

WFP – OEV

# **Technical Note on Efficiency Analysis**

Robrecht Renard and Stephen Lister

14 June 2013

## Contents

<b>About this guidance</b>	<b>iv</b>
<b>I. Concepts and definitions</b>	<b>1</b>
<b>II. Useful Tools for WFP Efficiency Analysis</b>	<b>9</b>
<b>III. Tool Guide 1: Unit Cost Benchmarking</b>	<b>11</b>
<b>A. Overview</b>	11
<b>B. Description</b>	11
<b>C. Efficiency analysis</b>	12
<b>D. How to benchmark unit costs</b>	12
<b>E. Example</b>	14
<b>F. Link with related WFP practice - alpha values</b>	15
<b>IV. Tool Guide 2: Cost-Effectiveness Analysis</b>	<b>18</b>
<b>A. Overview</b>	18
<b>B. Description</b>	18
<b>C. Efficiency Analysis</b>	18
<b>D. How to perform CEA</b>	19
<b>E. Example</b>	22
<b>F. Link with related WFP practice – omega values</b>	23
<b>V. Tool Guide 3: Specific Evaluation Questions</b>	<b>24</b>
<b>A. Overview</b>	24
<b>B. SEQs Description</b>	24
<b>C. How to develop SEQs on Efficiency</b>	25
<b>D. Examples of SEQs on Efficiency</b>	26
<b>Annex A Non-Standard Definitions of Efficiency</b>	<b>29</b>
The relationship between broad and narrow definitions	29
Why the broad definition is preferable	30
<b>Annex B Additional Efficiency Tools</b>	<b>32</b>
Cost Benefit Analysis (CBA)	33
Cost-Utility Analysis (CUA)	34
Multi-Attribute Decision Model (MADM)	35
Follow the Money	36
Benchmarking Other Indicators	37
<b>Useful Links/Guide for Further Reading</b>	<b>38</b>
<b>Acronyms and abbreviations</b>	<b>40</b>

## Figures

Figure 1	Efficiency and effectiveness in the logical framework.....	4
Figure A1	Efficiency and effectiveness in the logical framework.....	29
Figure A2	The “three Es” as interpreted by DFID .....	30
Figure A3	The “three Es” and Value for Money .....	30

## Boxes

Box 1	The DAC Evaluation Criteria .....	1
Box 2	Cost-effectiveness ranking reversal – an illustration .....	22
Box A1	The efficiency concept in economics .....	31
Box A2	Efficiency and Value for Money.....	31

## Tables

Table 1	Efficiency and Effectiveness indicators in the Logical Framework for WFP.....	7
Table 2	Characteristics of selected efficiency tools.....	10
Table 3	Benchmarking of school-meal-only \$ costs per recipient per year .....	14
Table 4	Alternative outcome rankings of three food modalities .....	22

## About this guidance

This Technical Note is a support to OEV's EQAS system. All five of the standard OECD DAC evaluation criteria are embedded in EQAS, but there has been little specific guidance on efficiency, and its treatment in WFP evaluations has been patchy. This note therefore provides guidance for evaluation managers (EMs) and evaluation teams which should lead to more systematic and rigorous treatment of efficiency in future.

The Technical Note comprises the following sections:

- Section I defines and explains the efficiency criterion, and shows how it relates to other evaluation criteria, especially effectiveness.
- Section II is a concise guide to the eight tools for efficiency analysis which have been identified as most relevant for OEV work.
- Sections III, IV and V provide a more detailed explanation of three of the tools – Unit Cost Benchmarking, Cost-Effectiveness Analysis and Specific Evaluation Questions respectively.
- Annex A clarifies competing definitions of efficiency.
- In o, five further tools are discussed – Cost-Benefit Analysis, Cost-Utility Analysis, Multi-Attribute Decision Models, Benchmarking Other Indicators, and Follow the Money. These tools are somewhat too complex in data needs and time, or less relevant for use in typical OEV work, but may nevertheless sometimes be appropriate.
- A short bibliography provides suggestions for further reading.

A background paper (Renard 2013), which is also available from OEV, provides more explanation of the concepts and tools summarised in this note.

## I. Concepts and definitions

### *Efficiency as an evaluation criterion*

1. Along with most aid agencies, WFP OEV uses the standard DAC evaluation criteria: relevance, efficiency, effectiveness, impact and sustainability (see Box 1 below).

#### **Box 1 The DAC Evaluation Criteria**

<b>Criterion</b>	<b>Definition and Sample Questions</b>
<b>Relevance</b>	<p><b>The extent to which the aid activity is suited to the priorities and policies of the target group, recipient and donor.</b></p> <p>Useful questions:</p> <ul style="list-style-type: none"> <li>• To what extent are the objectives of the programme still valid?</li> <li>• Are the activities and outputs of the programme consistent with the overall goal and the attainment of its objectives?</li> <li>• Are the activities and outputs of the programme consistent with the intended impacts and effects?</li> </ul>
<b>Effectiveness</b>	<p><b>A measure of the extent to which an aid activity attains its objectives.</b></p> <p>Useful questions:</p> <ul style="list-style-type: none"> <li>• To what extent were the objectives achieved / are likely to be achieved?</li> <li>• What were the major factors influencing the achievement or non-achievement of the objectives?</li> </ul>
<b>Efficiency</b>	<p><b>Efficiency measures the outputs -- qualitative and quantitative -- in relation to the inputs. It is an economic term which signifies that the aid uses the least costly resources possible in order to achieve the desired results. This generally requires comparing alternative approaches to achieving the same outputs, to see whether the most efficient process has been adopted.</b></p> <p>Useful questions:</p> <ul style="list-style-type: none"> <li>• Were activities cost-efficient?</li> <li>• Were objectives achieved on time?</li> <li>• Was the programme or project implemented in the most efficient way compared to alternatives?</li> </ul>
<b>Impact</b>	<p><b>The positive and negative changes produced by a development intervention, directly or indirectly, intended or unintended. This involves the main impacts and effects resulting from the activity on the local social, economic, environmental and other development indicators. The examination should be concerned with both intended and unintended results and must also include the positive and negative impact of external factors, such as changes in terms of trade and financial conditions.</b></p> <p>Useful questions:</p> <ul style="list-style-type: none"> <li>• What has happened as a result of the programme or project?</li> <li>• What real difference has the activity made to the beneficiaries?</li> <li>• How many people have been affected?</li> </ul>
<b>Sustainability</b>	<p><b>Sustainability is concerned with measuring whether the benefits of an activity are likely to continue after donor funding has been withdrawn. Projects need to be environmentally as well as financially sustainable.</b></p> <p>Useful questions:</p> <ul style="list-style-type: none"> <li>• To what extent did the benefits of a programme or project continue after donor funding ceased?</li> <li>• What were the major factors which influenced the achievement or non-achievement of sustainability of the programme or project?</li> </ul>
<p>Source: <a href="http://www.oecd.org/dac/evaluation/dacriteriaforevaluatingdevelopmentassistance.htm">http://www.oecd.org/dac/evaluation/dacriteriaforevaluatingdevelopmentassistance.htm</a>, June 2013.</p>	

### *Competing definitions of efficiency*

2. This Note uses the definition of efficiency that appears in the DAC multi-lingual “Glossary of Key Terms in Evaluation and Results Based Management”: “Efficiency is a measure of how economically resources/inputs (funds, expertise, time, etc.) are converted to results”. (DAC 2010a)<sup>1</sup> In line with the logical framework terminology, “results” in this definition can be either activities, outputs, outcome or impact.

3. The definition in Box 1 above (also from DAC) acknowledges that efficiency is a term derived from economics, but implies that efficiency analysis is more narrowly focused on the relationship between inputs and outputs. This ambiguity is pervasive, with many donors seeking to apply more restrictive terminology (as in the “three Es” terminology of economy, efficiency and effectiveness).

4. The broad definition is preferred because – as discussed below (see ¶18) – it allows a consistent application of efficiency techniques to answer both key questions: (a) “are we doing things right?” and (b) “are we doing the right things?”, and OEV is properly concerned with both these questions. At the same time, the broad definition helps to explain the relationship between the various partial definitions of efficiency, as well as the significance of “value for money”. Annex A provides a guide to the different efficiency concepts commonly encountered.

### *Complementarity between evaluation criteria*

5. All five DAC evaluation criteria cover important dimensions. Yet they are also interconnected. Let us give an example. Suppose that we have established that a planned in-kind food aid distribution is *relevant* from a development perspective. Suppose further that there are good reasons to expect that the intervention is going to be *effective* in what it sets out to do, and that there are equally good reasons to expect that it will have a positive development *impact*, some of which will be *sustainable* through time. The *efficiency* criterion then addresses the complementary question whether the benefits will outweigh the costs.

6. This is an important consideration. For even if the intervention scores well on relevance, effectiveness, impact and sustainability, the money poured into it could have been used to contribute to development in some different way, i.e. by fostering rural development. Would this have been a better use of scarce resources? And even if we establish that our planned intervention is in itself worth doing, compared with alternatives elsewhere in the economy, it may still be that there are alternative ways to provide food assistance in this particular context that achieve better results, or achieve the same result with fewer resources, for instance by a more focused targeting of beneficiaries (gender dimension), adopting another intervention design (vouchers), by another sourcing of some of the crucial inputs (local procurement), or by coupling food assistance with a health intervention.

7. Efficiency is in particular dependent on effectiveness and impact, in the sense that we have to establish first that an intervention is effective and has impact, before we can tackle the efficiency dimension.

---

<sup>1</sup> The same definition was used in the recent review of efficiency at WFP presented to the Executive Board (WFP 2011), and in the BMZ review of efficiency tools which was a starting point for the preparation of this technical note (Palenberg 2011).

### *Efficiency and the project life-cycle*

8. Efficiency analysis can be applied throughout the life cycle of an intervention:

- before (ex ante) – as an aid to project selection and design;
- during implementation – to help monitor and optimise performance;
- after (end of project, or ex post) – for learning and accountability.

9. In the case of OEV, evaluation analysis takes place during or after an intervention. Information is then available from implementation that was absent during project preparation. OEV evaluations can therefore test whether the assumptions made during project preparation about its efficiency have been verified during implementation. At the same time, ex ante and ex post efficiency analysis are two sides of the same coin, so it is important for WFP to adopt consistent approaches to efficiency analysis across the life cycle of its interventions.

### *Efficiency and equity*

10. Equity is not a separate criterion in the DAC list. This is not because it is not important, far from it. But it must be addressed within each of the five main criteria. This is certainly the case with effectiveness. WFP aims to reach the poorest and the neediest, and therefore a school feeding programme that would be biased in favour of children from better-off regions and social groups would not be considered effective. This would be captured by a measure of effectiveness such as the extent to which a food consumption deficit has been filled. In the case of children of well-to-do backgrounds without a food consumption deficit, effectiveness would be zero. Efficiency analysis must likewise be infused with an equity dimension. If WFP were to reduce costs by reducing its food assistance in difficult to reach, isolated regions, where most of the food insecure live, then it would become less efficient rather than more efficient. And efficiency measures must either capture this or be complemented by other criteria that do so.

