available financial resources in sectors that benefit children. In order to show the progress made in performance in budgeting for children since 2004, the scores values for the period 2007-2008 were compared with the corresponding performance scores for the period 2004-2005.

Scope of expenditures considered

Four operational categories of budgets for children have been identified for the analysis:

1. budget for social and child protection budgets under this category include national programmes and initiatives targeted to reduce poverty and vulnerability, as well as budgets allocated to support and protect children against harm and exploitation. These include projects specifically aimed at supporting orphans, children with disabilities, street children, child protection units, juvenile justice systems and related interventions
2. budget for health care: All funds allocated to programmes and initiatives targeted to provide health related services for mothers and children are included in this category
3. budget for the education of children: This category includes all public funds allocated for expenses related to primary and secondary level education
4. budget for child development Child development involves multi-sectoral interventions that include early childhood care, nutrition supplementation, pre-school and recreation services. It also covers direct financial and material support provided for children with disabilities, and other child-related programmes that do not fit in the above three categories.

The aforementioned budget categories are not the only ones that benefit children and ACPF recognises that budgets for other programmes that may seem unrelated to children could have a significant impact on their wellbeing. However, it is difficult to determine and quantify how much of the indirectly related budgets (such as budgets for food security and infrastructure development) go to children. In the African context, such disaggregated budget information is either non-existent, or publicly unavailable.

It is asserted that the four categories of the budget for children represent the bulk of the total national budget that is targeted mainly to provide for children and the analysis is based on these categories.

Spheres of expenditure

Public

Stock / Flow

Differs by indicator

Presentation

Ranking and categorisation according to whether country allocates the maximum/fair amount/the minimum of available resources for children

Data acquisition process

Not detailed, but data for the index appears to have been sourced from existing sources.

Strengths
• Attempts to address efficiency and effectiveness (whilst acknowledging limitations to this due to lack of information)

Limitations

• Budget lines that seem unrelated to children may have impact on children’s wellbeing. Good examples in this regard are investments in ensuring food security that have direct impact on nutritional status of children, and budgets related to infrastructure and community development. Investments in these areas, although not directly targeted to children, benefit them through improvements in family income and wellbeing, as well as in facilitating their access to basic services. However, it is methodologically difficult to disentangle the stream of benefits and identify those that can be attributed to child wellbeing, much less quantify them. As such the analysis pragmatically focused only on the four areas of investment: child development, health, education and social protection.

• It may not be possible to evaluate how effectively or efficiently budgets are being spent: Assessing budget efficiency and effectiveness, requires an in-depth look into the systemic, management, capacity, corruption and even attitude-related factors that cause inefficiency in the utilisation of budgets, which demands extensive data and information on administrative and process-oriented aspects. There is, however, a dearth of such detailed data and information on efficiency and effectiveness of budgets in most parts of Africa, and it is difficult to analyse budgets for children along these parameters. Alternatively, budget expenditures can be evaluated for efficiency and effectiveness in terms of objective child wellbeing outcomes achieved and services delivered. However, most budgets in Africa are not results-based, and present little information on what they objectively expect to achieve. Despite the limitations, however, an attempt is made in this report to assess budget efficiency by comparing budget expenditures against child wellbeing outcomes, and through a review of studies carried out on the subject.

• Data deficiencies complicate all facets of the budget cycle from budget preparation to execution and control and undermines budget analysis.

• It is not possible to evaluate whether the resources allocated are reaching the intended targets simply by analysing budget expenditures.

• budget analysis needs to be supplemented by detailed contextual information on the economy, population, governance, level of decentralisation and other systemic issues. Without such complementary information, budget analysis may not by itself show the full picture and serve its purpose of highlighting gaps in the budget cycle.

Sources:


World Education Indicators

Summary

Framework of indicators that compare national education systems across five thematic areas:

1. The outputs of education systems: Graduation from upper secondary and tertiary education
2. Sources and flows of education expenditure
3. Levels and uses of education expenditure
4. Access to education, participation and progression
5. Teachers and the learning environment

Participating countries include: Argentina, Brazil, Chile, China, Egypt, India, Indonesia, Jamaica, Jordan, Malaysia, Paraguay, Peru, Philippines, the Russian Federation, Sri Lanka, Thailand, Tunisia and Uruguay.

Primary Purpose

The WEI programme helps to assess progress and challenges in national education systems from an international perspective by facilitating international comparison (In particular, comparison of 19 countries participating in the World Education Indicators (WEI) programme against OECD averages). This analysis is

Annex p77
intended to provide a useful benchmarks for development, especially when national conditions, such as population size, are taken into account to provide context. The data also highlights where national policies have achieved positive results.

**Construction**

For sources and flows of education expenditure:

- total education expenditure over GDP
- total public expenditure over total private expenditure
- public expenditure on education as a percentage of total public spending
- also have an indicator on the flow of funds (share: Indirect public transfers and payments to the private sector/Direct public expenditure on private institutions/Direct public expenditure on public institutions

for levels and uses of education expenditure:

- Educational expenditure (public and private costs) per student
- Educational expenditure per student / GDP per capita
- Use of funds by nature of spending: breakdown between capital and current expenditure on educational institutions.

PPP is employed to enable international comparison.

**Scope of expenditures considered**

All types of students and all age groups are meant to be included in the data: children (including those classified as exceptional), adults, nationals, foreigners, as well as students in open distance learning, special education programmes and educational programmes organised by ministries other than the Ministry of Education, provided that the main goal of the programme is the educational development of the individual. However, vocational and technical training in the workplace, with the exception of combined school- and work-based programmes which are explicitly deemed to be part of the education system, are excluded from the education expenditure and enrolment data. Educational activities classified as “adult” or “non-regular” are covered, provided that the activities involve studies or have subject matter content similar to “regular” education studies or that the underlying programmes lead to qualifications similar to those gained through corresponding regular educational programmes. Courses for adults that are primarily for general interest, personal enrichment, leisure or recreation are excluded.

**Spheres of expenditure**

coverage extends, in principle, to the entire national education system regardless of the ownership or sponsorship of the institutions concerned and regardless of education delivery mechanisms. However only 8 out of 19 WEI countries can provide sufficient data of this type. This indicator excludes private expenditure outside schools and universities and, therefore, understates private investment in education.

