

# The 2008 Reader on Private Sector Development

By Jim Tanburn



# Measuring and Reporting Results



**The 2008 Reader  
on Private Sector Development**

**Measuring and  
Reporting Results**

By Jim Tanburn

Co-funded by SDC



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

**Swiss Agency for Development  
and Cooperation SDC**

Copyright © International Training Centre of the International Labour Organization 2008

This publication enjoys copyright under Protocol 2 of the Universal Copyright Convention. Applications for authorization to reproduce, translate or adapt part or all of its contents should be addressed to the International Training Centre of the ILO. The Centre welcomes such applications. Nevertheless, short excerpts may be reproduced without authorization, on condition that the source is indicated.

---

THE 2008 READER' ON PRIVATE SECTOR DEVELOPMENT  
MEASURING AND REPORTING RESULTS

ISBN 978-92-9049-437-9

First edition 2008

---

The designations employed in publications of the International Training Centre of the ILO, which are in conformity with United Nations practice, and the presentation of material therein do not imply the expression of any opinion whatsoever on the part of the Centre concerning i.a. the legal status of any country, area or territory or of its authorities, or concerning the delimitation of its frontiers. The responsibility for opinions expressed in signed articles, studies and other contributions rests solely with their authors, and publication does not constitute an endorsement by the Centre of the opinions expressed in them.

Publications of the Centre, as well as a catalogue or list of new publications, can be obtained from the following address:

Publications International Training Centre of the ILO  
Viale Maestri del Lavoro, 10  
10127 Turin, Italy  
Telephone: +39 - 011 - 6936.693  
Fax: +39 - 011 - 6936.352  
E-mail: [publications@itcilo.org](mailto:publications@itcilo.org)

---

Designed and printed by the International Training Centre of the ILO, Turin - Italy

## Summary

There is little information available about the impacts of programmes for private sector development (PSD), largely because:

- goals are very ambitious, and impacts costly to quantify – relative to the resources available; indeed, the cost of measuring impacts is often classified as an ‘overhead’, to be kept to a minimum; and
- systemic change in the private sector as a whole does not lend itself to the mechanistic model of inputs-outputs-outcomes-impacts in conventional thinking; attribution and timing issues are acute.

Besides, practitioners would need to accept the methodology, and to be rewarded for good performance, for results measurement to be adopted on a large scale. However, current indicators in common usage, such as leverage (to be maximised) and overhead (to be minimised), encourage perverse incentives and distract from the core task of achieving developmental goals. Furthermore, the many self-published ‘success stories’ leave most observers confused.

In the absence of much discussion on the subject, it remains rather sensitive, and one that people therefore try to avoid. Meanwhile, external pressures are growing, for more information; they are coming from donors (e.g. through the Paris Declaration, the MDG deadline), new players and aid models (e.g. social investors) and increased visibility (e.g. Live8). This Reader argues that practitioners need to seize the initiative and to develop answers, before someone else does it for them. In the absence of good data, critics will always be able to say: ‘if you cannot measure it, maybe it is not there’.

A brief overview is therefore given of current understanding in the field, including particularly the terms, indicators and methodologies in use. It is argued that multi-agency agreement in these areas would yield very important benefits, in addition to an approximate comparison of performance; for example:

- agencies could add impacts achieved across all of their country programmes, enabling them to report results for the agency as a whole
- agencies would also be able to make informed choices about which intervention strategies to fund

Examples are given of impacts measured in a standard format, including for example cost per job created; since the resulting numbers are very different in magnitude, they make a rational conversation about strategy choice possible – even if they are only correct to within +/- 50%. Agreement now needs to be built around the key parameters for formulating these numbers, including for example the multipliers to use for indirect impacts.

Approximate measures do not replace the need for rigorous impact assessments. But agreement between agencies on a small number of indicators, and their application across a wide range of interventions, would win recognition for the achievements of the PSD community. Affordable mechanisms are needed, to ensure that the numbers produced are credible – for example through certification of the methodologies used. Finally, rewards for cost-effectiveness will motivate and orient practitioners.

# Preface

This Reader comes at a particular point in the history of development, and of the development of value chains and service markets in particular. The tax-paying public in donor countries are wondering what their donor agencies are achieving, and some people are proposing that the answer is “not much”. They can do this, because there is little that is published about results, which is both convincing and comparable.

There are, of course, real challenges in measuring and comparing results. Those who have worked in the field for some time will already be familiar with them, and will be looking for fresh perspectives – rather than the usual agreement that “we ought to do more”.

This Reader aims to do exactly that, arguing that debates about rigour in methodologies have distracted from the more important institutional and human barriers to measuring results. These barriers need to be addressed in creative ways, as results can be estimated in ways that are affordable and not technically demanding. This Reader will have succeeded, if it leads to greater measurement and reporting of results; comments on the text are particularly welcome.

Ideally, there would be a crisp definition of the intended readership; in practice, however, the communities of practice are now fluid and overlapping. Certainly, it will be of interest to those developing value chains and service markets; it is also likely to be of interest to those engaged in broader reform of the business environment, and indeed in private sector development (PSD) more generally.

The format of this document is a break with the past, in that it does not seek to emulate the heroic efforts of Aly Miehlabrad and Mary McVay in profiling all recent developments in PSD. Nonetheless, Annex A includes a list of all of the entries that have been created or updated in the last year on [www.Value-Chains.org](http://www.Value-Chains.org), to give a quick overview of some of the most recent interesting publications.

My thanks to the many people who contributed substantial information to the preparation of this Reader, including Margrethe Holm Andersen, Geeta Batra, Alwyn Chilver, Nazia Habib-Mintz, John Marsh, Peter Roggekamp, Peter Schmidt, Don Sillers and Thom Sprenger and especially Aly Miehlabrad; the IFC workshops on this subject were a particular inspiration. Any errors are, of course, mine. Finally, my particular thanks to SDC for co-funding the document’s preparation and publication.