### *Efficiency and the logical framework*

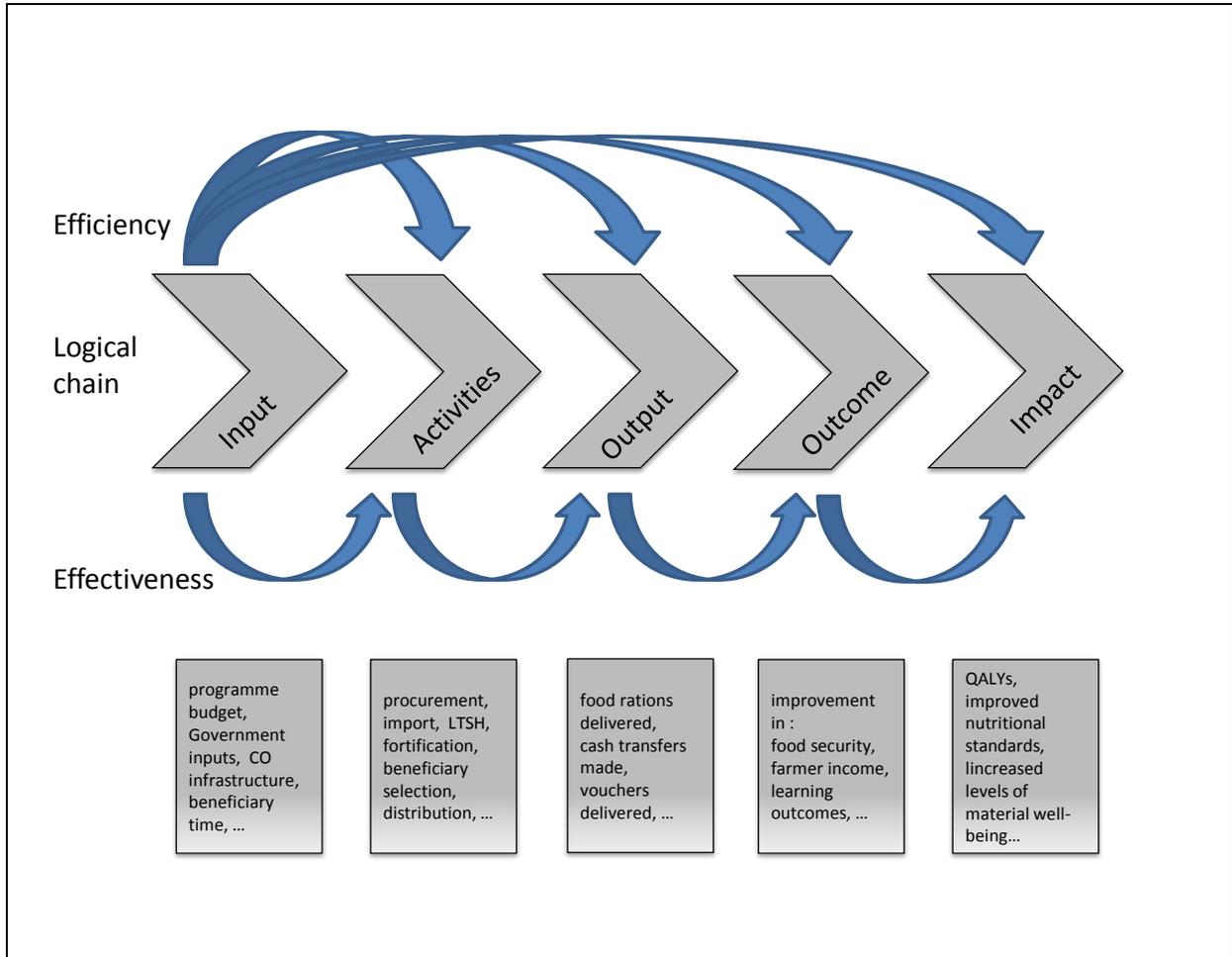
11. Following the DAC, efficiency is a measure that relates inputs to results. Inputs stand for resources provided by WFP, by national governments and by beneficiaries. Inputs are scarce and must be used well. In the logical framework terminology, “results” stand for either output (food assistance delivered to beneficiaries), outcome (nutritional improvement, improved learning outcomes, etc), or impact (lives saved or improved quality of life, etc). Efficiency also applies to the level just below output, i.e. activities (e.g. food delivered to warehouses, production of fortified food). In the resulting broad definition, efficiency can thus be applied to four hierarchical levels, as shown in Figure 1 below. As we shall see, efficiency analysis becomes progressively more complex as we move up the levels from activities to outputs, outcomes and impacts.

### *Efficiency and effectiveness*

12. It is important to understand the complementary relation between efficiency and effectiveness. The DAC definition of effectiveness (Box 1 above) refers to the extent to which a development intervention’s objectives are expected to be achieved (appraisal and interim evaluation), or have been achieved (end of project and ex post evaluation). In this definition “objectives” can again be best understood as pertaining to the different levels of the logical framework: activities, outputs, outcomes, and impact. To see the relevance of these different types of effectiveness, consider that a WFP intervention may have experienced “pipeline

breaks” at activities level, be nevertheless reasonably effective in securing output (the right food at the right time for the right beneficiaries), yet may not achieve its outcome (improved nutrition). Another intervention may contribute both to output and outcome, but without being successful in its intended impact (reduced child mortality).

**Figure 1 Efficiency and effectiveness in the logical framework**



13. Figure 1 above summarises these general concepts of efficiency and effectiveness and the links between them. Four levels of efficiency analysis are thus identified: *activities efficiency*, *output efficiency*, *outcome efficiency*, and *impact efficiency*. The corresponding concepts for effectiveness are *activities effectiveness*, *output effectiveness*, *outcome effectiveness*, and *impact effectiveness*. The top arrows express that in efficiency analysis a comparison is always made between inputs (costs) and a given result level, whereas the bottom arrows express that in effectiveness analysis the relationship is one of causality, where only those effects that can be attributed to the achievements of previous level count.

14. Two important implications follow from this. First, *an intervention that is ineffective cannot be efficient. But an intervention that is effective may not be efficient*. It will not be efficient if it achieves its results at too high a cost compared to other alternative courses of action. This establishes a hierarchy between effectiveness and efficiency analysis. Effectiveness analysis is important, but it needs to be complemented with efficiency analysis if we wish to establish that aid resources have been well used.

15. There is also a hierarchy between lower and higher levels of efficiency analysis. Ideally, efficiency analysis at impact level should be the main guide to ex ante decision making and interim and ex post evaluation of interventions. In practice, lower levels of efficiency analysis are often used because they are easier to apply, both in terms of skills and data requirements. But there is a trade-off, as the relevance of the results cannot be taken for granted.

16. A second implication is that *efficiency analysis builds on effectiveness analysis: efficiency cannot be assessed unless effectiveness has first been calculated*. The assessment of effectiveness is often very difficult, because we need to compare the results of an intervention with what would have happened in the absence of the intervention. But this “counterfactual” cannot be directly observed. It can at best be estimated indirectly. A random assignment of potential beneficiaries to an exposed and a control group, taking into account relevant characteristics such as age, gender, education, wealth, and income, is the best way to ensure that both groups are as similar to one another as possible. Such an experimental design is the gold standard in science, and has in the recent past been vigorously advocated in the field of development aid. In its absence, a host of less robust quasi-experimental designs are available, but they are less precise, not less complex analytically, and require data that is not routinely collected by either donor agencies or national governments.

17. The intrinsic difficulty of effectiveness analysis tends to increase as we move up in the logical chain from activities to output, outcome and impact. Thus, for example:

- We may be able to establish relatively easily whether WFP has been effective at activities level (e.g. whether food is available in a central warehouse in the country).
- It is more difficult to establish that the right food has been distributed at the right time to the right beneficiaries, and thus that output effectiveness has been achieved.
- It is even more difficult to demonstrate that beneficiaries are better nourished than they would have been in the absence of the food aid.
- Most difficult of all is to demonstrate whether a measure of impact such as stunting has improved as a consequence of the WFP intervention.

*Efficiency relates to both choice and implementation of interventions*

18. Efficiency analysis is both about “doing the right things” and “doing things right”. “Doing the right things” refers to strategic choices, for instance between alternative modalities of food assistance (food-in-kind, vouchers or cash), or between alternative ways of boosting learning outcomes (reducing class sizes, increasing teacher salaries, or providing school meals), taking each time equity issues on board (are we reaching the beneficiaries who need it most). Ideally such efficiency analysis of aid interventions is already performed at the decision making stage. The main efficiency tools to do this are outcome and impact efficiency analysis. But aid agencies must also “do things right”. This is about implementation, and the main tools to assess the operational efficiency of aid agencies are activities and output effectiveness and efficiency, mostly applied once the project is in operation or when it has come to an end. Note that an impact efficiency analysis performed ex post will take both dimensions on board.

*Examples of efficiency and effectiveness indicators for WFP*

19. In the logical framework of most WFP interventions a modality of food assistance (food-in-kind, vouchers, cash) is the output and improved food security the outcome. But often these interventions also pursue other objectives than nutrition, such as educational achievements (school feeding), better health through attendance of medical clinics (mother and child feeding) or farm income (P4P).

20. In Table 1 below we propose a generic logical framework for WFP interventions. Nutrition is a key link in the intervention theory, but the logical framework branches out to the other objectives. The table contains a list of often encountered indicators of effectiveness and efficiency at different levels of the logical framework, both for nutrition and (in italic) for the other dimensions.

21. Note that Table 1 is an important resource for evaluation managers as they develop TOR relative to efficiency – see the guidance in Section V below.

**Table 1 Efficiency and Effectiveness indicators in the Logical Framework for WFP**

	<b>Logframe</b>	<b>Effectiveness indicators</b>	<b>Efficiency indicators</b>
<b>Impact</b>	Nutritional status improved <i>Sustainable human development broadly defined</i>	$\Delta$ weight for height (short term measure) $\Delta$ height for age (long term measure) $\Delta$ body mass index (BMI) # deaths averted $\Delta$ Quality-adjusted life years (QALYs) $\Delta$ household income	Cost per quality-adjusted life year or QALY (Cost utility analysis) Net benefits using value of statistical life (VSL) estimates Cost per 1% improvement in nutritional indicator Beneficiary satisfaction scores Equity scores
<b>Outcome</b>	Food security enhanced <i>Improved education (learning outcomes)</i> <i>Mother-and-child health improvement</i> <i>Farmer income increased (P4P)</i>	$\Delta$ food diversity index $\Delta$ food consumption score $\Delta$ food consumption group $\Delta$ nutritional value score (NVS) <i><math>\Delta</math> farmer income per capita</i> <i><math>\Delta</math> school enrolment and dropout (gender differentiated)</i> <i><math>\Delta</math> end of schooling cognitive scores</i>	Cost per 1 % improvement in summary food security indicator Cost per kcal consumed Cost per nutritional value unit consumed Omega value <sup>2</sup>
<b>Output</b>	Aid (food, cash, vouchers) delivered <i>Assets created</i> <i>Disaster risk mitigation (DRM) measures in place</i>	# beneficiaries reached # MT food distributed # Kcal distributed # standardised daily nutrition rations distributed # cash/vouchers distributed # units of FBF as % total food distributed # <i>community assets created</i> # <i>of schools assisted, also as % of total</i> # <i>pupils trained</i> Beneficiary scores of project and staff Description of DRM measures realised	Cost per recipient Cost per standardised ration delivered Food market value of transfer of food commodity/cost of the transfer of food commodity (alpha value <sup>3</sup> ) Cost of transfer of \$1 of food or food purchasing power (1/alpha) Delivery cost per 1\$ of food or food purchasing power (1/alpha -1) Cost per kcal provided Cost per nutritional value unit provided Cost per nutritional value unit provided modality 1/cost per nutritional value unit provided modality 2 (omega ratio)

<sup>2</sup> See Section IV F for a discussion of omega values.

<sup>3</sup> See Section III F for a discussion of alpha values.