**Stock / Flow**

Annual: 2005

**Presentation**

Framework of indicators for the five different thematic areas

Can be disaggregated by level of education. Education expenditure can be disaggregated by public and private (where data exists)

**Data acquisition process**

The UNESCO Institute for Statistics (UIS) supports and maintains the secretariat for the WEI programme. The OECD provides key indicators for its Member States for inclusion in Education Counts, in addition to reporting similar data in parallel in the OECD Education at a Glance series. Data on graduates, personnel, entrants, enrolment and education finance are based on the annual UNESCO-UIS/OECD/Eurostat (UOE) data collection on education statistics. Data on educational attainment, teacher salaries and curricula are derived from the UOE questionnaires designed specifically for WEI countries.

**Strengths**

Annex p78
Limitations

Incomplete and inconsistent coverage of education expenditure data remains problematic. While many countries maintain relatively complete data on public expenditure, they do not collect information on expenditure by households, private schools or other private entities (e.g. foundations, enterprises, religious groups and labour unions). Currently, fewer than one-half of WEI countries provide data on private expenditure on educational institutions, and only six can do so by level of education. This information gap means that reported expenditure on education can be greatly understated in countries where the private education sector is sizeable or tuition fees are common.

Incomplete data coverage of education expenditure also persists in a number of areas in the public sector. For example, expenditure on pre-primary or adult education is not always included when such programmes are run by ministries other than education, e.g. family welfare or labour departments.

The paper recognises influence of contextual factors (e.g. coverage affects spending, and spending per capital to get the last 10% is more costly; efficiency of education system) but it doesn’t include these as interpretation aids.

Sources:

UNESCO, 2007 Education Counts: Benchmarking Progress in 19 WEI Countries

---

### OECD Social Expenditure Database (SOCX)

**Summary**

When the OECD Social Expenditure database (SOCX) was set up in the early 1990s, it was designed to capture the ‘social expenditure programme’ for all 34 OECD countries at the programme level. SOCX groups social spending by the nature of provision into public, mandatory private and voluntary private social expenditure, across nine different social policy areas.

**Primary Purpose**

Cross country comparison and in country diagnosis of deficiencies. The detail in SOCX allows for in-depth study of national and cross-national social protection policy.

**Construction**

The OECD Social Expenditure Database groups benefits with a social purpose in nine policy areas - Old-age, Survivors, Incapacity-related benefits, Health, Family, Active labour market policies, Unemployment, Housing, and Other social policy areas. What is included in each of these is detailed below.

In order to facilitate international comparisons this information is related to

- Gross Domestic Product (GDP)
- Deflator for GDP
- Gross Domestic Product at 2000 prices (GDPV)
- Gross National Income (GNI)
- Net National Income (NNI)
- Total General Government expenditure (GOV)
- Purchase Power Parities (PPP)
- Exchange rate (EXC)
- Population (POP)

**Scope of expenditures considered**

---

Annex p79
The OECD defines social expenditures as: “The provision by **public and private institutions** of benefits to, and financial contributions targeted at, households and individuals in order to provide support during circumstances which adversely affect their welfare, provided that the provision of the benefits and financial contributions constitutes neither a direct payment for a particular good or service nor an individual contract or transfer.”

Social benefits include **cash benefits** (e.g., pensions, income support during maternity leave and social assistance payments), **social services** (e.g., childcare, care for the elderly and disabled) and **tax breaks with a social purpose** (e.g., tax expenditures towards families with children, or favourable tax treatment of contributions to private health plans).

There are two main criteria which have to be simultaneously satisfied for an expenditure item to be classified as social. **First, the benefits have to be intended to address one or more social purposes.** Second, programmes regulating the provision of benefits have to involve either a) **inter-personal redistribution**, or b) compulsory participation.

**Demarcation of the policy areas**: Categories directly Relevant for CP

- Old-age – pensions early retirement pensions, home-help and residential services for the elderly;
- Survivors – pensions and funeral payments; **Allowances and supplements for dependent children of the recipient of a survivors’ benefit** are also recorded here.
- Incapacity-related benefits – care services, disability benefits, benefits accruing from occupational injury and accident legislation, employee sickness payments; **This excludes paid leave related to sickness or injury of a dependent child which is recorded under family cash benefits**
- Health – spending on in- and out-patient care, medical goods, prevention;
- Family – **This expenditure is often related to the costs associated with raising children or with the support of other dependants.** INCLUDES child allowances and credits, childcare support, income support during leave, sole parent payments;
- Active labour market policies – contains all social expenditure (other than education) which is aimed at the improvement of the beneficiaries’ prospect of finding gainful employment or to otherwise increase their earnings capacity. Includes employment services, training, employment incentives, integration of the disabled, direct job creation, and start-up incentives;
- Unemployment – includes all cash expenditure to people compensating for unemployment. Includes unemployment compensation, early retirement for labour market reasons;
- Housing – housing allowances and rent subsidies;
- Other social policy areas – includes social expenditure (both in cash and in kind) for those people who for various reasons fall outside the scope of the relevant programme covering a particular contingency, or if this other benefit is insufficient to meet their needs. non-categorical cash benefits to low-income households, other social services; i.e., support programmes such as food subsidies, are included, as is social expenditure related to immigrants/refugees and indigenous people are separately recorded in this category.

<table>
<thead>
<tr>
<th><strong>Spheres of expenditure</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Public and private</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Stock / Flow</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual.</td>
</tr>
</tbody>
</table>

To account for difference in FY: Adopting the same convention as for national accounts, year “n” is taken to mean the year in which a financial year begins, whether it starts on 1 January, 1 April, 1 July or 1 October. In cases where the financial year for social expenditure does not coincide with the calendar year, the relevant periods have been taken on a prorata temporis basis when using GDP (available for calendar years) and the GDP deflator.
Not an indicator, rather a database.

In order to facilitate international comparisons this information is related to GDP and other indicators, see above.

**Data acquisition process**

For all OECD countries data on public expenditure on health and public expenditure on active labour market policies (ALMPs) are taken from the OECD Health Data and the OECD database on Labour Market Programmes, respectively. Data on education of 3, 4 and 5 year olds in the OECD Education database feeds into the series on social spending on early care and education services. Data on unemployment compensation are taken from the LMP database for OECD countries that do not belong to the EU and from ESSPROS for EU countries.

For 10 non-European OECD countries, data delivered through the services of the delegates to the Working party on Social Policy of the Employment, Labour and Social Affairs committee responding to the SOCX Questionnaire.

For 24 European countries (EU-21, Iceland, Norway and Switzerland), data on social expenditure is provided by EUROSTAT as based on the information in their ESSPROS database.

**Strengths**

The detailed nature of expenditure data in SOCX constitutes an important form of quality control as the high level of transparency associated with detailed recording limits the scope for inappropriate recording (including double counting) of spending items in SOCX.