Jim Tanburn  
Jim@Tanburn.com

# Contents

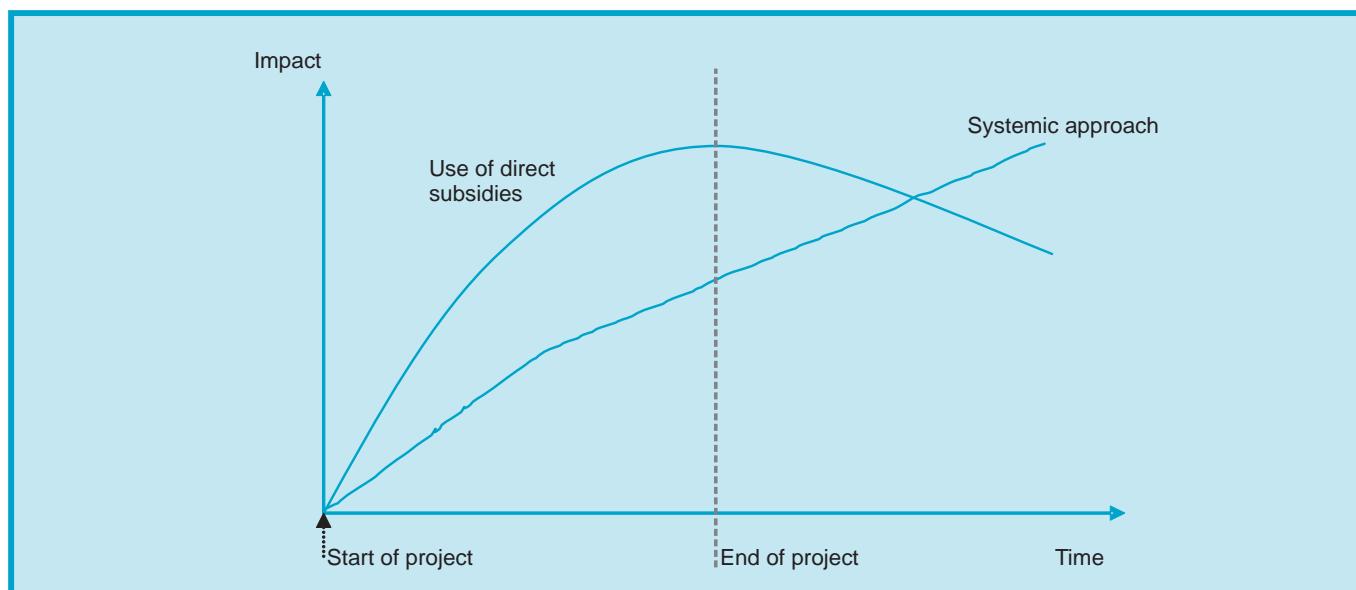
<b>Summary</b> .....	<b>iii</b>
<b>Preface</b> .....	<b>iv</b>
<b>A. Introduction</b> .....	<b>1</b>
<b>B. Is There Really a Lack of Information About Results?</b> .....	<b>3</b>
<b>C. Pressures for Change</b> .....	<b>6</b>
Pressures Within the Development Community .....	6
The Paris Declaration on Aid Effectiveness .....	7
New Players .....	8
<b>D. Methodologies in Use: A Primer</b> .....	<b>10</b>
Monitoring or Evaluation? .....	10
The Core Task .....	10
The Causal Model .....	11
Experimental and Quasi-Experimental Designs .....	13
Proxies .....	14
<b>E. A Discussion of the Methodological Options</b> .....	<b>15</b>
Randomised Controlled Trials (RCTs) .....	15
Alternatives to Experimental Designs .....	16
Agency Rankings of Own Programmes .....	17
Building Systems for Management .....	18
<b>F. Indicators ('What to Measure')</b> .....	<b>20</b>
Poverty Alleviation or Economic Growth? .....	20
Measuring Programme Costs .....	21
<b>G. PSD and the MDGs</b> .....	<b>22</b>

<b>H. Other Measurement Issues .....</b>	<b>24</b>
The Time Factor .....	24
Human and Institutional Factors .....	25
<b>I. The Benefits of Approximate Measures .....</b>	<b>28</b>
Bulking Up Programme Achievements Across the Whole Agency .....	28
More Rational Funding Choices .....	29
Comparison Between Programmes and Agencies .....	29
Early Adopters of Core Indicators .....	30
<b>J. Reporting on Scale.....</b>	<b>32</b>
<b>K. Impact Assessments for Technical Assistance in Value Chain Development.....</b>	<b>34</b>
SEDF in Bangladesh.....	34
The Prosperity Initiative and Bamboo in the Mekong.....	35
Cluster Initiatives in Sri Lanka.....	36
Katalyst in Bangladesh.....	37
Water Pumps and Other 'Appropriate Technologies' .....	39
<b>L. Impact Assessments for Management Training.....</b>	<b>40</b>
Management Training in India.....	40
Training of Peruvian Micro-finance Clients.....	40
<b>M. Impact Assessments for Challenge Funds, Matching Grants and Public Private Partnerships (PPPs) .....</b>	<b>42</b>
Programme for Cooperation with Emerging Markets (PSOM), Netherlands.....	44
DfID Challenge Funds .....	45
<b>N. Impact Assessments in Reform of the Business Environment ...</b>	<b>47</b>
<b>O. Summary of Results to Date .....</b>	<b>49</b>
<b>P. The Future.....</b>	<b>52</b>
<b>Annex A: New or Updated Entries on <a href="http://www.Value-Chains.org">www.Value-Chains.org</a> .....</b>	<b>54</b>