Understanding Efficiency – Definitions

	<b>Logframe</b>	<b>Effectiveness indicators</b>	<b>Efficiency indicators</b>
<b>Activities</b>	Needs assessment Preparation of cash, vouchers modalities Securing food (donor, triangular, local) Transport, LTSH Food fortification Capacity building Monitoring, ...	MT of food procured/warehoused # months of delays in construction Food “pipeline break” statistics Staff turnover statistics Staff recruitment statistics Timeliness of (donor and government) contributions in relation to the budget cycle	Cost per MT, broken down for commodities, transport, LTSH, DSC, ODOC, ISC, costs borne by government or beneficiaries CBA of production of FBF, financial analysis of warehouse construction, cost analysis of truck transport, ... Administrative costs of raising funds
<b>Input</b>	Donor financial support through WFP Donor food in kind donated through WFP Partner government resources Beneficiary contributions	Donor funding (actuals) Government contributions (actuals) Beneficiary contributions (actuals)	Not applicable <sup>4</sup>

---

<sup>4</sup> This is not to argue that inputs can be taken for granted in WFP interventions, like manna from heaven. On the contrary, securing sufficient and timely contributions by donors and governments is an important challenge, involving costs of its own, and failings in this respect may affect the efficiency of operations in a major way. For reasons of the syntax of the logical framework these considerations are here listed in the row above, of activities, the level at which they start to have an effect.

## II. Useful Tools for WFP Efficiency Analysis

22. Although there is a common set of underlying principles, there is no unique way of doing efficiency analysis. There are several reasons for this:

- (a) Inputs can be compared to results at various stages in the logical framework. The higher the stage, the more challenging and complex the efficiency analysis becomes.
- (b) The scope of the analysis can also differ. At one end, we may wish to compare two close alternatives, such as the provision of either cooked meals or biscuits in school feeding, from the perspective of a single objective, most likely nutrition. At the other end, we may wish to compare the efficiency, in terms of the twin objectives of nutrition and learning outcomes, of school feeding programmes with interventions that do not involve food assistance, such as selective lowering of tuition costs for poorer parents. The first is a much less complex question to address than the second.
- (c) Further, efficiency analysis can be performed with different degrees of sophistication, from detailed studies that involve primary data collection and the use of advanced economic techniques, down to back-of-an-envelope calculations that only involve simple arithmetic.

23. What type of efficiency analysis it is feasible and desirable to carry out in any particular case depends on several elements, such as the time and budget constraints, the availability of data, and the policy relevance of the efficiency question being addressed. For this reason, a broad range of tools have been developed over time that are all part of the efficiency toolbox.

24. OEV conducted a review of available tools and identified 8 that are the most likely to be useful in the context of WFP (Renard 2013). An overview of the different methods is provided in Table 2 below.

25. The 8 tools were selected from a larger 15-set toolbox proposed in a study for BMZ (Palenberg 2011). The main criteria were their relevance and applicability to OEV evaluations. In a second round 3 tools (shaded columns in Table 2) were identified as the ones most likely to be of immediate use. They are explained in Sections III, IV and V with reference to WFP situations, including through an example. The remaining 5 tools are briefly presented in o

.

**Table 2 Characteristics of selected efficiency tools**

		Cost-Benefit Analysis (CBA)	Cost-Effectiveness Analysis (CEA)	Cost-Utility Analysis (CUA)	Intuitive Multi-Attribute Decision Model (MADM)	Benchmarking Unit Costs	Benchmarking other indicators	Follow the money	Specific evaluation questions
features	analytical rigour (BMZ scores)	high	high	high	high	moderate	moderate	moderate	low
	location in logical framework	activities; impact	all levels	Impact	impact	inputs, activities	inputs, activities	activities	all levels
	intervention cycle	throughout	throughout	throughout	throughout	mid-term	mid-term	mid-term	mid-term
Requirements	data needs	moderate to very heavy	heavy to very heavy	very heavy	moderate	light	moderate	light	very light
	economic skills required	advanced	advanced	advanced	none	limited	none	limited	none
	WFP interventions	all	single objective	nutrition and/or health	all	all	all	all	all
Applicability	individual interventions	✓	✓	✓	✓	✓	✓	✓	✓
	country portfolios				✓	✓	✓	✓	✓
	role of OEV evaluators	verify	verify	Verify	do	interpret	interpret	do	do

Source: Based on Palenberg 2011.

### III. Tool Guide 1: Unit Cost Benchmarking

#### A. Overview

##### *Introduction*

26. The purpose of this guide is to show:
- how the benchmarking of unit costs can be a useful tool for assessing managerial efficiency in the WFP logistics chain;
  - how this tool differs from other tools, and in particular cost-benefit analysis (CBA) and cost-effectiveness analysis (CEA) with which it can sometimes be confused;
  - what are the challenges and pitfalls in getting relevant data, what calculations must be performed, and how the results can be interpreted.

##### *Why is this tool important?*

27. This tool focuses on managerial efficiency (“doing things right”). It compares different cost components of an intervention with those of similar interventions, in order to detect how an intervention scores against corporate or outside benchmarks. The result needs to be complemented by a careful contextual analysis, because not all negative deviations from benchmarks can be attributed to management failures, nor can all positive deviations be considered management successes.

#### B. Description

28. Unit costs are calculated as the ratio of costs to some relevant unit (food, staff, beneficiary, etc ) and at first sight can be confused with cost-effectiveness ratios (CERs) used in cost-effectiveness analysis (CEA – see Section IV below) or benefit-cost ratios (BCRs) used in cost-benefit analysis (CBA – see o). But this is a tool with more humble ambitions: CEA is used to compare two or more close alternatives, and CBA to assess whether an intervention is worth undertaking from a development perspective, but unit cost calculations only aim to test whether the costs of an intervention have been kept under control.

29. In the logical framework, this tool is therefore situated at the lower levels of inputs (food imported or secured locally, staff, etc) and activities (transport, warehousing, food fortification, delivery, etc), up to the level of output (food assistance delivered). The calculating of unit costs and the comparison with some external benchmark can be usefully performed during mid-term or ex post evaluation. As noted in ¶135 below, it may also be useful to monitor trends in unit costs, even if an external benchmark is not available.

30. Special care must be taken to establish appropriate benchmarks. The economic skills required are limited and do not require specific training, but an expertise in logistics may be needed to properly interpret the data and probe for underlying causes which may or not be linked to managerial efficiency.

### **C. Efficiency analysis**

#### *Efficiency dimensions captured*

31. This tool assesses how a given intervention has managed its costs in comparison to an average of similar WFP interventions. This is an important aspect of the broader efficiency issue, and one that WFP has invested a lot in addressing. The analysis can also be applied to some benchmarks outside WFP.

#### *Efficiency dimensions not captured*

32. The tool focuses on the input (cost) side. This may not be obvious, because costs are divided by some measure of results, either in monetary or non-monetary terms, and this may suggest that either CBA or CEA is performed. In effect, the division is only intended to standardise the costs and make them comparable to the benchmark, and does not have the aim of studying the efficiency with which inputs are transformed into results. Indeed, unit cost benchmarking does not aim to establish whether an intervention constitutes value for money compared with all other alternative courses of action, including the option of not intervening, as cost-benefit analysis (CBA) purports to do, nor does it gauge, more narrowly, whether an intervention is better at transforming inputs into results than some close alternatives, as CEA does. Because unit cost benchmarking algebraically resembles CERs or benefit-cost ratios (BCRs), and because unit cost benchmarking calculations are easier to perform, the temptation to interpret the results as informing about broader efficiency outcomes is understandable but erroneous. It is important that evaluators keep these limitations in mind.

33. It is also a tool that in general does not address equity issues well. The poorest people often live in remote areas and are more expensive to reach. A programme that has an unusually strong focus on the poor may therefore have higher unit costs.

### **D. How to benchmark unit costs<sup>5</sup>**

#### *Data requirements*

34. Two series of cost data are required, one for the intervention being studied, and one for a WFP-wide set of comparable interventions. For the intervention, both the costs and a measure of the results are required. The latter can be non-monetary (MT of food) or monetary (value of food ration, value of assets created, total funds resourced by CO per year). WFP has an elaborate system for monitoring the logistics chain, from resourcing of funds, through procurement, international and national transport, to warehousing and final distribution. Many basic data are collected in the process and reported in internal documents or registered in electronic databases. But this does not mean that such data have been suitably processed for the purpose of unit cost benchmarking, or that external consultants have access to them. In principle, external evaluators should be able to calculate, with information from the CO, unit costs for the intervention they are assessing, whereas WFP Rome should be able on request to provide relevant corporate benchmarks. Indeed, it is difficult to see how external evaluators could perform the analysis convincingly on their own.

---

<sup>5</sup> An example of how benchmarking unit costs could appear in an evaluation matrix is provided in Section V, subsection D below.

35. A somewhat less ambitious form of benchmarking is not to compare with interventions in other countries, but with the trend in the country in question. Within its management results framework, WFP has in place a number of key performance indicators to measure cost-efficient delivery, i.e. the average annual cost of beneficiary, average cost of food assistance (food rations, cash and vouchers), and average cost of technical assistance delivered. Currently, country offices and regional bureaux will get key data elements to calculate the relevant values, and eventually the values will be automatically generated and populated into the Performance and Risk Management Organizational Information System (PROMIS). PROMIS will facilitate aggregation and show trends across different countries and regions. Evaluators may usefully access these trends as a reference, and furthermore base medium to long term annual trends within a given country as a benchmark for unit costs of that particular country.

#### *Calculations*

36. Unit costs can be meaningfully calculated per type of intervention, such as emergency feeding (EMOP), food assistance during protracted relief and recovery (PRRO), including school feeding or food for assets (FFA). In addition unit costs can be calculated for country portfolios as a whole, in particular for the cost of staff, infrastructure and other overheads. Suitable measures must in each case be identified, for instance: \$ cost per MT of rice of a given standard delivered into a central warehouse, \$ cost per \$ of asset created in FFA, \$ cost per \$ of food ration provided in school feeding, \$ cost of CO staff per \$ of aid resources secured. Both costs and results must be measured over the same period of time.

#### *Interpretation*

37. Calculating unit costs makes sense if a benchmark is available for WFP as a whole to compare with. The interpretation is then straightforward. If unit costs are lower than the WFP-wide benchmark, then (other things equal) the intervention scores well on cost-efficiency. In order for this interpretation to hold, benchmarks must be sufficiently disaggregated. For instance, comparing school feeding programmes without any further qualification is problematic. A cost study on school feeding distinguished between "meals only", "meals + take-home rations", "take-home rations only", and "biscuits only" (WFP 2012). As cost items vary considerably across these alternatives, disaggregation is necessary if we wish to compare like with like.

38. A complication is that other things are often not equal, and only part of the deviation from the global average can be attributed to the managerial efficiency of the individual intervention or the country portfolio. There are indeed many exogenous drivers of cost differences such as whether this is a pilot project or not, project size, country political situation, geography of the country and distance from major food supplying countries, and all of these influences should be netted out.<sup>6</sup>

---

<sup>6</sup> This can be done through a statistical regression analysis in which the influence of these exogenous variables on unit costs is calculated. It is the variation not explained by these exogenous variables that can be attributed to differences in managerial efficiency. More simply, but less robustly, evaluators may make comparisons within a subset of countries that are more similar in key features.

39. When benchmarking within the same country, the trend data will show that the costs are either going up or down. There may be good reasons for this, which the evaluators have to discuss with the country office before they make any efficiency claims on this basis.

40. Finally, it should be taken into account that setting WFP averages as the standard of performance makes good sense for internal management purposes, but may not satisfy external accountability needs. Only if it is assumed that WFP on the whole is doing well at keeping its costs under control, does average performance across countries constitute a standard of good practice. For advocacy purposes, a comparison with other donors delivering similar interventions will be more convincing.