**Limitations**

There remain weaknesses in spending data, not least because local governments often play a key role in financing childcare services. In some federal countries, it is much more difficult to get a good view of public support for childcare across a country. This is because local governments may use different funding streams to finance childcare services, e.g., non-earmarked general block-grants, as in Canada, or because information on spending by local governments on childcare is not reported to national authorities, e.g., Switzerland. These issues are not restricted to federal countries. In the Netherlands, municipalities can provide childcare support for (groups) of their inhabitants, and they may finance this out of the general block-grant to municipalities. They can also use the central government funding stream to municipalities to support labour market integration for income support recipients, to finance, for example, childcare support for social assistance clients.

**Sources:**

OECD Social Expenditure Database (SO CX),

Child Well-Being Indicators - OECD

Summary
Child well-being indicators measure the quality of children’s lives. These data compares 21 policy-focused measures of child well-being in six areas, chosen to cover the major aspects of children’s lives: material well-being; housing and environment; education; health and safety; risk behaviours; and quality of school life. Each dimension is a composite of several indicators, which in turn have been selected in part because they are relatively amenable to policy choices.

Primary Purpose
Cross country comparison: There are two main reasons to identify differences in country performance across these child well-being dimensions. First, it shows the dimensions of child well-being where countries are comparatively successful or unsuccessful. It highlights where significant improvement in child well-being may be possible and so provides countries with information that can help in developing child policy priorities. Second, it allows comparative leaders and laggards to be identified. The question of how leaders arise, and why laggards fall behind can then begin to be addressed, and examples of best country practices can be drawn for future policy changes.

Construction
Six dimensions of child well-being have been identified here to cover the major aspects of children’s lives: material well-being, housing and the environment, education, health, risk behaviours, and quality of school life. Each dimension has roots in the international standards agreed for children in the United Nations Convention on the Rights of the Child (United Nations, 1989).

Across the six dimensions, 21 indicators of child well-being have been selected. A number of ideal selection requirements were borne in mind in choosing indicators.

- The child is taken as the desirable unit of analysis, rather than the family. A child-centered approach is now the norm in studies of child poverty and child well-being. Indicators should be as up-to-date as possible. Indicators cannot reliably inform comparative policy unless they paint a picture of child well-being reasonably close to the here-and-now.
- Indicators should be taken from standardised data collections which collect comparable cross-country information. If data is not reasonably comparable, it will fail to meet one of the most basic needs of a cross-country, data-driven study.
- Indicators should cover all children from birth to 17 years inclusive. The United Nations definition of a child as a person under age 18 is used here. Given evidence about the importance of the intrauterine environment for the child’s future health and development and the fact that in most countries a foetus legally becomes a child in utero, it may also be desirable to extend the definition of childhood to the period before birth. Indicators need a policy focus.
- As child well-being measures in this chapter are policy-focused, indicators with a relatively short causal chain from government action to improvements in well-being are favoured over indicators for which relationships between policy actions and outcomes were more speculative and the causal chain was longer.
- Indicators should cover as many OECD member countries as possible.

Scope of expenditures considered
NA

Spheres of expenditure
NA

Stock / Flow
Most annual, some multi-year average.

Presentation
No summary indicator is provided, or overall country ranking for child well-being is presented. The well-being indicators are presented in an index by dimensions, but not aggregated into a single over-arching child well-being index. No over-arching index is presented due in part to the limitations in the coverage of

Annex p82
available data. In addition there is little theory to guide which aggregation method to use. Given a lack of good theory and data, it was considered that creating an over-arching index would distract the focus towards discussion of the aggregation method, and away from more important practical issues of improving child well-being.

The data is reported by country and, where possible, disaggregated by sex, age and migrant status.

An interactive dashboard allows users to:
- Select an indicator and have results displayed in GIS format on a map
- Organise data on scatter plot diagram, by selecting indicators for x and y axis to show the relationship between indicators across the different countries
- Rank countries according to selected indicator
- Generate a range of charts including histograms and parallel coordinates, to show where countries fall on the distribution in relation to others.

**Data acquisition process**

All indicators presented in the framework are already publically available.

There has been no attempt to collect new data.

**Strengths**

Data presentation and Visualisation

**Limitations**

Despite a desire to cover all the OECD countries, there was incomplete coverage for the majority of indicators.

Complete country coverage was possible for eight of the 21 indicators.

**Sources:**

OECD 2009 Comparative Child Well-being across the OECD

http://stats.oecd.org/childwellbeing/
Annex 3: Discussion of options from option proposal

Options for construction of the benchmark
Annex 3.1. The mid-assignment report set out the various options for the construction of architecture of a benchmark. In this annex we reflect the discussion of options provided in the Options Report as background to the specific content proposals provided in the main body of this report.

Standardised benchmark content for assessment
Annex 3.2. The standardised content proposed in below used for the benchmark construction options assessment. The ‘content’ provided here should not be seen as a considered proposal on content against each option. Its purpose was to illustrate how the same content could be accounted for in different construction options.

Annex 3 Table 1  Standardised benchmark content for option assessment

<table>
<thead>
<tr>
<th>Content purpose</th>
<th>Illustrative datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assess quantum of child protection expenditure</td>
<td>Quantum of qualifying expenditure as defined</td>
</tr>
<tr>
<td>Assess expenditure distribution</td>
<td>Disaggregation by main economic classification</td>
</tr>
<tr>
<td>(indicative of expenditure efficiency, quality,</td>
<td>Disaggregation by service component</td>
</tr>
<tr>
<td>effectiveness, equity)</td>
<td>Disaggregation by type of risk/harm</td>
</tr>
<tr>
<td></td>
<td>Disaggregation by prevention and response</td>
</tr>
<tr>
<td></td>
<td>Disaggregation by funder (internal and external financing)</td>
</tr>
<tr>
<td>Assess equity of expenditure</td>
<td>Disaggregation by income group, gender and/or geographic location</td>
</tr>
<tr>
<td>Assess expenditure relative to need</td>
<td>Expenditure relative to a socio economic data series, for illustrative purposes number or proportion of children in poverty</td>
</tr>
<tr>
<td>Assess expenditure relative to country financial</td>
<td>Expenditure relative to GDP or GDP per capita</td>
</tr>
<tr>
<td>capacity</td>
<td></td>
</tr>
<tr>
<td>Assess expenditure relative to country status</td>
<td>Expenditure relative to other countries with similar income status, emergency and conflict affected country status,</td>
</tr>
<tr>
<td>Assess priority given to child protection</td>
<td>Government expenditure relative to total general government expenditure</td>
</tr>
<tr>
<td>Assess quality of expenditure governance</td>
<td>PEFA measures</td>
</tr>
</tbody>
</table>

Annex 3.3. The illustrative use of this content by benchmark construction option below assumes that in application all benchmarks will be presented in two parts:

- The core benchmark – this is the core indicator, or framework of indicators that comprise the actual benchmark. When users of the benchmark refer to a country’s performance against the benchmark, this is the score(s) they would primarily use.
- The benchmark presentation – this is a standardized presentation of the benchmark and allows for interpretive information that is not accounted for in the construction itself, to be presented. This can be either in standardized quantitative form (for example the World Bank Social Protection benchmark), or in standardized narrative form, e.g. the PEFA assessment reports.
Annex 3.4. In addition, over time, UNICEF could think of building up a database of assessments, that would present country benchmark scores comparatively on a dashboard, allowing various sorting and analysis options.