## A. Introduction

1. Development practitioners are given ambitious goals; they are expected to achieve very large and measurable results, with resources that are actually very modest, relative to the economies they are hoping to influence. Indeed, some of the expectations are arguably not even realistic, within the tight timeframes and budgets of most development agencies.
2. Furthermore, the goals are multi-layered, usually including sound commercial performance in the market, and a wide variety of social and development goals; the developmental goals are often diverse, including for example a focus on the poorest, gender concerns, maybe a rural flavour, mention of youth, and so on. Within this setting, it can be difficult to identify and maintain clear priorities, and many practitioners are therefore juggling multiple priorities, in order to satisfy as many of the stakeholders as possible.
3. And there are many stakeholders: donor staff in the field and at headquarters, colleagues in the implementing organisation, and counterparts in government and elsewhere. In addition, practitioners have to establish credibility with potential partners in the private sector, who may already be suspicious of donor-branded programmes. All this – before questions of measuring results arise.
4. Most people in the field are fully committed to getting things done; arguably, they have to be, in such a complex environment, if they are to achieve anything. They are not, by nature, statisticians or academics, and generally find the task of rigorous results measurement daunting. And it cannot be denied that rigorous measurement of results is expensive, with sums in excess of \$250,000 being mentioned. Besides, most donor money is intended for use in making a difference in the world, rather than in measuring it; measurement is often classified as an ‘overhead’, with the associated pressures to reduce the cost to an absolute minimum.
5. The increasing focus on ‘systemic’ approaches, where practitioners are expected to understand entire systems, has made measurement more challenging. Rather than just meeting the internal needs of the system within their own agency (and its funders), managers also have to design and implement interventions that make sense also to the people in a completely different system: the one within which the intended target group lives and works. Practitioners are no longer exerting a calibrated influence over a carefully-controlled and limited set of players, but seeking to influence an entire sector of the economy.
6. One rationale for such systemic approaches is that, by building on local dynamics, wishes and ownership, they will be much more likely to achieve sustainable improvements. If those improvements are sustainable, then the ultimate impacts will be much greater, as they will continue to accrue (and perhaps also to grow) long after the intervention has come to an end. According to this logic, the longer-term impacts – and particularly the impacts of spontaneous replications or ‘copy-cats’ – will add up to paint a truly impressive picture of value for money.
7. Much has been written about this shift already (not least in previous years’ Readers) so it will not be explored in more detail here. However, it is important with respect to the theme of measuring results, as it raises interesting questions about when to measure those results. Measurement at the end of a project may find better impacts generated by an intervention that used subsidies liberally – but these impacts will probably decline rapidly once the external financing comes to an end (as in the illustration, below).



8. Systemic approaches, meanwhile, may take longer to understand local dynamics and demand, and to establish the credibility needed to catalyse long-term changes in the market as a whole. They do, actually, raise questions about the conventional logic of measuring results, which usually involves an implicit, mechanistic model of achievement: funds paid at one end will ultimately lead to defined results that come out at the other end of the 'machine'. In practice, the desired impacts of a systemic approach may change over time, as the aspirations of the target group evolve, and as new market-based opportunities arise during implementation.

9. In addition, there are so many influences being exerted on the market system that the results achieved by any particular development programme

are unlikely to be replicated anywhere else, or at any other point in history. Unfortunately, however, the elegance of these arguments has distracted from the original logic of systemic approaches: to demonstrate greater impact. Even in programmes where managers could have argued very convincingly for explicit and likely impacts in the medium term, they have very rarely done so. Arguably, this has led to a decline in interest in the paradigm.

10. Recent years have seen some important new pressures being brought to bear on practitioners, to demonstrate results in a more effective way, based partly on the impression that the current state of affairs is not satisfactory. The following Chapter explores that perception in more detail.

## B. Is There Really a Lack of Information About Results?

*The previous Chapter argued that development agencies are working in areas that are complex – particularly with respect to systemic approaches. The pressures to demonstrate results have also grown – but is there really a lack of information?*

11. Last year's Reader was sub-titled: "Striving for Tangible Results for the Poor"; it noted a "continued failure to measure, document and disseminate significant results in eradicating poverty, or even reaching large numbers on a sustainable basis, despite strong anecdotal evidence of significant impact".

12. Other synthesis documents have been making similar observations for some time; one such document for USAID in 2004, for example, reviewed 50 evaluations from various agencies<sup>1</sup>, and concluded that "very few studies use control groups or time series data that would allow studying change over time and comparison of participants and non-participants ... Self-selection was an outstanding issue in all studies, since none of them used random experimental design or corrected for the problem."

13. Similarly, the Director of the Shell Foundation recently wrote: "what always stuns me at events like the Global Philanthropy Forum is that so few of the NGOs and charities presenting offer any sort of independent validation of their impacts or present these in a comparative framework against the performance of other organisations in the same field.

14. "Likewise, the assembled donors didn't seem too interested in documented results or in the nature of the accountability of the presenting organisation. Indeed in a three-day-long conference with dozens of individual sessions, there was only one very poorly-attended session that considered the issue of measuring impact – and I never heard the phrase 'customer service' mentioned!!"<sup>2</sup>

15. In the specific field of private sector development (PSD), there are also many signs of a lack of information. Altenburg and von Drachenfels, for example, have written: "Although the BDS debate has been under way for almost ten years and has received a lot of attention among donor agencies, there is still almost no empirical evidence of sustainable BDS programmes."<sup>3</sup>

16. Actually, these comments demonstrate a misunderstanding, in that they were looking for evidence that donor-funded programmes had achieved sustainability. There are many reasons why BDS programmes that were started with donor funding are unlikely to become sustainable. The debate has been rather about how donor-funded programmes could enhance the value addition of BDS providers who are already operating sustainably in the private sector; the lack of credible research to document achievements does, however, create the 'space' in which such assertions can be made.