### ***E. Example***

41. We constructed a hypothetical example on school feeding, presented in Table 3 below.

**Table 3** Benchmarking of school-meal-only \$ costs per recipient per year

	WFP benchmark Uncorrected		WFP benchmark corrected <sup>b</sup>		Actuals country	
	\$	%	\$	%	\$	%
	(1)	(2)	(3)	(4)	(5)	(6)
commodity	19,2	48%	23.5	42%	23.8	36%
sea transport	1.6	4%	2.3	4%	3.7	6%
LTSH	8.0	20%	17.3	31%	19.4	29%
ODOC	2.4	6%	3.1	6%	10.5	16%
support	8.8	22%	9.6	17%	9.4	14%
<b>total unit costs</b>	<b>40.0</b>	<b>100%</b>	<b>55.8</b>	<b>100%</b>	<b>66.8</b>	<b>100%</b>
market value <sup>a</sup>	30.0		39.0		41.0	
alpha value	<b>0.75</b>		<b>0.70</b>		<b>0.61</b>	
CTR <sup>c</sup>	<b>0.33</b>		<b>0.43</b>		<b>0.63</b>	

notes

*a: food plus imputed costs of home cooking*

*b: exogenous cost drivers removed*

*c: cost transfer ratio (delivery costs as a ratio of the value of the transfer)*

42. Interpretation is as follows:

- The upper part of the table presents the unit cost of school feeding (school meals only) per pupil per year. Column (1) gives the cost benchmark for all WFP school-meal-only interventions per pupil per year, together with a breakdown along the usual lines of WFP cost accounting. The actual dollar costs in the country are presented in column (5). We assume that this information was obtained by the consultants by studying project documents and interviews with CO staff.
- At \$66.8, the actual unit cost is \$26.8 (or 67%) higher than the \$40.0 benchmark. It would thus appear that this particular school feeding programme is managed very poorly. But this comparison is misleading. Column (3) provides a corrected set of

benchmarks where exogenous factors have been filtered out in the way described above. This leads to a considerably higher new benchmark of \$55.8.

- Even so, the unit costs in our country are still \$11.0 (or 20%) above the corrected WFP average. The commodity cost is not to blame, so we must look into the non-commodity costs. The cost breakdown allows to identify the major culprit in this case: ODOC. Equipped with this information, the evaluation team can start discussing with the CO, government officials and partner organisations to understand what is the reason for this higher than average cost, and to assess whether it is justified for some reason that the average comparison may have overlooked.

43. Note that the correction proposed in column (3) is at the moment of writing hypothetical, because no such corrections are being made available by WFP Rome. It is thus unlikely that consultants will have this additional information. The numerical example is meant as a warning: it illustrates how misleading one-size-fits-all benchmarking may be.

### ***F. Link with related WFP practice - alpha values***

#### *Calculation of alpha values*

44. In the past, WFP promoted the use of “alpha values”, but this tool has fallen out of favour. Alpha values are the local market value of a food-in-kind transfer, expressed in dollars, divided by the total dollar cost to WFP. The concept can be illustrated with the previous example:

- The total dollar cost per pupil per year is presented in the upper part of Table 3 above. In the lower part of the table the value of the intervention is estimated at market value. This is the value in the local market of the food that is prepared for school meals. The imputed cost of home cooking must be added, because that is what parents also save when their children get a meal at school.
- Note that the cost of preparing the food, usually not borne by WFP but by the schools, should also be included in the cost side. We suppose that this is the case here. The evaluators, after talking to a few schools, have estimated the annual value of the food plus the saving on cooking costs at \$41 per year.
- A benchmark value for all school-meal-only programmes of WFP is also available (\$30), and a corrected market value (\$39) has been calculated by WFP Rome using the same method of regression analysis discussed above. This allows to calculate the uncorrected alpha benchmark for school-meal-only interventions at WFP (\$0.75), a similar corrected alpha benchmark (\$0.70), and the actual alpha value for the intervention in the country (\$0.61).
- As alpha values express \$ benefits per 1\$ of total costs (transfer value plus administrative costs), the higher the alpha value the better.
- The empirical results confirm our earlier findings: the costs of the intervention do not seem to have been kept under control very well. The cost overrun, compared to the whole population of similar projects at WFP, is most striking when an uncorrected benchmark is used, but this is not very meaningful. When the corrected benchmark is used the intervention does better, but remains more costly than the benchmark.

*Interpretation of alpha values*

45. There is some confusion in the literature over the actual interpretation to be given to alpha values. One possible interpretation is to consider the alpha value as a BCR. First, both denominator and numerator are expressed in the same monetary units (either \$ or local currency). Second, the numerator can indeed be regarded as a measure of the benefit to beneficiaries. For it expresses what, in the absence of the food aid package, they would have paid for the same food, or what they would obtain if they sold the food. Third, the denominator expresses the costs of the food package, including both the transfer itself (the food) and other costs of bringing it to its beneficiaries. As both numerator and denominator are expressed in the same monetary units, the ratio can indeed be interpreted as a simple BCR at output level.

46. Seen from this angle, alpha values may help decide when to buy locally rather than import food, or even stay out altogether. It is however problematic to interpret alpha values in this fashion, because observed market prices are poorly suited for estimating the benefits at impact level:

- First, the use of market prices assumes that current food prices are not affected by the intervention. But this is often an untenable assumption. If we are for instance comparing a food-in-kind with a cash transfer, we must take into consideration that in tiny and poorly integrated markets a cash transfer programme may cause a considerable jump in the local price of food.
- Second, even the changed market price is a good approximation of the value of a commodity only for people who, in the absence of the intervention, have purchasing power and access to the market. In WFP emergency interventions, these conditions are typically not satisfied, and the benefit of providing one unit of food aid is higher than its market value.
- Third, and importantly, in all interventions, whether humanitarian or developmental, market prices are a poor guide for another reason: the underlying merit good argument. Explicitly or implicitly, food is treated in WFP interventions as what economists call a “merit good”. In such a case consumer preferences are discarded in favour of paternalistic standards set by nutritional experts. The nutritional content of food is something that beneficiaries may not be aware of or may not care about, but that WFP judges to be very important.<sup>7</sup>

47. For all these reasons alpha values should not be interpreted as belonging to full CBA, but be treated as an application of the more narrow tool of unit cost benchmarking.

---

<sup>7</sup> This is also why WFP is promoting omega values as a replacement of alpha values; on omega values see Section IV-F below.

*Cost-transfer ratio*

48. A variant of the alpha value, the *cost-transfer ratio (CTR)*, has gained some currency in the recent donor literature. It is also shown in the lower part of Table 3 above. It is calculated as  $(\frac{1}{\alpha} - 1)$ , the ratio of administrative costs to transfers, or the administrative cost of making a 1\$ transfer to a beneficiary. In our country it is \$0.63, against only \$0.43 for WFP as a whole (corrected benchmark). The CTR can be used for all food assistance interventions as a tool of unit cost benchmarking.

## IV. Tool Guide 2: Cost-Effectiveness Analysis

### A. Overview

#### *Introduction*

49. The purpose of this guide is to show:

- which aspects of efficiency can be assessed by cost-effectiveness analysis (CEA);
- what the data requirements are;
- what calculations must be performed; and
- how the results can be interpreted.

#### *Why is this tool important?*

50. Food assistance comprises a range of modalities and instruments, and WFP aims to select the most efficient individual interventions and country portfolios on the basis of a careful comparison of the advantages (benefits) and disadvantages (costs) of the possible courses of action. CEA is well suited for making informed choices between mutually exclusive alternatives, such as food, cash and vouchers. External evaluators commissioned by OEV may be invited to use this tool to verify whether the calculations that were the basis for the initial decision (planned costs and results) are confirmed by the data that are available after implementation (actuals). When performed after implementation, CEA assesses both the development efficiency of the interventions (“has WFP done the right thing?”) and the management efficiency of WFP and its partners (“Have WFP and its partners done it the right way?”).

### B. Description

51. Cost-effectiveness analysis (CEA) is a technique that calculates and compares the cost per unit of results, called cost-effectiveness ratio (CER), across alternatives. All the relevant costs incurred to achieve a particular result must be included. The results can be situated at four stages in the logical framework: activities (food delivered to a central warehouse near the harbour), output (food delivered to beneficiaries), outcome (improved nutritional status of beneficiaries), and impact (lives saved). The higher up in the logical framework, the more meaningful the measure of efficiency becomes: the cost per MT imported into the country and warehoused tells us less about the development contribution than the improved nutritional status of beneficiaries, which in turn tells us less than the cost per life saved.

52. The most powerful form of CEA uses quality-adjusted life years (QALYs) as a results indicator at impact level. QALYs capture both the quantitative and qualitative aspects of gains in life years. This version of CEA is also called cost-utility analysis (CUA) and is briefly described in Annex B. QALYs are difficult to estimate, and the application of CUA is beyond what can be expected of external evaluators working for OEV.

### C. Efficiency Analysis

#### *Efficiency dimensions captured*

53. CEA identifies the most efficient among alternatives. This tool can help WFP move from food aid to food assistance by informing the choice between food aid, cash and vouchers. It can also be used in other circumstances where alternatives are being compared, such as different modalities for school feeding.

54. A strong point of CEA is that it is potentially much better at addressing equity issues than full cost-benefit analysis (CBA) from which it is derived. In CBA benefits are expressed in money terms. This starts from market information, and this gives it a bias against the poor: benefits as measured by what people are willing to pay or the cost of income forgone are in general lower for the poor. In CEA benefits are expressed in a non-monetary metric, which can be chosen so as to be equity sensitive. To give an example: the cost per ration per beneficiary is not sensitive to equity, for the beneficiary can be poor or not so poor. Cost per improvement in the food security score, by contrast, is very sensitive to equity. When food assistance is provided to people who are already well fed, they will have an acceptable score both before and after the intervention. and so feeding them does not improve the effectiveness score.

*Efficiency dimensions not captured*

55. *CEA does not tell us whether an intervention is worth doing, only how it ranks in comparison with other interventions.* In this sense it is a less complete efficiency tool than cost-benefit analysis (CBA).

56. It is difficult to imagine a single non-monetary metric that simultaneously expresses nutritional, health (MCHN), educational (school feeding), and income (P4P) benefits. CEA is therefore most suitable when the focus is on alternatives that all pursue a common single objective – for example when the three modalities of food, cash and vouchers are mutually exclusive alternatives that focus mostly on nutritional improvements.

57. Even if we stick to WFP interventions that achieve their development impact through nutritional improvements, CEA faces some limitations. It is, for instance, unsuited to compare interventions that pursue humanitarian relief with those that pursue longer term development. The reason is that in humanitarian interventions timeliness is of the greatest importance, and this dimension is not captured by indicators of nutritional improvement. In this case CEA offers a solution (in principle), as the QALY measure is influenced by both nutritional intake and timeliness of an intervention.

***D. How to perform CEA<sup>8</sup>***

*Data requirements*

58. The costs that enter into the numerator of the CER include the whole gamut, from procurement to international and national transport, storage, distribution, technical assistance, contracts for supporting services, to staff and equipment at country and regional office and headquarters. Much of this data is kept somewhere in the financial accounting and reporting systems of WFP.

59. On the other hand, the results that enter into the denominator of the CER calculations have to be expressed in quantitative but non-monetary terms. Examples include number of beneficiaries reached, MT of food delivered, nutritional value units consumed, weight-for-height and height-for-age measures of malnutrition, and numbers of deaths avoided.

---

<sup>8</sup> An example of how CEA could appear in an evaluation matrix is provided in Section V, subsection D below.