**Scoring benchmark construction options**

**Agreed criteria and approach to scoring**

Annex 3.5. Annex 3 Table 3 below sets out criteria for scoring the construction options, using a simple scoring methodology of 0 to 3, where 3 implies the highest score. In order to make the assessment more transparent, we developed identifying descriptions against each score for each criterion.

Annex 3.6. The criteria proposed in the mid-assignment report were applicable against all the key questions in the design framework. Some of the criteria however, are less relevant to the construction options and have been omitted in this assessment. Omitted criteria are:

- **Objective of the benchmark**: the expenditure selection methodology proposed in the previous sections fulfilled the stated UNICEF objectives, insofar as it allows the exogenous definition of the scope of the benchmark, and is state-centric. The further objective of cross-country comparability is assessed in various dimensions through the remaining criteria. We therefore do not assess the options against this overall objective.

- **In country monitoring and diagnosis, and ownership of the benchmark**: given that (i) the UNICEF reference group has elected cross-country comparability as the priority objective (rather than in-country monitoring and diagnosis); and (ii) that the scope of services to be included in the benchmark and other content is defined in a standard way across the construction options for the assessment with the result that all options will carry the same level of country relevance and ownership, we do not assess these criteria.

Annex 3.7. We have however refined the criteria given clarity provided in response to the mid-assignment report on the objectives of the benchmark. The criteria, their definition, and identifying descriptions are provided in the Annex 3 Table 3. In reading the assessment of options, it is important to remember that the assessment is of the marginal difference that the construction of the core benchmark makes to performance of the benchmark (thus constructed) against the criteria relative to other options. It does not assess the benchmark as a whole. The assessment also does not, except for the “ease of presentation” criterion, take into account what will be in the presentation. For example, the practicality and cost of data collection on expenditure itself is the same across options, but different constructions would have different implications for the cost and practicality of data collection and analysis, leaving out the presentation.

Annex 3.8. The scoring approach was used to identify possible marginal differences between the suitability of different options for the UNICEF Child Protection Financial Benchmark. The approach however is not intended to mechanically select the approach to be used. Rather it is an aid to generate a proposal, which can then focus discussion with clear, agreed criteria.
Annex 3.9. The criteria are presented into two sets: one set assesses the potential of options in terms of the potential and quality of comparison between countries and depth of information, the other set assesses the ease of presentation, understanding and application. In order not to weight the set over the other, the average score of indicators within each set is calculated, and the two averages summed.

**Annex 3 Table 2  Assessment criteria sets**

<table>
<thead>
<tr>
<th>Set</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparability and depth of information</td>
<td>Multi-country comparability</td>
</tr>
<tr>
<td></td>
<td>Multi-context interpretability</td>
</tr>
<tr>
<td></td>
<td>Depth of understanding relative to child protection systems</td>
</tr>
<tr>
<td>Ease of presentation, understanding</td>
<td>Ease of understanding and interpretation of differences in scores</td>
</tr>
<tr>
<td>and application</td>
<td>Simplicity of methodology</td>
</tr>
<tr>
<td></td>
<td>Relative cost and practicality, all countries equal</td>
</tr>
<tr>
<td></td>
<td>Cost and practicality of application relative to country context</td>
</tr>
</tbody>
</table>

Annex 3.10. The UNICEF reference group may want to weight the one set higher than the other: for example, given that the collection of expenditure data already will require some specialist input, 'simplicity' indicators overall could be ranked lower as the options overall may not add that much cost depending on for example the exact number of availability of data sets used. Also, individual criteria within the sets may be weighted differently.
### Annex 3 Table 3  Construction option assessment criteria and scores