<sup>1</sup> Zandniapour et al, 2004. Review of Impact Assessments of Selected Enterprise Development Projects. DAI for AMAP BDS Knowledge and Practice Research Task Order. [http://www.microlinks.org/ev02.php?ID=7102\\_201&ID2=DO\\_TOPIC](http://www.microlinks.org/ev02.php?ID=7102_201&ID2=DO_TOPIC)

<sup>2</sup> [http://www.shellfoundation.org/newsletter/14\\_05\\_2007/directors\\_message.htm](http://www.shellfoundation.org/newsletter/14_05_2007/directors_message.htm)

<sup>3</sup> The 'New Minimalist Approach' to Private Sector Development: A Critical Assessment, by Tilman Altenburg and Christian von Drachenfels. Development Policy Review, 2006, 24 (4): 387-411

17. Another factor supporting such assertions is the way in which interest in a particular type of intervention lasts perhaps 3-5 years, before agencies move to the next 'big thing'<sup>4</sup>. This is again a result of the lack of information about results; agencies therefore must run with concepts that are elegant and attractive, rather than with approaches that are proving to be effective (since proof is rarely available, within the timescale required).

Unfortunately, programme managers may therefore find themselves evaluated against new and more 'up-to-date' yardsticks, rather than against the programme document they had been trying to implement.

18. In the meantime, and in the absence of any commonly accepted methodology for measuring and reporting results, every agency – and indeed every project – makes its own measurements and does its own reporting. This is often under the heading of 'success stories', already reinforcing the perception that the contents are not impartial or objective.

19. There is also the perception that the indicators have been carefully chosen, to tell the story of the results in the best possible light. The reader must therefore think hard about the contents, in order to work out what the reasonable questions might be. Actually, many 'success stories' contain little or no 'hard' information; some include no quantified results in any form at all, preferring to tell the story of one or two carefully-chosen beneficiaries.

20. An example is given on the following page of the genre – representing one of the most convincing available, in the sense that it does provide a clear story and some quantified impacts. While not relating the results to the cost of the intervention, this summary does give a clear sense of the scale and the incomes resulting from the new seeds. The problem is that the reader, unless she is particularly expert in this field, cannot get a sense of whether the intervention was successful, relative to other, comparable programmes. It sounds good, but how good is good?

21. Finally, there is surprisingly little dis-aggregation of results by sex, so it is often difficult to learn much about the impact on gender issues. Wherever data are available from the cases referred to in this Reader, it is included in the text. Otherwise, and until dis-aggregation is more thoroughly implemented, most commentators instead note that poverty alleviation is likely to favour women in particular, since the majority of those living in poverty are women. Clearly, however, more needs to be done to understand the gender dynamics within the overall impact 'stories'.

22. There are pressures to improve on the current situation, and these are explored in more detail in the following Chapter.

<sup>4</sup> See Reusse, 1999, *The Ills of Aid* (University of Chicago Press) for a more detailed discussion on this – particularly in Chapter 3.1, "The Paradigm Life Cycle".

## Case Study: September 2006

### CHALLENGE

Farmers in Uganda, like in most of the under-developed economies, are often trapped in a vicious cycle of poverty. They often obtain low yields resulting in low farm incomes. The low yields mean that the farmers need unrealistically high prices in order to make profits. Yet the farmers have no control on prices and are often vulnerable to market shocks. Like several other crops, sunflower, which has been produced in Uganda for over a decade and is a basic feedstock in the oilseed processing industry, was found to have yields that were insufficient to encourage farmers to continue growing the crop and thus expand production, despite Uganda having a major and growing deficit in vegetable oils.

### INITIATIVE

When [the project] commenced operation in 2004, it was recognized that the Sunfola variety the farmers were planting had degenerated and this had resulted in low farmer confidence. [The project] advocated for suitable hybrid varieties and collaborated with NARO to screen imported hybrids and open pollinated varieties.

In partnership with A.K. Oils & Fats (U) Ltd, [the project] promoted a new hybrid sunflower variety, *PAN 7351* to farmers in Lira, Apac, Masindi and Sironko through the establishment of an average of 850 farmer field demonstration sites per season in these districts. [The project] promoted an outgrower scheme (OGS), where the farmers grow the crop knowing there is a ready market.



A hybrid sunflower field in Kyatiri, Masindi district belonging to Mr. John Kyomya. John is one of the commercial farmers in the OGS who have adopted the recommended production practices.

### RESULT

Over the past three years the OGS has reached a total of 31,300 farmers and generated gross earnings of US\$5,367,000 (US\$2,146,800 in net incomes). Farmers have been so encouraged by the presence of a guaranteed market that all growers are willing to pay for seed in advance of stocks arriving in-country, in order to secure planting material. A.K. Oils & Fats (U) Ltd uses site coordinators as the aggregation point for all seed sales (no commission paid on seed sales) as well as output aggregation (US\$ 10 per kg paid to site coordinator as incentive). Site coordinators now are micro enterprises in their own right – building stores, developing farmer skills in production (since farmer output directly influences their revenues) and working closely with A.K. Oils & Fats (U) Ltd team.

## C. Pressures for Change

*The previous Chapter looked at the lack of comparable information about the impacts of development work – particularly in private sector development (PSD). This lack gives rise to a number of problems, including rapid swings in interest of development agencies, from one paradigm to another. This Chapter examines whether the situation is now different, relative to the situation 5 or 10 years ago.*

### Pressures Within the Development Community

23. Recent years have seen some important developments; for example, the incidence of poverty in Asia has fallen from 32% in 1990, to 19% today<sup>5</sup>. But the public perception is generally that such achievements have not been due to the work of development agencies. The Green Revolution was seen as a developmental achievement, but since then, development agencies have not scored many major ‘hits’.

24. Rather, their efforts are often portrayed as insignificant – and occasionally as doing more harm than good. Indeed, there have been several books in recent years, arguing this case, with eloquent titles such as ‘Despite Good Intentions: Why Development Assistance to the Third World has Failed’ and ‘The Road to Hell: The Ravaging Effects of Foreign Aid and International Charity’.

25. William Easterly has written another in the genre: ‘White Man’s Burden: How the West’s Efforts to Aid the Rest Have Done So Much Ill’; he asks: “why have aid interventions where it is so difficult to

tell whether they are working or not? What incentives follow from objectives for which you can’t hold anyone accountable if they are not met?”