60. Care must be taken that indicators are correctly measured. The following are illustrations, and evaluators must take sufficient time to apply the same principles in their work. When cost components are expressed in more than one currency, an appropriate exchange rate has to be used, either the dollar or the local currency. Numbers of beneficiaries should be standardized for the level of benefits obtained per person, for instance by converting to standard rations per person. And all data must be collected for the same time period. Note also that most cost data will be obtained from WFP sources, but costs incurred by national government and beneficiaries can be important and should be included in the analysis whenever possible.

61. On the results side, indicators express aspects of effectiveness. They measure the difference between the situations with and without the intervention. There is a considerable literature on how to measure with/without differences rigorously. The problem is that the without situation cannot be observed because the intervention has taken place, and the situation before the intervention is not always a good guide. At the lower levels in the logical framework this is often not difficult. For instance, the tonnage of food that the project has imported will be an appropriate with/without indicator. But at higher levels measurement is increasingly complex. In such cases a counterfactual has to be carefully established to estimate what would have happened in the absence of the intervention. At impact level, the gold standard for estimating the counterfactual is to randomly assign a sufficient number of potential beneficiaries to an exposed and a control group, taking into account relevant characteristics such as age, gender, education, wealth, and income. If done properly, differences in averages between the exposed and control groups will express, apart from an unbiased statistical error term, the effect of the intervention. In reality, such randomised experiments are seldom conducted, and most with/without comparisons in WFP evaluations rely on informal and casual estimates. This makes it sometimes hard to convincingly claim that the results measured are attributable to the intervention only. Evaluators should be attentive to this methodological question, as the with/without comparison is crucial in performing CEA. In the absence of a robust methodology, they should argue why they think that a particular counterfactual seems to them probable.

#### *Calculations*

62. The calculation takes the form of a simple division, with the costs in the numerator and the benefits in the denominator. The fact that costs and benefits are each expressed in different units does not pose a problem, as long as the data in the numerator (costs in dollars or local currencies) and the denominator (non-monetary metric of results) are expressed in the same unit across all the interventions being compared. If this condition is satisfied, the CER of all the alternatives can be compared.

63. It is sometimes argued that both the data in the numerator and in the denominator should be discounted through time, to take account of the fact that the resources invested in the intervention have an opportunity cost (numerator) and that the value per unit of benefit decreases with rising income levels through time (denominator). The selection of appropriate discount rates however, especially of the non-monetary benefits in the denominator, is complex, and the use of different discount rates in numerator and denominator is controversial. No discounting need be applied in CEA calculations performed for OEV evaluations.

*Interpretation*

64. The decision rule for CEA is a simple ranking, with lower CERs indicating higher levels of efficiency. The alternative with the lowest CER has the highest efficiency. There are however a number of limitations the evaluator should take into account:

- (a) A first caveat is that the scale of interventions is neutralised when benefits are calculated per unit cost. This limitation should not normally cause major problems in WFP evaluations, because food assistance interventions are typically scalable to the same level across alternatives. But exceptions may occur. Imagine a voucher programme that is shown by CEA to be more efficient than either cash or food alternatives, but that can only be applied in urban areas with large commercial food outlets, and for this reason reaches only a fraction of the beneficiaries of the food and cash alternatives. It may then be justified to abandon the CEA ranking rule and select the second ranked alternative, say food, because in this way many more beneficiaries are being reached and the total benefits are higher. The problem can be overcome if a mixed modality (vouchers in larger urban areas combined with food elsewhere) is introduced as an additional alternative.
- (b) A second caveat relates to the fact that CEA can be practised at different levels of the logical framework. This flexibility is in itself one of the advantages of CEA. There is however a trade-off between the ease of calculation and the reliability of the result. Indeed, CEA calculations at activities or output level are typically easier than at outcome and impact level, but are also typically less convincing as indicators of “Value for Money” from a development perspective. More formally, lower level CERs are good efficiency indicators only if no ranking reversal occurs when the calculation is repeated at higher levels of the logical framework, but this cannot be taken for granted –see the illustration in Box 2 below. Evaluators who make extrapolations from lower levels of CEA should therefore provide additional, qualitative arguments why in their opinion no such reversal is likely to occur in the case under study.
- (c) A third caveat is that CEA has been conceived for the comparison of interventions that all have the same single objective, e.g. nutritional improvements. For this reason CEA is more appropriate for EMOP than for multiple-objective WFP interventions such as school feeding, where both nutrition and school feeding are pursued, or Food for Assets (FFA) programmes, where both nutrition and asset creation are pursued, or P4P, where both nutrition and farmer incomes are pursued. For this reason CEA is not suited for evaluating country portfolios. A typical portfolio is composed of several interventions (some combination/sequence of EMOPs, PRROs, DEVs) that cannot be reduced to one single objective, and thus cannot be aggregated in one single overarching non-monetary metric.

65. To address the complication of multiple objectives, other efficiency tools than CEA have to be used. CUA using QALYs or DALYs is more powerful in this regard than CEA, as it can integrate both health and nutrition objectives. CUA is in turn less powerful than CBA, which uses money as the common yardstick. But both CUA and CBA are much more demanding in data and more complex in terms of methodology, and for this reason are not often within reach for external evaluators working for OEV. Another possibility is the use of Multi-Attribute Decision Models (MADM) a method that is not a substitute for but a complement to the use of CEA. All these alternatives are briefly reviewed in Annex B.

### Box 2 Cost-effectiveness ranking reversal – an illustration

Lower level CERs are good efficiency indicators only if no ranking reversal occurs when the calculation is repeated at higher levels of the logical framework. Yet it is not possible to know for certain whether such reversals will occur until we do the actual calculation.

Consider two alternative interventions A and B. A is cheaper to deliver per unit per beneficiary. In other words, at output level it has the lower CER. The preferred impact indicator in this case is cost per life saved. On what basis can we claim that modality A will also have a lower cost per life saved than modality B? A reversal can occur for several reasons: for example, if modality B is more timely, if it is better at targeting beneficiaries, if it leads to less sharing with non-targeted family members, or if it leads to the consumption of more nutritious food.

Evaluators who make extrapolations from lower levels of CEA should therefore provide additional, qualitative arguments why in their opinion no such reversal is likely to occur in the case under study.

#### E. Example

66. The data in the following example are taken from a WFP document on the choice between food, cash, and vouchers modalities (Ryckembusch 2012). The information draws on household surveys, and measures nutritional intake, an outcome measure.

**Table 4 Alternative outcome rankings of three food modalities**

Treatment	$CER_{\text{outcome, kcal}}$	$CER_{\text{outcome, kcal}}$ ranking	nutritional value score (NVS)	$CER_{\text{outcome, NVS}}$	$CER_{\text{outcome, NVS}}$ ranking
	(1)	(2)	(3)	(4)	(5)
Cash	1317	1	6.15	214	3
voucher	1450	2	7.70	188	1
in-kind	1717	3	8.70	197	2

Where:

$CER_{\text{outcome, kcal}}$  : cost in local currency of providing 1300 kcal/person/day.

$CER_{\text{outcome, NVS}}$  : cost in local currency of providing one unit of NVS/person/day.

67. The interpretation is as follows:

- Column (1) provides information in the form of a CER: the monthly cost (in local currency) of inducing the consumption of an additional daily ration of 1300 kcal/person under each of the three modalities. Column (2) provides the ranking of the three modalities in terms of CEA. Cash is the clear winner.

- The WFP study calculated not only the energetic but also the nutritional value of the rations that beneficiaries consume under the three modalities. Column (3) provides the corresponding nutritional value scores (NVS) per 1300 kcal/person/day. In column (4) the corresponding CERs, expressing the cost in local currency per nutritional unit consumed, are presented.
- The ranking, presented in column (5), is now different: cash is the least attractive alternative, vouchers the most attractive, with food-in-kind in the middle position. WFP favours nutrients (NVS) over energy (kcal) as the approach to measure nutritional outcomes. Vouchers are thus the preferred option.

68. The evaluators must still argue why no rank reversal is likely to occur if we move to the level of impact efficiency, where ideally we would like to perform our CEA.

### ***F. Link with related WFP practice – omega values***

69. The Cash and Voucher Division of WFP promotes what are labelled “omega values” to perform the efficiency comparison of food aid modalities. An analysis of this approach has been included in a special issue of World Development on “Impacts of Innovative Food Assistance Instruments” (Ryckembusch et al. 2013). Omega values are obtained by the division of two CERs at outcome level. This can be illustrated in Table 4 above:

- If we wish to compare cash and in-kind transfer modalities in that example, the two CERs in column (4) can be divided to obtain  $214/197=1.08$ . Similarly, if we wish to compare vouchers and in-kind transfer modalities, we obtain  $188/197=0.95$ .
- The interpretation of omega values is as follows: if the ratio exceeds unity the modality in the denominator (food-in-kind) is the better option, if less than unity it is the modality in the numerator (in the first calculation cash, in the second vouchers).
- Using omega values leads to the same conclusion as a direct comparison of CERs, but the interpretation is less intuitive. Note also that omega values can only be compared two at the time, whereas any number of CERs can be ranked at the same time.

## V. Tool Guide 3: Specific Evaluation Questions

### A. Overview

70. Specific evaluation questions on efficiency (SEQs) are identified in the BMZ study (Palenberg 2011) as a potential method/tool, but one which isn't anywhere very systematically set out. Importantly, though, this is distinguished from mere "expert judgement", since SEQs should be precisely defined in a way that ensures different evaluators will arrive at similar judgements. This is important from the point of view of transparency and credibility of the judgements, and also to enable comparisons across similar evaluands.

71. The BMZ treatment of SEQs focuses on systematic theory-based questions that can be used to test the validity of the intervention logic from an efficiency perspective. However, it is possible to take a broader view and consider good practice in the selection of efficiency-related evaluation questions that could, in principle, bring in elements of other identified tools (such as unit cost benchmarking or "follow-the-money") as appropriate. The point is that the careful formulation of EQs should help to focus evaluations where efficiency analysis can be most useful, and ensure a consistent approach that enhances the likelihood of drawing valid conclusions across suites of evaluations.

72. In this approach, evaluation managers (EMs) have a pivotal role, since the key effort in developing SEQs is at the TOR stage. OEV needs to decide on the relative importance of efficiency issues in particular types of evaluation, while individual EMs need to ensure that such issues are then appropriately reflected in specific TORs and subsequent evaluation matrixes. In turn, the same concepts and criteria need to be applied in reviewing draft evaluation reports.

### B. SEQs Description

#### *Narrow approach*

73. The BMZ survey envisages an SEQ approach that is carefully linked to theory-based evaluation. The intervention logic (reconstructed if necessary by the evaluators) is examined in order to understand at each level (activities, output, outcome, impact) the expected links between inputs and results. What are the key assumptions that affect these links? Correspondingly, what evidence would indicate that a link is not working at all, or is working inefficiently (less-than-expected results per unit of input)? Relevant evidence could be quantitative (most likely drawn from project records of inputs and results) or qualitative (e.g. interviews with staff, observers and beneficiaries) which provide evidence on the input chain (wastage in the delivery chain, pipeline breaks, whether food is eaten by the right people – issues of taste preferences and targeting etc). Preliminary examination and prima facie evidence at the TOR stage might suggest particular links in the intervention logic that merit special examination. These might be particular sub-activities (FBF production, or the quality of FFA activities), or specific elements of the implementation chain for all activities (e.g. food procurement and handling).

### *Cross-cutting approach*

74. Most likely, once EMs start considering efficiency issues systematically in this way, they will also identify a variety of tools that might be employed to address them. Accordingly, the formulation of SEQs for efficiency is presented here as a cross-cutting approach, since, whatever the tools proposed, the same set of good practice principles should apply to the formulation of TOR and EQs.