<table>
<thead>
<tr>
<th>Definition</th>
<th>Score of 0</th>
<th>Score of 1</th>
<th>Score of 2</th>
<th>Score of 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-country comparability</td>
<td>The option has no potential to allow easy comparison between countries</td>
<td>The option has the potential to allow more than two countries to be compared at a time, but with limited understanding of their relative performance.</td>
<td>The option has the potential to allow comparative analysis or presentation of more than two country’s performance towards satisfactory understanding of relative performance.</td>
<td>This option has potential to allow countries to be ranked &amp; therefore allow an immediate &amp; clear understanding of their relative performance on the benchmark.</td>
</tr>
<tr>
<td>Multi-context interpretability</td>
<td>The degree to which the option has the potential to allow CP expenditure in countries to be compared relative to their context</td>
<td>The option has no potential to pay attention to how contexts differ between countries when comparisons are made.</td>
<td>The option has potential to allow too few context factors to be included. Therefore comparison between countries will take context into account highly insufficiently.</td>
<td>The option has the potential to include enough context factors in the benchmark itself. Therefore the comparison between countries will take context into account only to a limited degree.</td>
</tr>
<tr>
<td>Depth of understanding relative to child protection systems</td>
<td>The degree to which the option has the potential to provide information on the adequacy &amp; quality of expenditure relative to Child Protection Systems</td>
<td>This option has no potential to provide depth on the adequacy &amp; quality of child protection expenditures</td>
<td>This option has potential to provide some depth on the adequacy &amp; quality of child protection expenditures</td>
<td>This option has the potential to provide significant depth on the adequacy &amp; quality of child protection expenditures</td>
</tr>
<tr>
<td>Ease of presentation</td>
<td>The degree to which the option will allow disaggregation and clear, succinct presentation, for example through graphics</td>
<td>This option has no potential to allow disaggregation or clear succinct presentation</td>
<td>Components of this option could be disaggregated, and presented clearly and succinctly</td>
<td>This option can be disaggregated for the most part, and presented clearly and succinctly, but it may present methodological difficulties.</td>
</tr>
<tr>
<td>Ease of understanding and interpretation of differences in scores</td>
<td>The degree to which the option has potential to allow users of the benchmark easy understanding of what better or poorer performance on the benchmark means</td>
<td>This option has the potential that differences in performance will be understood but mostly only by highly numerate stakeholders</td>
<td>This option has the potential that differences in performance will be understood but mostly only by stakeholders with at least ten years of schooling</td>
<td>This option has the potential that differences in performance will be understood but mostly only by stakeholders even with limited or no schooling.</td>
</tr>
<tr>
<td>Simplicity of Methodology</td>
<td>The degree to which the methodology is robust &amp; defensible, contributing to the validity of the indicator</td>
<td>The option is likely to contain many assumptions or judgements, increasing the potential that almost all stakeholders or readers will challenge it / is unlikely to be considered credible by most user-groups</td>
<td>The option is likely to contain many assumptions or judgements, increasing the potential that at least key stakeholder groups will challenge it &amp; may consider it not credible</td>
<td>The option is likely to contain hardly any assumptions that are not commonly acceptable. It is unlikely that any stakeholders will challenge it &amp; most will consider it credible.</td>
</tr>
<tr>
<td>Cost &amp; practicality of application, all countries equal</td>
<td>The degree to which the option has the potential to be practical &amp; cheap to implement</td>
<td>The option is likely to require a significant investment in country application by a team of specialists contracted to apply it</td>
<td>The option is likely to require investment in country application by specialists, but some collection &amp; analysis can be done by UNICEF</td>
<td>The option is likely to require some country-specific application, but collection &amp; analysis can be done by UNICEF</td>
</tr>
<tr>
<td>Cost and Practicality of application relative to context</td>
<td>The complicatedness of data collection &amp; degree of cost under most circumstances</td>
<td>The option has the potential to be complicated &amp; costly in almost all circumstances</td>
<td>The option has the potential to be complicated &amp; costly in all countries except countries with advanced public</td>
<td>The option has the potential to be complicated &amp; costly in only few countries where public financial</td>
</tr>
<tr>
<td>Definition</td>
<td>Score of 0</td>
<td>Score of 1</td>
<td>Score of 2</td>
<td>Score of 3</td>
</tr>
<tr>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>financial management &amp; statistical systems</td>
<td>where public financial management &amp; statistical systems function poorly</td>
<td>management &amp; statistical systems function poorly</td>
</tr>
</tbody>
</table>
Description and assessment of benchmark options

Annex 3.11. The paragraphs below provide an illustrative description of each construction option identified in the mid-assignment report. It is important to note that many of the judgments made and the content proposed are not robust and can be challenged technically: the purpose at this stage is not to propose the actual content, but to illustrate how the different options could be used to communicate the content delineated. Annex 3 Table 3 on page 87 provides the assessment across indicators, and highlights the highest scoring indicators.

Annex 3.12. OPTION 1: Unitary indicator

The only logical unitary indicator option would be the quantum of qualifying expenditure on child protection, pushing the remainder of the content into the presentation.

- **Type description**: A single numerical integer
- **Illustrative CP Benchmark description**: Quantum of expenditure on child protection in US dollars (PPP 2005)
- **Presentation description**:
  - Presentation of the country ranking, overall, and among countries of a similar status assessed in the same period. This would require judgment on country characteristic identifiers, but not as part of the benchmark itself.
  - Standardised disaggregation of the benchmark would be possible (by economic classification, by component and type of service, by type of risk/harm, by funder). However, the benchmark itself does not require disaggregated expenditure data collection.
  - Disaggregation by gender and income group would require data on the incidence of expenditure by gender and income group, but not as part of the benchmark itself.
  - Presentation of benchmark as a percentage of GDP, and per child in need as a percentage of per capita spending. This would require additional data collection and analysis, but not as part of the core benchmark.
  - Presentation relative to other countries assessed in the same period, on a scatter plot with HDI ranking. This would require judgment on country characteristic identifiers, but not as part of the benchmark itself.
  - Presentation of an aggregate PEFA score for the PEFA measures selected, or the scores separately with discussion. If an aggregate score is provided it opens up the option to do a matrix analysis of benchmark scores relative to PEFA scores, grouping the country (e.g. high expenditure with low overall PEFA, or low expenditure but high governance). This would require additional data collection and analysis, but not as part of the core benchmark.

Annex 3.13. OPTION 2: Unitary summary indicator simple

The illustrative benchmark contextualizes the CP expenditure by country need and country financial capacity. Other information is pushed to the presentation.

- **Type description**: A single indicator but summary in nature. A simple numerator/denominator construction with 2 to 4 integer variables.
- **Illustrative CP Benchmark description**: Quantum of expenditure on child protection per child as a percentage of GDP per capita (conversion to USD PPP not required)
- **Presentation description**:

---

23 See paragraph 66 in the revised mid-assignment report for a brief discussion on a methodology to convert PEFA scores to numerical scores

Annex p89
annex p90

o The option allows presentation of the country ranking, overall, and among countries of a similar status assessed in the same period. This would require judgment on country characteristic identifiers, but not as part of the benchmark itself.

o Standardised disaggregation of the benchmark. This option does not allow the same ease of disaggregation as Option 1. Only some of the disaggregation can be done for the benchmark numerator as a whole without additional data collection (e.g. by economic classification, by funder). Others would require similarly disaggregating all the variables (the no of children in poverty and GDP per capita (e.g. by region)). For the expenditure by harm or risk or prevention or response, both the variables in the numerator would need to be disaggregated, but not the numerator.

o The benchmark numerator could be disaggregated by gender, region or income group, but it would require disaggregating both components of the numerator, and information on the incidence of expenditure for gender and income group. This is however not part of the core benchmark itself.

o Graphic presentation relative to other countries assessed in the same period, same as for Option 1.

o Presentation of PEFA scores as for Option 1 as interpretive data.


The illustrative benchmark contextualizes the CP expenditure by country need and country financial capacity, as well as by the balance of services offered in addressing risks versus responding to harm. It therefore builds in a global policy concern with regards to the balance. Countries at the same income level that spend more on prevention relative to protection, will have a higher score.

- **Type description**: A single indicator but summary in nature. A complicated indicator with 5 or more integer variables.

- **Illustrative CP Benchmark description**: Total expenditure on child protection times the ratio of total expenditure on prevention versus protection, divided by children in need, as a percentage of GDP per capita (conversion to USD PPP not required).