26. He concludes: “1) Don’t do things that can’t be evaluated. 2) Don’t design an aid program such that there are no consequences of a negative evaluation. 3) Don’t use the word ‘evaluation’ when what you are describing is not an independent evaluation of a specific intervention for which somebody can be held accountable.”<sup>6</sup>

27. Indeed, agencies have for years published manuals about how to measure results – usually involving a menu of options of what to measure, and advice on how to measure them. The following was published in 1996:

28. “The development of objective indicators of performance is .. essential for the public accountability of the MDBs [Multilateral Development Banks] and their ability to justify their use of public resources to shareholder governments, parliaments, and the public. Currently, it is not possible to compare their operational results, or even to describe them in a common language. Major public sector institutions like the MDBs must be able to account for their efforts in readily understood terms.

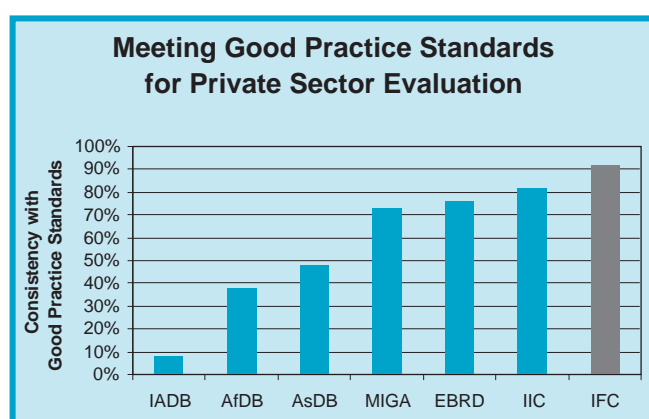
29. “A common methodology for evaluating their portfolios should be developed and kept up to date over time, with best practices in evaluation techniques being identified and disseminated. A determined effort should be made to harmonize performance indicators and evaluation criteria, taking into account the differing circumstances of each institution. The lessons learned from these

<sup>5</sup> Mallon, 2006. Asia Overview Paper. Prepared for the Donor Committee for Enterprise Development Conference, Bangkok. [www.businessenvironment.org/dyn/be/docs/132/Day1PlenaryMallon.pdf](http://www.businessenvironment.org/dyn/be/docs/132/Day1PlenaryMallon.pdf)

<sup>6</sup> IFC Outcomes newsletter, March-June 2007 [www.ifc.org/ifcext/sme.nsf/AttachmentsByTitle/OutcomesNewsletter.pdf/\\$FILE/OutcomesNewsletter.pdf](http://www.ifc.org/ifcext/sme.nsf/AttachmentsByTitle/OutcomesNewsletter.pdf/$FILE/OutcomesNewsletter.pdf) See also <http://www.nyu.edu/fas/institute/dri/Easterly>

evaluations should be shared among the MDBs with a view to applying them quickly in new operations.”<sup>7</sup>

30. In practice, however, little progress has been made on an inter-agency basis in developing a common methodology. Most agencies apparently struggle, even to conform to their own good-practice guidelines – although this could be interpreted as suggesting that the current guidelines do not address core concerns of either the agency or its staff. The chart, below, gives IFC’s assessment of the extent to which this is the case<sup>8</sup>.



31. Researchers continue to try to define which types of intervention will give the best value for money. One example is the Copenhagen Consensus<sup>9</sup>, which identified some interventions as being the ‘best’, including for example HIV/Aids prevention, promotion of dietary supplements and the liberalisation of trade. Similarly, another study reported that the traditional donor priority sector of education did not seem to be effective (finding no correlation between increased access to education, and increases in prosperity)<sup>10</sup>. These types of analysis illustrate the interest of many people to define the

most cost-effective ways to invest their development funds; they remain, however, controversial.

32. The pressures continue, though, to do more; for example, when Robert Zoellick was recently appointed as President of the World Bank, he was “told by Treasury Secretary Henry Paulson, Jr. and others that the bank must do more to provide definitive measures of the effects of its \$23 billion in lending to poor countries, in part to assure Congress about how the money is being spent.”<sup>11</sup>

## The Paris Declaration on Aid Effectiveness

33. Donors have certainly been challenged to respond to the accusations that they are not as effective as they might be. One of the better-known responses recently has been the Paris Declaration on Aid Effectiveness, in March 2005<sup>12</sup>. This Declaration proposed that funds increasingly be channelled through partner governments. It also pledged to reduce by one-third “the proportion of countries without transparent and monitorable performance assessment frameworks”, by the year 2010.

34. There is, however, concern about progress in consolidating aid assistance in this way; the OECD wrote recently that “at country level, the 2006 Survey also raises *serious concerns about the high costs of delivering and managing aid* [emphasis in original text]. In 2005, the 34 developing countries covered by the survey received 10,507 donor missions, more than one for each working day. Even those that explicitly asked for “quiet periods” to get on with their day-to-day work were not always spared.”<sup>13</sup>

<sup>7</sup> Development Committee, Task Force on Multilateral Development Banks, “Serving a Changing World—Report of the Task Force on Multilateral Development Banks,” March 15, 1996, p. 18. As cited in MDB Evaluation Cooperation Group Working Group on Private Sector Evaluation: Good-Practice Standards for Evaluation of Private Sector Investment Operations, 2006 [www.adb.org/Evaluation/GPS-3rd-edition.pdf](http://www.adb.org/Evaluation/GPS-3rd-edition.pdf)

<sup>8</sup> <http://www.businessenvironment.org/dyn/be/docs/141/Michelitsch.pdf>

<sup>9</sup> [www.copenhagenconsensus.com](http://www.copenhagenconsensus.com)

<sup>10</sup> Lewis, 2004. *The Power of Productivity: Wealth, Poverty and the Threat to Global Stability*. McKinsey and Co. / the University of Chicago Press

<sup>11</sup> International Herald Tribune, 1<sup>st</sup> June 2007. Zoellick signals plans to revamp World Bank

<sup>12</sup> Available from many sites, including <http://www1.worldbank.org/harmonization/Paris/FINALPARISDECLARATION.pdf>

<sup>13</sup> <http://www.oecd.org/dataoecd/29/3/38435815.pdf>

35. There is also concern about what this proposal means for PSD; the Paris Declaration does not identify an explicit role for the private sector in donor-funded assistance. PSD practitioners would expect the private sector to be closely involved in both the design and the implementation of that assistance, implying a more nuanced approach than simply channelling all funds through the partner government. In principle, this can be taken into account through the participation of the private sector in the preparation of the national Poverty Reduction Strategy; in practice, however, such inputs are not always achieved. Anyway, the participation of the private sector should be continuous, rather than through a one-off exercise.