### **C. How to develop SEQs on Efficiency**

75. The DAC guidance (see Box 1 above) offers some illustrative “useful questions”:

- Were activities cost-efficient?
- Were objectives achieved on time?
- Was the programme or project implemented in the most efficient way compared to alternatives?

76. However, these questions are very general. For real evaluations the EQs should be sharpened as follows:

- Distinguish clearly between efficiency and effectiveness.
- Ensure a proper sequence between effectiveness and efficiency analysis.
- Decide at which levels of the logical framework efficiency analysis is required.
- Prioritise – both between efficiency and other issues, and amongst possible efficiency questions.
- Identify appropriate techniques of efficiency analysis and reflect them in the evaluation design.

### *Distinguishing efficiency and effectiveness*

77. A glance at past WFP evaluations – and those of many other agencies – throws up numerous examples where “efficiency-and-effectiveness” is treated as a single term and/or findings appear under the wrong heading. This is unfortunate, not least because there are often important real-life trade-offs in the simultaneous pursuit of efficiency and effectiveness; efforts to constrain costs may undermine effectiveness. Here is a good example, which uses the terms correctly:

Finally, there are examples where the tradeoffs between efficiency gains and effectiveness issues have worked poorly. In its efforts to secure best value for money in PAT [programme assistance team] recruitment for instance, WFP selected teams on the basis of the cheapest tenders. The lowest price did not always correspond to capacity to carry out the work, and WFP is aware that the price paid in declining effectiveness has outweighed the cost savings incurred. Similarly, the decision to purchase HEB [high energy biscuits] from an Indian supplier for the school feeding programme encountered major difficulties in quality and supply. (From Afghanistan CPE 2012)

78. Issues of equity and timeliness are tricky: they are not exclusively about efficiency (cf. ¶10 in Section I), but they are very likely to have an important efficiency dimension and may influence the attention to efficiency at TOR stage. *Prima facie* indicators of inefficiency may include wastage or spoilage of food, pipeline breaks, and issues about inclusion or exclusion errors in targeting. Whether interventions reach the right people is a question of

effectiveness; maximising outputs and outcomes for a given level of resources is a question of efficiency.

#### *Sequencing analysis of effectiveness and efficiency*

79. Efficiency cannot be assessed if results (effectiveness) are not known, so make sure that the evaluation pays proper prior attention to results and then links its analysis of efficiency to the effectiveness findings. Refer to Table 1 above for an extensive list of relevant effectiveness indicators and efficiency indicators at each level of the logical framework (reference to this table may also help to avoid confusion between efficiency and effectiveness questions).

#### *Levels of ambition*

80. In turn, this requires clarity about which logical framework levels (activities, outputs, outcomes, impact) require efficiency analysis. As already noted, the analytical and data requirements for efficiency analysis are less at lower levels of the logical framework, but the findings may be correspondingly less powerful. A more ambitious evaluation will be more demanding in terms of data and analytical resources. Again Table 1 above provides a checklist of the results that may be sought at each level of the logical framework, and the corresponding effectiveness and efficiency indicators that might be relevant.

#### *Prioritise*

81. The investigation of efficiency questions will have a direct cost (in terms of evaluation team time and effort) and an opportunity cost (other questions they could have focused on instead). As well as the balance between efficiency and other topics, decide the scope of efficiency analysis to be used. For example, the evaluators may be asked to consider efficiency in implementation of a given instrument (e.g. looking at unit costs of school feeding), or comparison between alternative instruments (e.g. the efficiency of food assistance vs. cash or vouchers).

#### *Specifying efficiency analysis techniques*

82. Be clear about which techniques of efficiency analysis are proposed. This in turn requires judgments about:

- which available data ought to be deployed (and hence gathered at the preparatory stage of the evaluation);
- the level of effort to be allocated to efficiency analysis (vs. alternative uses of evaluation time);
- the skills to be represented on the evaluation team (particularly if the use of more technical tools is envisaged, but also if there is to be special attention to logistics, etc).

### ***D. Examples of SEQs on Efficiency***

83. In an extensive scan of recent WFP CPEs and OpEvs, efficiency was usually conspicuous by its absence from TOR (except in the obligatory reference to the 5 DAC evaluation criteria). Systematic attention to efficiency in evaluation reports themselves was also rare, although in some cases, sections or annexes on logistics were a partial exception. It is likely that this will change: increased emphasis on “value for money” will entail more attention to all dimensions of efficiency in future.

84. There is a similar scarcity of efficiency-specific questions in Evaluation Matrixes. However, the Laos CPE (WFP 2009) provides one good example, reproduced below.<sup>9</sup>

Evaluation Questions	Means of Verification Indicator(s)	Source of Verification and Data Collection Strategy of the Evaluation Team
<b>Q3: How has the portfolio performance and what have been its results?</b>		
How efficient has the programme delivery been (for FFR, FFW, school feeding, nutritional education programmes)?	Least cost method Comparators' cost Cost of other options	Try to establish the unit cost (e.g. the cost per child fed) and compare it with <ul style="list-style-type: none"> <li>• overall sector expenditure (e.g. amount of money government spends per child on basic education)</li> <li>• cost of programmes run by others, if any</li> <li>• alternative options (type of food delivered, cash or vouchers if markets exist and function)</li> <li>• alternative implementation strategies and/or partners</li> </ul> Cost information from WFP Country Office, and other information through discussion with partners.

85. Below is a draft EQ to show how cost-effectiveness analysis (in this case applied at outcome level) might appear in an evaluation matrix.

Evaluation Question	Means of Verification Indicator(s)	Source of Verification and Data Collection Strategy of the Evaluation Team
How cost-effective is the choice of the selected food assistance modality (commodity-based vouchers) compared to the other alternatives considered (food-in-kind, cash or value-based voucher alternatives)?	Cost-effectiveness analysis, using nutritional; value score (NVS) as the metric of output	Go back to the calculations (cost/NVS) that were made at appraisal to rank the alternatives and verify whether the numerical data still stand. <ul style="list-style-type: none"> <li>• For the alternatives not chosen, stick with the original data and calculations, except when there is evidence that some of the prices were wrongly estimated (source: project documents; CO)</li> <li>• For the alternative chosen, verify the different cost components used during appraisal against actuals (for instance the cost of the commodities, the overhead charged by the retail companies, or the cost to beneficiaries if fewer shops participate in the scheme and they have to travel a longer distance to get the food parcels) (source: CO)</li> <li>• For the alternative chosen, test for weaknesses in the implementation (a change in the composition of the rations, a mismatch with consumer preferences so that certain items are not actually consumed), that will alter the nutritional value (source: CO)</li> <li>• Rerun the calculations and rank (or apply omega values pair-wise)</li> </ul>

<sup>9</sup> Evaluators may choose to adopt more elaborate matrix formats, but this simple version highlights the main analytical and data collection issues well, and so it is echoed in the further examples of EQs below.

86. The final example below is a draft EQ to show how unit cost analysis might appear in an evaluation matrix. Ideally it would be possible to be more specific about which existing WFP metrics and data series could be used for benchmarking, and it is expected that ongoing improvements to WFP budget and data systems will allow more systematic cost analysis and benchmarking in future.

Evaluation Questions	Means of Verification Indicator(s)	Source of Verification and Data Collection Strategy of the Evaluation Team
Are the costs for the schooling programme kept under control?	Unit cost analysis	<ul style="list-style-type: none"> <li>• Calculate the total cost of the school feeding, broken down in costs borne by WFP, government and beneficiaries. The costs of WFP should be further broken down by category (commodity, transport and warehousing, etc) (source: project documents and reports, CO).</li> <li>• Use an appropriate metric to calculate the amount of school feeding delivered (e.g. number of meals, or number of children/year who received the full foreseen school meals), so that you obtain an accurate number that allows comparison with some benchmark. (source: project documents, interviews with staff involved, field visits as a check).</li> <li>• Compute unit costs, and compare with benchmark (source WFP benchmarks: WFP Rome or regional bureau).</li> </ul>

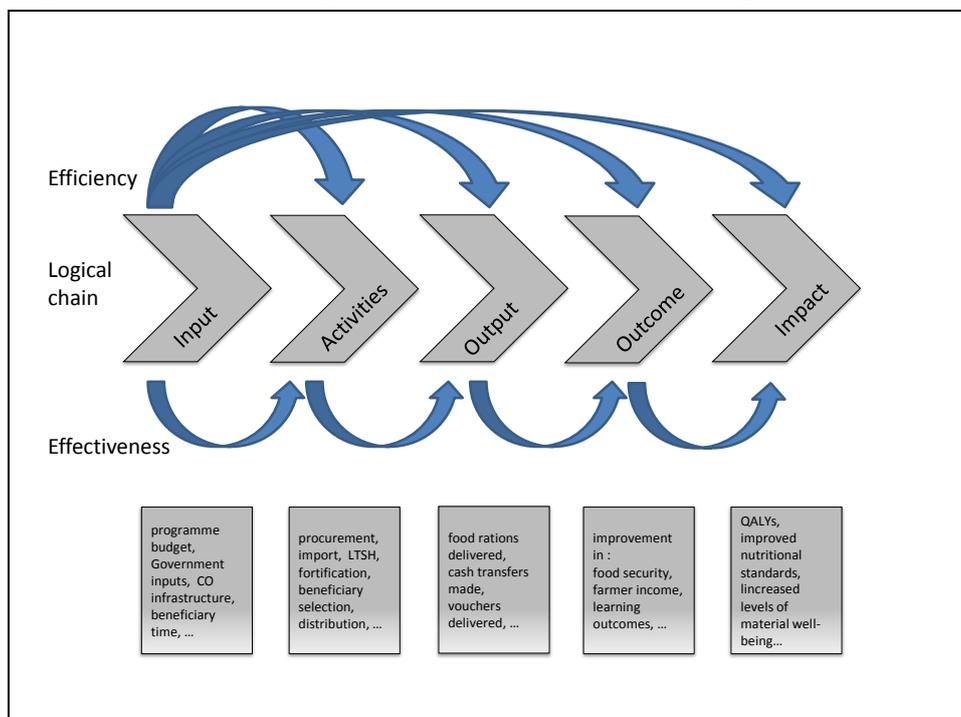
## Annex A Non-Standard Definitions of Efficiency

A1. This annex (a) shows how the “three Es” and similar variants of efficiency terminology map on to the broad DAC definition used in this Technical Note, and (b) provides a fuller explanation of why the broad definition is the right one for WFP/OEV.

### *The relationship between broad and narrow definitions*

A2. The figure showing how efficiency and effectiveness both relate to the logical framework is reproduced below for convenience.

**Figure A1 Efficiency and effectiveness in the logical framework**

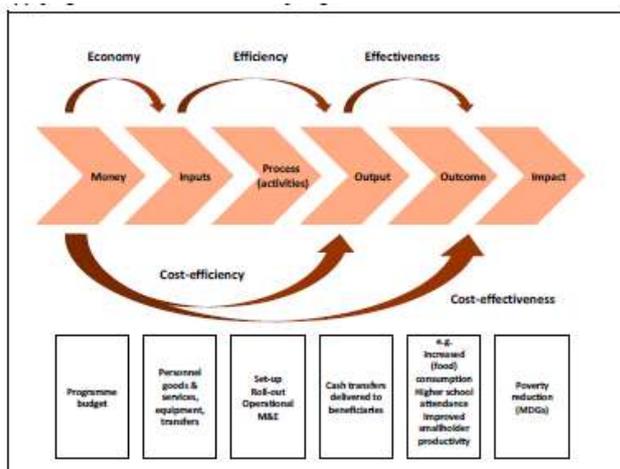


A3. Some donors have adopted narrower definitions, and the DAC itself also uses a different definition of efficiency, that for instance appears in the on-line summary of Key Norms and Standards: “Efficiency measures the outputs – qualitative and quantitative – in relation to the inputs” (DAC 2010b, echoed in Box 1 at the beginning of this Note). This narrow definition is adopted by ALNAP (ALNAP 2006:44) and many donors.