- **Presentation description**:
  o The option allows presentation of the country ranking, overall, and among countries of a similar status assessed in the same period. This would require judgment on country characteristic identifiers, but not as part of the benchmark itself.
  o Standardised disaggregation of the benchmark. As for Option 2 and compared to Option 1, the higher complexity of the benchmark will make disaggregation more complex.
  o The benchmark numerator could be disaggregated by gender, region or income group, but it would require disaggregating all the components of the numerator, requiring information on the incidence of expenditure for gender and income group. This is however not part of the core benchmark itself.
  o Graphic presentation relative to other countries assessed in the same period, same as for Option 1.
  o Presentation of PEFA scores as for Option 1 as interpretive data.

Annex 3.15. OPTION 4: Unitary index: change over time

This illustrative benchmark emphasizes the growth in countries expenditure on child protection over time. If growth (relative to need and financial capacity) is the key global policy priority, this benchmark will provide a suitable comparative measure.
• **Type description:** A single statistical measure of change relative to a selected benchmark.

• **Illustrative CP Benchmark description:** Expenditure on child protection per child as a % of GDP per capita expressed as an index with a specific year as a base point for all countries (conversion to USD PPP not required).

• **Presentation description:**
  - The option allows presentation of the country ranking, overall, and among countries of a similar status assessed in the same period. This would require judgment on country characteristic identifiers, but not as part of the benchmark itself.
  - Standardised disaggregation of the benchmark. If the benchmark were simply an index of growth in child protection qualifying expenditures over time, full disaggregation would be relatively simple to undertake. However, for the indicator used here, disaggregation becomes more complicated.
  - The benchmark could be disaggregated by gender, region or income group, but it would require information on the incidence of the underlying expenditure for gender and income group. This is however not part of the core benchmark itself.
  - Graphic presentation relative to other countries assessed in the same period, same as for Option 1.
  - Presentation of PEFA scores as for Option 1 as interpretive data.

Annex 3.16. **OPTION 5: Unitary index: relative standing**

This illustrative benchmark contextualizes the countries expenditure on child protection relative to need and financial capacity, taking into account different overall poverty levels.

• **Type description:** A single statistical measure of relative standing to a selected benchmark.

• **Illustrative CP Benchmark description:** Expenditure on child protection per child in need as a % of GDP per capita expressed as an index of the average across countries in the same quartile of countries ranked by proportion of the population living on less than USD2 a day.

• **Presentation description:**
  - The option allows presentation of the country ranking, overall. Judgment on some country characteristic identifiers however, in this case becomes part of the benchmark itself. In presentation it would be possible to rank countries of a similar status against characteristic not used in the core benchmark. For example, the presentation may look at benchmark scores of countries affected by emergencies or conflict in the same period separately.
  - Standardised disaggregation of the benchmark. If the benchmark were simply an index of standing in child protection qualifying expenditures relative to similar countries, full disaggregation would be relatively simple to present. However, for the indicator used here, disaggregation becomes more complicated.
  - The benchmark could be disaggregated by gender, region or income group, but it would require disaggregation of all components and information on the incidence of the underlying expenditure. This is however not part of the core benchmark itself.
  - Graphic presentation relative to other countries assessed in the same period, same as for Option 1.
  - Presentation of PEFA scores as for Option 1 as interpretive data.

Annex 3.17. **OPTION 6: Composite indicator: average of indexes** (discussed here, but a composite indicator can also be a composite of individual indicators rather than indexes)
This illustrative benchmark contextualizes comparative country expenditure on child protection relative to
country financial capacity and takes into account the balance between prevention and response, and the
governance of expenditure.

- **Type description**: A summary indicator calculated from a standardised set of components, such as an average of indexes or an average of indicators.

- **Illustrative CP Benchmark description**: The illustrative benchmark has three components, measuring different aspects of child protection expenditure.
  - Component 1 measures expenditure on prevention as a percentage of GDP relative to a minimum and maximum value set in the indicator (all in USD PPP).
  - Component 2 measures expenditure on response as a percentage of GDP relative to a minimum and maximum value set in the indicator design (all in USD PPP).
  - Component 3 measures the governance of child protection expenditure, and it expresses the average PEFA score on the selected measures relative to a minimum value (1) and maximum value (4).

  The nominator in each of the three indices is the actual value for the country minus the minimum value. The denominator is the maximum value minus the minimum value. The composite Child Protection Expenditure Index is then calculated as the geometric mean of the three dimension indexes.

- **Presentation description**:
  - The option allows presentation of the country ranking, overall. In presentation it would be possible to rank countries against other countries of a similar status. Judgment on status characteristics would not form part of the benchmark.
  - Standardised disaggregation of the benchmark. The benchmark already contains a disaggregation. Also, in order to disaggregate the minimum and maximum values of expenditure per GDP would need to be adjusted to reflect the disaggregation.
  - The benchmark could be disaggregated by gender, region or income group, but it would require disaggregation of all components and information on the incidence of the underlying expenditure and adjustment of the minimum and maximum value assumptions for the financial indexes. This is however not part of the core benchmark itself.
  - Graphic Presentation relative to other countries assessed in the same period as for Option 1.

**Annex 3.18. OPTION 7: Unitary composite indicator – set of assessments with standardized scoring methodology**

This illustrative benchmark contextualizes comparative country expenditure on child protection relative to
country financial capacity and takes into account a number of other factors that may be important to expenditure.

- **Type description**: A summary indicator calculated from a standardised set of components: a series of assessments with a standardised scoring methodology and summarised in one number.

- **Illustrative CP Benchmark description**: The illustrative benchmark has four components, measuring different aspects of child protection expenditure. Once all the scores have been converted to numerical scores as required by the specific component or sub-component, the scores are averaged to one score out of a 100.

  The number of components and their composition can be set as required. A component theme may also have more than one question. All components do not need to be of the same kind. Calculating a unitary score can first calculate the score for components, and then average the component scores. The unitary score can also have one leading component. For example the
lead indicator may rate the total child protection expenditure. All other components are averaged. The arithmetic mean of this average and the lead indicator score can then be calculated, meaning that half of the summary indicator is determined by the lead indicator and the remainder by the rest.