36. Another problem in channelling PSD support exclusively through partner governments is that there are usually many Ministries involved in one way or another in PSD. Again, therefore, PSD practitioners are challenged to show that the results of a more nuanced approach justify the extra complication.

## New Players

37. The private sector has itself become a major influence recently, and this influence is likely to grow further, in the coming years. Many companies are seeing development issues increasingly as part of their core business model, for a variety of reasons.

38. New social investors, such as Acumen, are also emerging. Some of these new players are vocal and media-savvy – including for example the Shell Foundation, quoted above. The Bill and Melinda Gates Foundation, Google.org and other ‘thought leaders’ are having an influence far beyond their (substantial) funds; the Gates Foundation, for example, notes as one of its Guiding Principles: “delivering results with the resources we have been given is of the utmost importance – and we seek and share information about those results”.

39. Some additional examples of relatively new organisations in this trend are given below:

- the Sustainable Food Lab, [www.sustainablefoodlab.org](http://www.sustainablefoodlab.org); launched by the Kellogg Foundation and Unilever, now with 70 members – including both companies (such as General Mills, Ahold, Starbucks and JPMorgan Chase) and Foundations (such as Gates, Oxfam, Shell and Technoserve)
- the Sustainable Agriculture Initiative (SAI)([www.saiplatform.org](http://www.saiplatform.org)); SAI was founded to support sustainable agriculture, by Danone, Nestle and Unilever; members now include Coca-Cola, Findus, Kraft and McDonalds
- the World Business Council for Sustainable Development (WBCSD) ([www.wbcsd.org](http://www.wbcsd.org)), whose membership now includes 190 companies
- the Business Social Compliance Initiative ([www.bsci-eu.org](http://www.bsci-eu.org)) for ethical procurement, which has about 70 retailers, importers and manufacturers in Europe as its members
- EurepGAP ([www.eurepgap.org](http://www.eurepgap.org)) is a private sector body that sets voluntary standards for the certification of agricultural products
- A ‘Private Sector in Development Initiative’, launched recently to (among other things) “define common standards for measuring and reporting economic, social and environmental impacts”<sup>14</sup>
- Business Action for Africa (BAA)([www.businessactionforafrica.org](http://www.businessactionforafrica.org)) has 150 members (80% coming from the business sector)

40. Indeed, very many of the people involved in these initiatives, coming from the private sector, pride themselves on performing against agreed metrics; they bring with them the assumption that this will be both possible and desirable in the field of PSD. While not all their expectations will necessarily be realised, they are nonetheless much more oriented towards achieving measurable results, than some practitioners are used to. The pressure to report results is, therefore, likely to increase in the near future.

<sup>14</sup> [www.aspeninstitute.org/site/apps/nl/content2.asp?c=huLWJeMRkPH&b=2463765&ct=3529087](http://www.aspeninstitute.org/site/apps/nl/content2.asp?c=huLWJeMRkPH&b=2463765&ct=3529087)



41. Other initiatives have grown up to support this trend. The Financial Times, for example, recently partnered with Dalberg (a consultancy firm) and the UN Global Compact to invite businesses to rate their non-profit partners<sup>15</sup>; this initiative is explained in more detail in a subsequent Chapter. Similarly, there has been a rise in training courses on poverty, business and development<sup>16</sup>.

42. Similarly, the Clinton Global Initiative and Dalberg recently published a report accusing mainstream development agencies of insufficient demand orientation, being costly and slow, and lacking innovation and accountability. Recommendations included creating transparent reporting and standards for programme delivery and supply chain performance, no longer “rewarding failure”, and using accountability to create a system that is responsive and dynamic<sup>17</sup>.

43. Finally, development aid has enjoyed a higher profile in recent years, for example through the Live8 series of concerts. While this has increased substantially the political support for development aid, it has also made the issue of results more challenging – public opinion is not particularly

interested in the finer points of evaluation methodology or systemic change.

44. The international landscape is also changing, with new donors emerging; this year, for example, “China pledged \$20 billion to finance trade and infrastructure across [Africa] over the next three years”<sup>18</sup>. Indeed, China’s trade with Africa already exceeded \$55 billion in 2006. Countries such as Mexico, South Korea and Poland are also becoming significant donors, potentially bringing their domestic PSD experiences to the discussion.

45. All of these pressures have led donor agencies to place a high priority on measuring results; in a recent survey, members of the Donor Committee for Enterprise Development voted the broad category of “impact assessment, benchmarking of results” as the highest priority for the Committee in the coming years<sup>19</sup>.

46. The following Chapter reviews the current methodologies in use, and is intended as a Primer on the subject. Seasoned practitioners, therefore, may prefer to go directly to the following Chapter.