A4. DFID, in particular, has propagated the “three Es” approach, and has linked to it the increasingly prominent concept of Value for Money – as illustrated in Figure A2 and Figure A3 below. However, the various terms in this approach can easily be understood in relation to the logical framework levels of efficiency depicted in Figure A1 above. Thus:

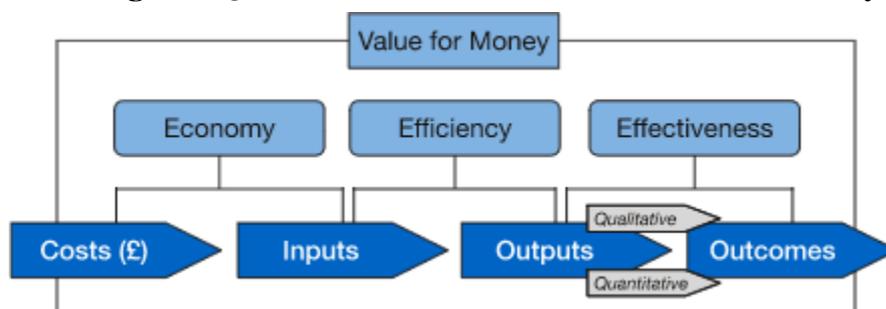
<i>Standard terminology</i>	<i>“3 Es” terminology</i>
activity level efficiency	“economy” in figures A2 and A3
output efficiency	“efficiency” in figures A2 and A3
outcome efficiency	“cost-effectiveness” in figure A2
impact efficiency	“value for money” in figure A3

**Figure A2 The “three Es” as interpreted by DFID**



Source: Hodgson et al 2011.

**Figure A3 The “three Es” and Value for Money**



Source: Barnett et al. (2010: 6)

*Why the broad definition is preferable*

A5. This mapping shows that the broad definition is actually rather simple. There are additional reasons to prefer it:

- (a) It is consistent with the economics literature on efficiency (see Box A1).
- (b) Conversely, the definition that limits efficiency strictly to a relationship between inputs and outputs suggests that the relation between inputs and outcomes and the relation of inputs and impact is not in the realm of efficiency analysis. The narrow definition fails to acknowledge the simple but important fact that the notion of efficiency applies to all stages of the logical framework beyond input, i.e. output, outcome and impact, rather than just to output, and that there is some inherent hierarchy between these stages, with efficiency assessment at higher stages being by definition preferable.
- (c) Furthermore, the narrow definition obfuscates the distinction in evaluation theory between two questions: what is the effect of the intervention at a certain stage in the logical framework (effectiveness), and could we have achieved the results at lower cost by choosing another alternative (efficiency).

### Box A1 The efficiency concept in economics

The literature on welfare economics and cost-benefit analysis is concerned with decision making in the public sector. In the field of development aid, the stated objectives of public donors are broad-ranging, from saving lives in emergencies, to constructing social safety nets for the poor, enhancing pro-poor economic growth, empowering women, and enhancing personal and political freedom.

The tools developed by economists are more limited, and useful only in those cases where the inputs can be expressed in monetary terms, and the results either in monetary terms (cost-benefit analysis) or in some non-monetary metric (cost-effectiveness analysis). However, this efficiency toolbox can be applied both ex ante and ex post:

- When used ex ante, it is quite sensibly assumed that the intervention will be decently managed, and the analysis concentrates on the strategic part of decisions making. This means that the underlying efficiency definition used in ex ante analysis is the one that relates inputs to impact.
- If used ex post, the calculations will not only reveal whether the right strategic decisions were made, but also whether the intervention was well managed. It will thus also address efficiency at lower echelons in the logical framework: inputs to activities, inputs to outputs, inputs to outcome, as well as inputs to impact.

Given that the economics toolbox is designed for both ex ante and ex post evaluation, the broad DAC definition is appropriate.

A6. Moreover, attention to the broad definition makes it clear that “*value for money*” is properly understood as *efficiency at the impact level*. Indeed, this is consistent with the approach to Value for Money that is being adopted by WFP – see Box A2. Note that the Executive Board review of efficiency at WFP (WFP 2011) explicitly used the broad DAC definition of efficiency, and this is echoed in the efficiency definition presented in Box A2.

### Box A2 Efficiency and Value for Money<sup>10</sup>

WFP has adopted the OECD-DAC definition of Value for Money as the umbrella term to summarize achieving an optimal balance between Economy, Efficiency and Effectiveness, defined as:

- **Economy:** Absence of waste for a given output. (An activity can be considered economical when the costs of the scarce resources used approximate the minimum needed to achieve planned objectives.)
- **Efficiency:** A measure of how economically resources/inputs (funds, expertise, time, etc.) are converted to results.
- **Effectiveness:** The extent to which WFP intervention objectives were achieved, or are expected to be achieved, taking into account their relative importance.

<sup>10</sup> Based on a 2013 EMG discussion paper on Value for Money at WFP.

## **Annex B Additional Efficiency Tools**

This annex provides a brief guide to 5 further tools that may be of (more limited) use to OEV. A fuller discussion of their strengths and weaknesses is available in Renard 2013. See also the comparison of characteristics in Section II, Table 2 of this Technical Note.

The tools described are:

- Cost Benefit Analysis (CBA)
- Cost-Utility Analysis (CUA)
- Multi-Attribute Decision Model (MADM)
- Follow the Money
- Benchmarking Other Indicators

## **Cost Benefit Analysis (CBA)**

### *Features of CBA*

1. CBA is an advanced technique from economics, designed to measure all the inputs (costs) and results (benefits) of an intervention, using a common yardstick – money. Once all costs and benefits have been correctly estimated, they are aggregated into a measure of project worth, such as Net Present Value (NPV), Benefit-Cost Ratio (BCR), or Internal Rate of Return (IRR). This in turn allows to establish whether a project is worth doing, and to rank alternative projects. CBA been conceived for application at the highest level in the logical framework, i.e. impact.

### *CBA requirements*

2. In spite of the intellectual sophistication of the instrument that was developed, and despite the efforts of many countries and international institutions such as the World Bank, UNIDO and OECD, to promote its use, CBA has never become widespread in developing countries, except by researchers and academics, and since its heyday in the 1970s its use has declined, even at the World Bank.

3. CBA's limited use in practice reflects its heavy theoretical and data requirements. Advanced economics skills are required to perform CBA In terms of data needs, CBA is the most demanding of all the tools discussed. Like all efficiency analysis, it can only be applied if the corresponding results from effectiveness analysis are available. And even if we know the results, putting money values on them remains a complex task.

### *CBA applicability for WFP*

4. Almost all WFP interventions have an important nutrition component that at the level of impact translates into indicators such as lives saved or improved quality of life. This severely limits the application of CBA because of the difficulty of putting a monetary value on human life. It does not help to go down in the logical framework and apply CBA at outcome or output level, for the problem does not go away. What for instance is the monetary value of an increased nutritional intake? There seems to be no meaningful way to asses this without linking up with the effect on the value of human life.

5. However, if we go even further down in the logical framework, CBA re-emerges as a distinct possibility. CBA can indeed be applied to good effect at the level of activities, such as investment in warehouses, locally produced fortified food, or emergency road improvements, that can be regarded as sub-projects. In these cases, benefits typically consist of foreign exchange or fuel savings that can be more readily measured in money terms. The whole issue of valuing human life then disappears from view. Such CBA is not only far less complex, but also not controversial.

6. As for the role of evaluation teams contracted by OEV to assess CPEs and OpEvs, it is not realistic to expect that they will be able to perform such analysis. What they may do, is verify whether appropriate efforts have been made by the CO to make use of the tool in selecting interventions and modalities, and where CBA has been applied, to verify whether the recommendations were followed and whether its major assumptions are still valid.

## **Cost-Utility Analysis (CUA)**

### *Features of CUA*

1. Cost-utility analysis can be best understood as a special case of cost-effectiveness analysis. The cost calculations are identical to those for CEA (see section IV of this Technical Note), but it is distinct in two ways: it is situated at impact level only, and it uses a particular metric (Quality-Adjusted Life Years – QALYs) that captures both the effect of an intervention on life expectancy and on the quality of life. (For poorer countries Disability-Adjusted Life Years, a somewhat simplified version of QALY, may be used, but this metric is methodologically controversial.)
2. CUA can be used throughout the intervention cycle. At the ex ante stage, results from theoretical studies and empirical research can be used to generate predictions about what impact a certain intervention is expected to have if well implemented. During mid-term reviews, evaluators can correct for assumptions that have proved too optimistic or pessimistic or for implementation failures. During ex post evaluations, actuals can be substituted for planned results.

### *CUA requirements*

3. On the cost side CUA is comparable to CEA. We need to know and measure in monetary terms all the relevant costs, whether borne by WFP, by the government, by non-governmental partner organisations, or by beneficiaries. And again, cost items must be identified and measured with care and precision. But whereas CEA is often applied at activity or output level, CUA is always at impact level, and thus the analyst must include a wide set of costs. As for the results side, WHO publishes annual data on DALYs, but these cannot be used directly. The analyst must estimate what the impact of a certain WFP intervention is on particular diseases in order to calculate its effect on DALYs. This is work for specialists. The economic skills required are advanced, identical to those for CEA.

### *CUA applicability for WFP*

4. Although WFP is not an agency specialised in health, the large majority of its interventions aim at nutritional outcomes, and these can be captured at impact level in QALYs or DALYs. This means that all WFP operations where the objective is nutrition, health, or a combination of both, can be subjected to CUA. This makes the method suitable to some interventions with multiple objectives, such as MCHN. Note that QALYs and DALYs look at the effect on the whole lifespan of a beneficiary. From this perspective, it will be more efficient to save the lives of younger people than of elderly people. If the moral principle is that all human beings, independent of their life expectancy, have the same right to be helped, then it is better to use as a metric lives saved, which is an application of CEA.
5. CUA is unsuited for evaluating country portfolios because WFP portfolios should not be assessed solely on its nutrition and health aspects, but also on learning outcomes (school feeding), assets created (FFA), or farmer income (P4P). Evaluators can play a role in verifying existing calculations made at the appraisal stage. But when they find that no such study was performed, they do not have the qualifications, time or the resources to do CUA themselves.

## Multi-Attribute Decision Model (MADM)<sup>11</sup>

### *Features of MADM*

1. MADM is a tool for scoring interventions on multiple dimensions. The principle is simple and widely applied (think of the way athletes are compared by aggregating separate scores for the different running, jumping, throwing disciplines of the decathlon). WFP interventions indeed often have several objectives, and they may not all be captured by one single efficiency tool. Apart from food consumption, we may for instance consider beneficiary preferences, equity, gender, and environment effects as important dimensions (the list is illustrative and not exhaustive). If we use the same scoring system (say a scale between 1 and 10), we can then add the scores across all the dimensions. We need not attach equal weight to all the dimensions. For instance, OEV might decide that in a certain country, nutritional value gets a weight of 1, beneficiary preferences a weight of 0.2, equity and gender a weight of 0.3, and environmental effects a weight of 0.4.
2. In the scientific literature MADM is presented as a way to bring dimensions to bear that are important to decision making but are not, or are not sufficiently, taken into account in the more quantitative economic analysis. It has been developed mainly as a decision tool, and therefore its first and most important level of application is during ex ante appraisal. But it can be applied at mid-term and ex post to verify whether the right choice was effectively made and whether it is still valid. The technique is simple in itself, but scoring consistently across interventions and evaluators is not straightforward.