<table>
<thead>
<tr>
<th>Expenditure relative to financial capacity</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Component 1 Qualifying child protection expenditure relative to GDP</td>
<td></td>
</tr>
<tr>
<td>A: Child protection expenditure is 0.3% of GDP and larger</td>
<td>A score = 100</td>
</tr>
<tr>
<td>B: Child protection expenditure is smaller than 0.3% of GDP but 0.2 0.2% and larger</td>
<td>B score = 75</td>
</tr>
<tr>
<td>C: Child protection expenditure is smaller than 0.2% of GDP but 0.1% and larger</td>
<td>C score = 50</td>
</tr>
<tr>
<td>D: Child protection expenditure is smaller than 0.1% of GDP but larger than 0</td>
<td>D score = 25</td>
</tr>
<tr>
<td>E: There is no child protection expenditure</td>
<td>E score = 0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Priority given to child protection</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Component 2 Qualifying child protection internally financed expenditure as a proportion of tax revenue</td>
<td></td>
</tr>
<tr>
<td>A: Child protection expenditure is larger than 1% of tax revenue</td>
<td>A score = 100</td>
</tr>
<tr>
<td>B: Child protection expenditure is smaller than 1% of tax revenue but 0.75% and larger of tax revenue</td>
<td>B score = 75</td>
</tr>
<tr>
<td>C: Child protection expenditure is smaller than 0.75% of tax revenue but 0.5% and larger of tax revenue</td>
<td>C score = 50</td>
</tr>
<tr>
<td>D: Child protection expenditure is smaller than 0.5% of tax revenue but larger than 0</td>
<td>D score = 25</td>
</tr>
<tr>
<td>E: There is no child protection expenditure</td>
<td>E score = 0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Existence of adequate and efficient governance and accountability components</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Component 3 Expenditure on the governance and accountability components of child protection services (or personnel and capital etc)</td>
<td></td>
</tr>
<tr>
<td>A: Child protection expenditure on governance and accountability is between 12 and 15% of total expenditure on child protection</td>
<td>A score = 100</td>
</tr>
<tr>
<td>B: Child protection expenditure on governance and accountability is smaller than 12 but 9% or larger of total expenditure on child protection</td>
<td>B score = 75</td>
</tr>
<tr>
<td>C: Child protection expenditure on governance and accountability is smaller than 9% but 6% or larger of total expenditure on child protection</td>
<td>C score = 50</td>
</tr>
<tr>
<td>D: Child protection expenditure on governance is more than 15% or less than 6% of total expenditure on child protection</td>
<td>D score = 25</td>
</tr>
<tr>
<td>E: There is no child protection expenditure on governance and accountability</td>
<td>E score = 0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Governance of child protection expenditure</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Component 4 Governance of child protection expenditure</td>
<td></td>
</tr>
<tr>
<td>A: The average PEFA score for the selected measures is between 3 and 4.</td>
<td>A score = 100</td>
</tr>
<tr>
<td>B: The average PEFA score for selected measures is smaller than 3 but 2 or larger</td>
<td>B score = 67</td>
</tr>
<tr>
<td>C: The average PEFA score for selected measures is smaller than 2 but larger than 1</td>
<td>C score = 33</td>
</tr>
<tr>
<td>D: A PEFA has not been done</td>
<td>D score = component is not calculated</td>
</tr>
</tbody>
</table>

- **Presentation description**: Much of the interpretive information can be included in the benchmark.
- The option allows presentation of the country ranking, overall. In presentation it would be possible to rank countries against other countries of a similar status. Judgment on status characteristics would not form part of the benchmark.
- Standardised disaggregation of the benchmark: The benchmark will use some of the disaggregation elements (as illustrated in its description above). In principle it could lend itself to disaggregation of the financial components in dimensions not assessed in the
  
Annex p93
benchmark itself, for example benchmarks for expenditure on prevention could be reported separately. This however would require a re-setting of some of the standards implied by the scoring descriptions, depending on the criteria used.

- The benchmark could be disaggregated by gender, region or income group, but it would require disaggregation of some components (depending on the nature of the component elements) and information on the incidence of the underlying expenditure. This is however not part of the core benchmark itself.
- As it is a summary score, the benchmark could be plotted against relevant other indicators for countries.

Annex 3.19. OPTION 8: Framework of standardised indicators with standardised scoring methodology, not summarized

- **Type description:** A framework of indicators, with a standardised scoring methodology against each component, but not summarised into a single value
- **Illustrative CP Benchmark description:** Same as Option 7, but without summarizing the scores into a single value
- **Presentation description:** As the option is not summarized, it will be difficult to rank countries. Pairwise comparison, or comparison by component would be options to compare countries.
  - Standardised disaggregation – same as for Option 7, but component by component.
  - Disaggregation by gender, region or income group: same as for Option 7.
  - As there is no summary score, the benchmark does not lend itself to simple graphic presentations.

Annex 3.20. OPTION 9: Framework of indexes, but not summarised

- **Type description:** A framework of indexes with a standardised scoring methodology against each component, but not summarised into a single value
- **Illustrative CP Benchmark description:** same as Option 6 but without summarizing the scores into a single value
- **Presentation description:** As the option is not summarized, it will be difficult to rank countries. Pairwise comparison, or comparison by component would be options to compare countries.
  - Standardised disaggregation – same as for Option 6, but component by component.
  - Disaggregation by gender, region or income group: same as for Option 6.
  - As there is no summary score, the benchmark does not lend itself to simple graphic presentations.

Annex 3.21. Option 10: Framework of indicators/indexes, but the scoring methodology is not standardized

The illustrative benchmark measures a number of important factors for child protection expenditure, but does not attempt to standardize the scoring.

- **Type description:** A framework of indexes and indicators, that measure in a standardised way for each indicator, but different scoring values and calculation methods are used for each indicator. Some may be numerical and others not
- **Illustrative CP Benchmark description:** The illustrative benchmark has several components, each with a specific indicator proposed.
Component 1: Adequacy of child protection expenditure
Child protection expenditure per child in need as a percentage of GDP per capita

Component 2: Child protection expenditure relative to financial capacity
Child protection expenditure as a percentage of GDP

Component 3: Balance of prevention and response child protection expenditure
Ratio of expenditure on prevention to expenditure on response

Component 4: Adequacy of investment in policy planning and oversight
Expenditure on governance and accountability mechanisms as a percentage of total expenditure

Component 5: Priority given to child protection
Child protection expenditure financed by government as a percentage of tax revenue

Component 6: Governance of child protection expenditure
Country latest PEFA scores against identified measures

Presentation description: As the option is not summarized, countries cannot be ranked. Pairwise comparison, or comparison by component would be options to compare countries.
- Standardised disaggregation – component by component where possible.
- Disaggregation by gender, region or income group – component by component, provided that incidence of expenditure is known for gender and income groups.
- As there is no summary score, the benchmark does not lend itself to simple graphic presentations.

Annex 3.22. Annex 3 Table 5 overleaf sets out the scoring the indicators using the scoring methodology described in paragraph Annex 3.5 to paragraph Annex 3.10. On this methodology and relative weighting of individual criteria, the ranking of options is as follows.