<sup>15</sup> [www.ft.com/reports/philanthropy2007](http://www.ft.com/reports/philanthropy2007)

<sup>16</sup> See for example Cambridge University’s Business and Poverty Leadership Programme, [www.cpi.cam.ac.uk/programmes/poverty\\_development/bplp/about\\_the\\_programme.aspx](http://www.cpi.cam.ac.uk/programmes/poverty_development/bplp/about_the_programme.aspx)

<sup>17</sup> <http://www.dalberg.com/pdfs/taskforce.pdf>

<sup>18</sup> <http://www.iht.com/articles/2007/08/20/asia/zambia.php>

<sup>19</sup> [www.enterprise-development.org/resources/item.asp?resourceid=387](http://www.enterprise-development.org/resources/item.asp?resourceid=387)

## D. Methodologies in Use: A Primer

*The previous Chapter argued that the pressures for development agencies to report on their relative achievements are now greater than before. This Chapter provides a Primer on methodologies for measuring impacts. It is not intended to be comprehensive, as there are many guides now published, which cover similar ground, more comprehensively<sup>20</sup>. Experienced practitioners can probably move directly to the following Chapter.*

### Monitoring or Evaluation?

47. Discussions about measurement of results often refer to “monitoring and evaluation” or M and E, whereas these two terms refer to two separate concepts. Monitoring is about on-going measurement of performance, particularly looking at parameters like efficiency; are things being done right? Ideally, these kinds of measurements can lead to improvements during implementation; they are often conducted internally by the intervention team.

48. Evaluation, however, is the focus of this Reader; it is often conducted by external consultants, and is about proving impacts (rather than improving interventions). It is answering questions about whether the right things are being done, and tends to be rather more sensitive – because it is potentially a measure of the performance of the programme design, and of the implementing team, and of the implementing agency. Reputations and careers are potentially at stake.

49. Evaluations conducted at the end of programmes are referred to as ‘*ex-post*’, while assessments of anticipated impacts conducted during the programme design process are referred to as ‘*ex-ante*’. Ideally, *ex-ante* evaluations would make rational spending decisions possible.

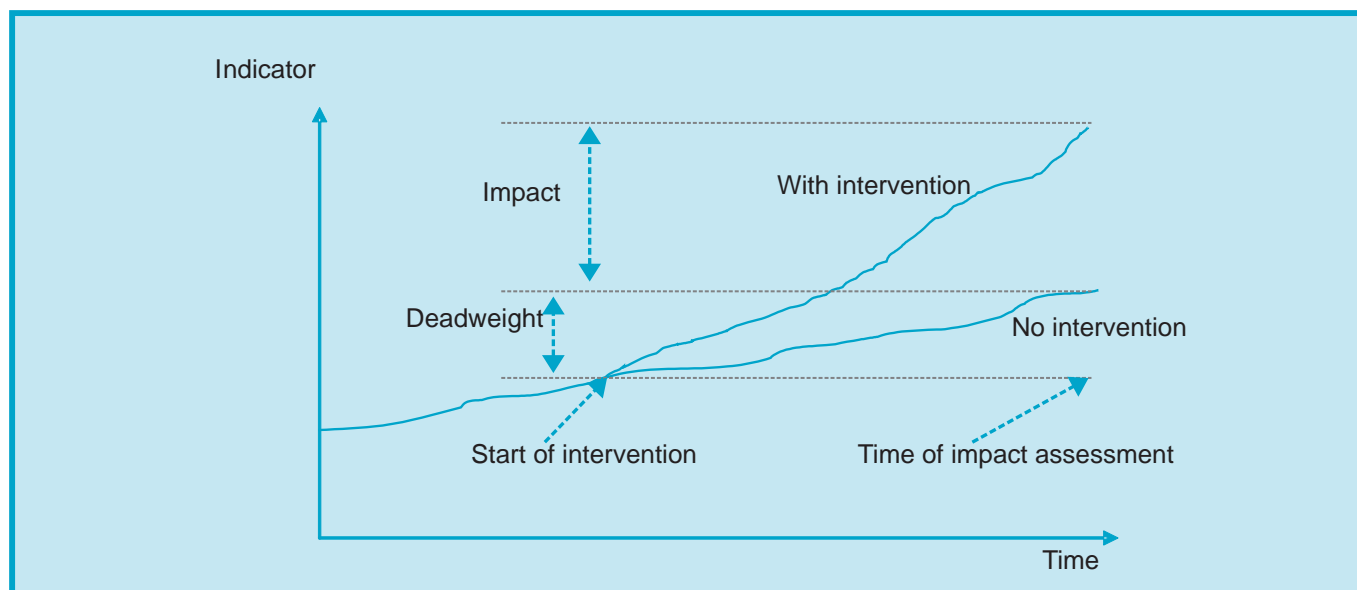
50. The following discussion will focus on evaluation, rather than on monitoring. Historically, evaluations have tended to look at outputs (e.g. number of people trained), but are now increasingly looking at outcomes (e.g. changes in behaviour as a result of the training). The term ‘impact’ refers to the developmental results that donors ultimately seek as a consequence of the outcomes (e.g. jobs created, people lifted out of poverty). Impacts are the most difficult to measure with rigour, and to report in ways that would convince someone who is sceptical about the intervention.

### The Core Task

51. The core task in measuring impacts is to ‘establish the counter-factual’: to discover what would have happened, if the intervention had not taken place at all. All reporting of results implicitly covers the counter-factual; the message is always: ‘without our intervention, this would not have happened’. In practice, of course, the counter-factual is tricky to establish, for various reasons:

- ‘*attribution*’: one has to demonstrate that the measured impacts resulted from the intervention, rather than from other interventions (perhaps by another agency), or from something completely different
- ‘*displacement*’: one also has to demonstrate that those who did not benefit directly from the intervention did not suffer, at the expense of those who did benefit (the ‘treatment group’); for example, did management training enable some entrepreneurs to do better, penalising those who did not receive the training?
- ‘*deadweight*’: a sceptic might wonder if the observed impacts would not have happened anyway, for example just as a continuation of prior trends (see illustration, below)

<sup>20</sup> See for example the two Papers prepared by Eric Oldsman for the Donor Committee for Enterprise Development, which can be downloaded from <http://www.enterprise-development.org/groups/group.asp?groupid=5>



- *'Hawthorne effect'*: people who are being studied may change their behaviour – just because they are being studied<sup>21</sup>.
- *substitution* is similar, in considering the possibility that the treatment group changed its behaviour as a result of the prospect of treatment.