### *MADM requirements*

3. Before applying MADM we must decide the relevant objectives to include and their weights. Insights from the scientific literature on development in general and food assistance in particular, will be helpful, but in the end such choices should reflect the strategic priorities of WFP. OEV therefore would need to establish a scoring system to be applied across a set of comparable evaluations and incorporate use of MADM into the TOR. From here on all depends on whether evaluators can find sufficient information to meaningfully score the different dimensions. If such information is available, the additional efforts to perform MADM are minimal. More probably, evaluators will have some serious work to do to fill in some of the scores. No specific economic skills are involved to perform MADM.

### *MADM applicability for WFP*

4. MADM is very suitable to compare WFP interventions for which partial indicators are available, for instance on nutrition, education and asset creation. The aggregation that MADM provides is especially useful in conjunction with CEA, because the latter can only capture the elements that can be expressed in the one results metric chosen for CEA (e.g. food ration per person per month, change in food consumption score, unit decrease in malnutrition). For this reason this tool could also be used to assign an overall score to country portfolios. But MADM requires WFP to decide on the objectives to be assessed, and also on the scale and the weights to be used by evaluators. At present, no such practice exists. In devising a MADM care must always be taken to avoid double counting so that all the partial indicators should measure different dimensions, without overlapping.

---

<sup>11</sup> The BMZ study (Palenberg 2011) distinguishes between “scientific” and “intuitive” MADM. WFP has judged that the “scientific” version is impractical, it is the “intuitive” version that is described here.

## **Follow the Money**

### *Features of “follow the money”*

1. “Follow the money” is a pragmatic approach, involving the detailed recording of all the steps in the implementation of an intervention. The investigation may reveal waste of resources such as equipment that was bought but never used, or training that was provided but did not have any effect on how trainees do their job, or a supporting activity that proved useless because it was implemented too late. This analysis is applicable at the activities level of the logical framework, up to the output level. Although it stops where the expenditures of the intervention stop, the analysis should help in assessing the likelihood that intended outcomes and impact are achieved. This tool can be useful for interventions that have come to an end, but also during mid-term evaluations.

### *Follow the money – requirements*

2. This tool starts from the detailed description of the sequence of inputs and activities, and how they lead to the outputs. The corresponding expenditures are then tracked. At each turn, the question is asked whether the same results could have been achieved with a lower expenditure, or whether a given expenditure could have produced better results. Following the sequence of implementation, the robustness of the underlying programme theory is also tested. Was this expenditure really useful or necessary in reaching the next level? And if not, what would have been the better alternative?

3. Such a step-by-step investigation requires a detailed description of how an intervention was conceived and then implemented, and detailed financial data about budgets and actual expenditures. Interviews with intervention managers and implementing staff, including those of government and implementing NGOs, will have to be conducted. A participatory approach where beneficiaries are also involved in the assessment, would be a bonus. No advanced economic skills are required. An understanding of the complex logistics of food assistance is probably important.

### *Follow the money – applicability*

4. Applicable at the level of individual WFP interventions and country portfolios. This tool can be used at different levels of detail and complexity. It can be focused on the whole intervention or on some critical section of the logical framework only. Even if not selected as a separate tool, it could be a useful add-on in other approaches. To some extent this tool gives a name to an approach that has already been followed in some elements of WFP evaluations: in particular evaluations of WFP logistics may follow such an approach. Equally, it may overlap with unit cost approaches, since unit costs may be checked during the same process. On the other hand, it should not be confused with WFP’s audit functions.

## **Benchmarking Other Indicators**

### *Features of benchmarking other indicators*

1. Technically, the benchmarking of partial indicators other than unit costs is similar to the benchmarking of unit costs discussed at length in Section III. It is also a quantitative tool involving numerical data. The difference is that in this case the indicators are non-monetary. Some useful indicators are generated as part of the annual CO performance plans, such as share of funding needs secured, share of funding unspent at the end of the intervention, share of projects starting within targeted time, share of operations with full ration size distributed during the reporting period. It will be then up to the evaluators to assess these results against corporate benchmarks and to interpret why shortfalls have occurred. This tool is located at the lower levels of the logical framework (inputs and activities). It can be usefully performed during mid-term and ex post evaluations.

### *Benchmarking other indicators: requirements*

2. Data needs are to some extent already satisfied through the WFP monitoring and reporting system. But the problem is that often there is no reliable and relevant benchmark to compare with, and single observations without a comparison are not very meaningful. It is hardly the task of external evaluators to establish such benchmarks. They may draw on studies by other donors who have worked in similar interventions, but WFP should take an active role in establishing appropriate benchmarks and make them available for external evaluators. If indicators are selected for which no information is readily available from WFP, it will be a challenge for external evaluators to construct relevant indicators with limited time and resources. No economic skills are required to use this tool.

### *Benchmarking other indicators: applicability*

3. This tool is applicable for all types of WFP interventions and for evaluating country portfolios. TOR should indicate which are relevant indicators to benchmark. The major problem is that benchmarks are not often available. Evaluators must be aware that they are dealing with very partial indicators, and that often trade-offs with other results should be considered. For instance, timeliness may be achieved but at a very high delivery cost, and this may explain why the subsequent distribution cycle has been seriously reduced. In this case the result on timeliness may be very good, but at the expense of size of the rations. WFP staff have to continuously make such trade-offs, and whether they made the right balance in a particular case cannot be read from using just one benchmark. But judicious use of such benchmarks could assist the consistency and transparency of evaluators' judgments.

## Useful Links/Guide for Further Reading

- ALNAP 2006 *Evaluating humanitarian action using the OECD-DAC criteria*, Overseas Development Institute, London, 2006.
- Bailey & Hedlund 2012 *The Impact of Cash Transfers on Nutrition in Emergency and Transitional Contexts - A Review of Evidence*, Bailey, S., Hedlund, K., Humanitarian Policy Group, Overseas Development Institute, London, 2012.
- Barnett et al 2010 *Measuring the Impact and Value for Money of Governance & Conflict Programmes*, Barnett, C., Barr, J., Christie, A., Duff, B., Hext S., ITAD, London, 2010.
- Boardman et al 2011 *Cost-Benefit Analysis – Concepts and Practice*, Boardman, A.E., Greenberg, D.H., Vining, A.R., Weimer, D.L. (2011), fourth revised edition, Pearson.
- DAC 2010a *OECD Glossary of Key Terms in Evaluation and Results-Based Management*, Paris, OECD Development Assistance Committee (2002, reprinted 2010)  
<http://www.oecd.org/development/peer-reviews/2754804.pdf>
- DAC 2010b *Evaluating Development Co-operation Summary of Key Norms and Standards, Second edition*, OECD DAC Network on Development Evaluation (undated but approved 2010)  
<http://www.oecd.org/dac/evaluation/dcdndep/41612905.pdf>
- Hall et al 2011 *The Impact and Effectiveness of Emergency Nutrition And Nutrition Related Interventions: A Review Of Published Evidence 2004-2010*, Hall, A., Blankson, B., Shoham. J., Emergency Nutrition Network. Oxford UK, 2011
- Harvey et al 2012 *Food aid and food assistance in emergency and transitional contexts: a review of current thinking*, Harvey, P., Proudlock, K., Clay, E., Riley, B., Jaspars, S. , HPG – ODI, London, 2012.
- Heider et al 2009 *Country Portfolio Evaluation of WFP Assistance to the Lao PDR: Final Evaluation Report*, Caroline Heider (Team Leader) et al, WFP Office of Evaluation, 28 August 2009.
- Jackson 2012 *Value for money and international development: Deconstructing myths to promote a more constructive discussion*, Penny Jackson, OECD Development Co-operation Directorate, May 2012  
<http://www.oecd.org/dac/effectiveness/49652541.pdf>
- Hodgson et al 2011 *Guidance for DFID country offices on measuring and maximising value for money in cash transfer programmes - Toolkit and explanatory text*, Hodgson, A., White, P., Greenslade, M. DFID, London. 2011.
- Lister et al 2011 *WFP's School Feeding Policy: A Policy Evaluation*, Stephen Lister, Stephen Anderson, Mirella Mokbel Genequand, Anthea Gordon, Judith Sandford and Stephen Turner. Mokoro. (30 November 2011, Report number: OE/2012/002.)

- Palenberg 2011 *Tools and Methods for Evaluating the Efficiency of Development Interventions*, Evaluation Working Papers, Bonn, Bundesministerium Für Wirtschaftliche Zusammenarbeit und Entwicklung (BMZ), Palenberg, M. (2011)
- Renard 2013 *WFP Efficiency Analysis Guidance Materials Background Paper – Definitions, Concepts and Methodology*, Robrecht Renard, June 2013 (final).
- Ryckembusch 2012 *Transfer Modality Selection*, powerpoint, Ryckembusch, D., WFP. Rome, 2012
- Ryckembusch et al 2013 *Enhancing Nutrition: A New Tool for Ex-Ante Comparison of Commodity-Based Vouchers and Food Transfers*, Ryckembusch, D., Frega, R., Guilherme Silva, M., Gentilini, U., Sanogo, I., Grede, N., Brown, L., World Development, 2013.
- WFP 2009 *Country Portfolio Evaluation of WFP Assistance to the Lao PDR, Final Evaluation Report*, Caroline Heider et al, 28 August 2009.
- WFP 2011 *Efficiency at WFP*, Board Paper WFP/EB.2/2011/5-E/1, 18 October 2011
- WFP 2012 *School Feeding Cost Analysis – Final Report*, powerpoint, WFP, Rome. 2012.

## Acronyms and abbreviations

AIDS	Acquired Immunodeficiency Syndrome
BCR	Benefit Cost Ratio
BMZ	Federal Ministry for Economic Cooperation and Development, Germany
CBA	Cost Benefit Analysis
CEA	Cost Effectiveness Analysis
CER	Cost Effectiveness Ratio
CO	Country Office
CPE	Country Portfolio Evaluation
CTR	Cost-Transfer Ratio
CUA	Cost-Utility Analysis
DAC	Development Assistance Committee
DALY	Disability Adjusted Life Year
DEV	Development Operation
DR	Discount Rate
DRM	Disaster Risk Mitigation
DSC	Direct Support Costs
EM	Evaluation Manager
EMOP	Emergency Operation
EQ	Evaluation Question
EQAS	Evaluation Quality Assurance System
ER	Evaluation Report
FBF	fortified blended food
FFA	Food For Assets
FFR	Food for Relief
FFW	Food For Work
HIV	Human Immunodeficiency Virus
IRR	Internal Rate of Return
ISC	Indirect Support Costs
LTSH	Land-side Transport, Storage, Handling
MADM	Multi-Attribute Decision Model
MCHN	Maternal and Child Health and Nutrition
M&E	Monitoring and Evaluation
MT	Metric Ton
NGO	Non-governmental organization
NPV	Net Present Value
NVS	Nutritional Value Scores
ODOC	Other Direct Operating Costs
OECD	Organisation for Economic Co-operation and Development
OEV	WFP Office of Evaluation
OpEv	Operational Evaluation
P4P	Purchase for Progress
PROMIS	Performance and Risk Management Organizational Information System
PRRO	Protracted Relief and Recovery
QALY	Quality-Adjusted Life Year
SEQ	Specific Evaluation Question
TOR	Terms of Reference
UNICEF	United Nations Children's Fund
UNIDO	United Nations Industrial Development Organisation
WFP	World Food Programme
WHO	World Health Organisation