Annex 3 Table 4 Ranking of options

<table>
<thead>
<tr>
<th>Option No</th>
<th>Short description</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Option 7</td>
<td>Unitary composite indicator: set of assessments with standardised scoring methodology</td>
<td>1</td>
</tr>
<tr>
<td>Option 2</td>
<td>Unitary summary indicator simple</td>
<td>2</td>
</tr>
<tr>
<td>Option 6</td>
<td>Unitary composite index / indicator: average of indexes</td>
<td>3</td>
</tr>
<tr>
<td>Option 8</td>
<td>Framework of standardised indicators with standardised scoring methodology, not summarised</td>
<td>4</td>
</tr>
<tr>
<td>Option 1</td>
<td>Unitary indicator</td>
<td>5</td>
</tr>
<tr>
<td>Option 4</td>
<td>Unitary index: change over time</td>
<td>6</td>
</tr>
<tr>
<td>Option 3</td>
<td>Unitary summary indicator complicated</td>
<td>7</td>
</tr>
<tr>
<td>Option 5</td>
<td>Unitary index: relative standing</td>
<td>7</td>
</tr>
<tr>
<td>Option 9</td>
<td>Framework of indexes, but not summarised</td>
<td>7</td>
</tr>
<tr>
<td>Option 10</td>
<td>Framework of indicators/indexes, but the scoring methodology is not standardised</td>
<td>10</td>
</tr>
</tbody>
</table>

Annex 3.23. The top three were assessed as follows:

Option 2 (Ranked 2): Simple Unitary Summary Indicator. This option scored well because it is easy to understand; is not likely to add significant cost to apply relative to the other options; involves fewer assumptions and judgments than other options besides those commonly accepted; and allows countries to be ranked. Given the simplicity of indicator construction, it will also lend itself to disaggregation, and presentation through standardized graphics. As is discussed below, this may allow more context-relevant information to be brought into comparisons. It however scored poorly against criteria assessing the degree to which context is taken into account comparing countries, and the depth of understanding it provides on the adequacy and quality of CP expenditure.
Option 7 (Ranked 1): A Unitary Composite Index / Indicator: this option scored well because it has the potential for country context to be taken into account in comparisons; to allow countries to be ranked; to provide understanding of the adequacy and quality of child protection expenditures as a part of the benchmark, and it can be understood reasonably well. As the option presents a unitary measure, it would lend itself to presentation through graphics against datasets not included in the core indicator itself, but may not easily disaggregate, depending on its components. It scored poorly on the potential that methodology will be questioned, and the cost and practicality of application.

Option 6 (Ranked 3): Unitary composite indicator: set of assessments with standardised scoring methodology: Similarly to unitary composite index/indicator, this option scored well because it has the potential for country context to be taken into account in comparisons; to allow countries to be ranked; to provide understanding of the adequacy and quality of child protection expenditures as a part of the benchmark, and it can be understood reasonably well. Similarly to the unitary composite index, countries’ scores on the benchmark could be presented in graphics utilizing other datasets in a standardized way, but not all its components may lend themselves to disaggregation. The indicator scored poorly on the potential for judgments and assumptions in the methodology to be questioned affecting its credibility (on account of involving more assumptions and judgments and requiring judgments about weighting), and the cost and practicality of its application.
### Annex 3 Table 5  
**Assessment of construction options**

<table>
<thead>
<tr>
<th>Indicator option</th>
<th>Indicator description</th>
<th>SET 1</th>
<th>SET 2</th>
<th>Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OPTION 1: Unitary indicator</strong></td>
<td>A single numerical integer</td>
<td>3 0 0 3 3 3 3</td>
<td>1 3</td>
<td>4.00</td>
</tr>
<tr>
<td><strong>OPTION 2: Unitary summary indicator simple</strong></td>
<td>A single indicator but summary in nature. A simple numerator/denominator construction with 2 to 4 integer variables.</td>
<td>3 1 2 2 3 3 3</td>
<td>1.667 2.8 4.47</td>
<td>2</td>
</tr>
<tr>
<td><strong>OPTION 3: Unitary summary indicator complicated</strong></td>
<td>A complicated indicator with 5 or more integer variables.</td>
<td>3 2 2 0 2 1 2</td>
<td>2.333 1.4 3.73</td>
<td>7</td>
</tr>
<tr>
<td><strong>OPTION 4: Unitary index: change over time</strong></td>
<td>A single statistical measure of change relative to a selected benchmark.</td>
<td>3 0 1 2 3 2 3</td>
<td>1.333 2.6 3.93</td>
<td>6</td>
</tr>
<tr>
<td><strong>OPTION 5: Unitary index: relative standing</strong></td>
<td>A single statistical measure of relative standing to a selected benchmark.</td>
<td>2 1 1 3 2 2 2</td>
<td>1.333 2.4 3.73</td>
<td>7</td>
</tr>
<tr>
<td><strong>OPTION 6: Unitary composite index / indicator</strong></td>
<td>A summary indicator calculated from a standardised set of components; whether indexes or indicators</td>
<td>3 3 3 2 2 0 1 2 3 1.4 4.40</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td><strong>OPTION 7: Unitary composite indicator: set of assessments with standardised scoring methodology</strong></td>
<td>A summary indicator calculated from a standardised set of components: a series of assessments with a standardised scoring methodology and summarised in one number.</td>
<td>3 3 3 2 2 1 1 2 3 1.6 4.60</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>OPTION 8: Framework of standardised indicators with standardised scoring methodology, not summarised</strong></td>
<td>A framework of indicators, with a standardised scoring methodology against each component, but not summarised into a single value</td>
<td>1 3 3 2 2 2 1 2</td>
<td>2.333 1.8 4.13</td>
<td>4</td>
</tr>
<tr>
<td><strong>OPTION 9: Framework of indexes, but not summarised</strong></td>
<td>A framework of indexes with a standardised scoring methodology against each component, but not summarised into a single value</td>
<td>1 3 3 2 1 1 1 2</td>
<td>2.333 1.4 3.73</td>
<td>7</td>
</tr>
<tr>
<td><strong>OPTION 10: Framework of indicators/indexes, but the scoring methodology is not standardised</strong></td>
<td>A framework of indexes and indicators, that measure in a standardised way for each indicator, but different scoring values and calculation methods are used for each indicator. Some may be numerical and others not.</td>
<td>0 3 3 1 2 3 0 0</td>
<td>2 1.2 3.20</td>
<td>10</td>
</tr>
</tbody>
</table>

Annex p97