52. To illustrate the question of attribution, consider the example of a bridge built to link an island to the mainland; each span of the bridge has been built by a different agency – a good example of donor coordination. Once the bridge is completed, trade between the island and the mainland improves greatly, benefiting many islanders<sup>22</sup>. All 3 donors might claim to have achieved the total impact, since without their part of the bridge, there would have been no impact at all. Besides, others involved in boosting the trade, such as the banks and the State, deserve some of the credit. How much, therefore, of the total impact can each individual donor claim?



53. None of these aspects are trivial, and the frequent practice of measuring the performance of enterprises before and after the intervention is not

effective in addressing any of them (for example, their performance might have improved, but the performance of enterprises who did not receive any help might have improved even more). At the minimum, the team responsible for implementation should always be clear about their causal model: what the implementing agency expects to happen as a result of the intervention. This is explored in more detail in the following Section.

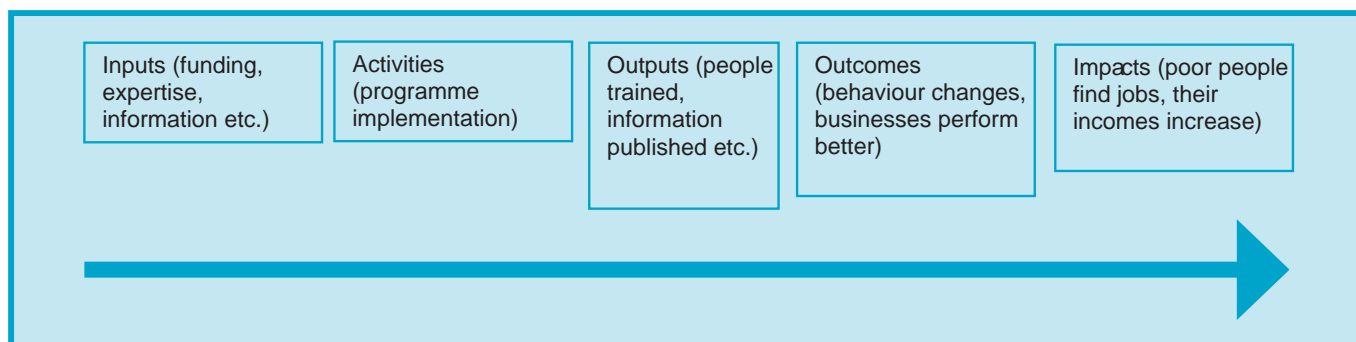
## The Causal Model

54. In the mind of any practitioner is an implicit causal model: the expectation that the various inputs (funding, expertise, information) will ultimately lead to certain developmental impacts, through a sequence of events. Each event or change is the direct result of the previous one; each step is required to achieve the next one, and (ultimately) the final impact expected.

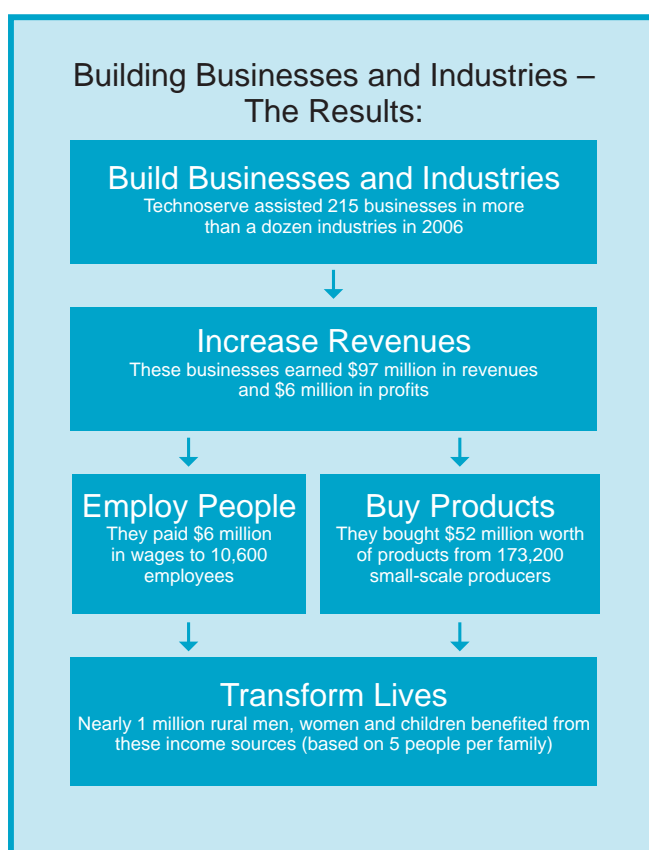
55. The following diagram illustrates a generic causal model, with each stage leading to the next; once this is made explicit, the measurement of results becomes more manageable. The practitioner only has to measure the achievement at each stage, and to show that the achievement was the result of the achievement of the previous stage.

<sup>21</sup> [http://en.wikipedia.org/wiki/Hawthorne\\_effect](http://en.wikipedia.org/wiki/Hawthorne_effect)

<sup>22</sup> Thanks to Peter Roggekamp for this example



56. Another illustration of a causal model is given below; this is an extract from the Technoserve Annual Report for 2006, reporting on impacts achieved<sup>23</sup>. This sort of reporting has probably contributed to the recent placing of Technoserve in the top 5 organisations, as rated by businesses for quality of partnership<sup>24</sup>; “companies rated Accountability and Execution as being the most important performance categories in their evaluation of partners”.



57. Historically, many evaluations have reported on outputs, but struggled to say much about either outcomes or impacts; the introduction of logical framework approaches (log frames) has been an attempt to require programme designers, at least, to be explicit about the sorts of outcomes and impacts that they were expecting. They also require some estimation of the likely chances of success, and specification of the possible risks at each stage.

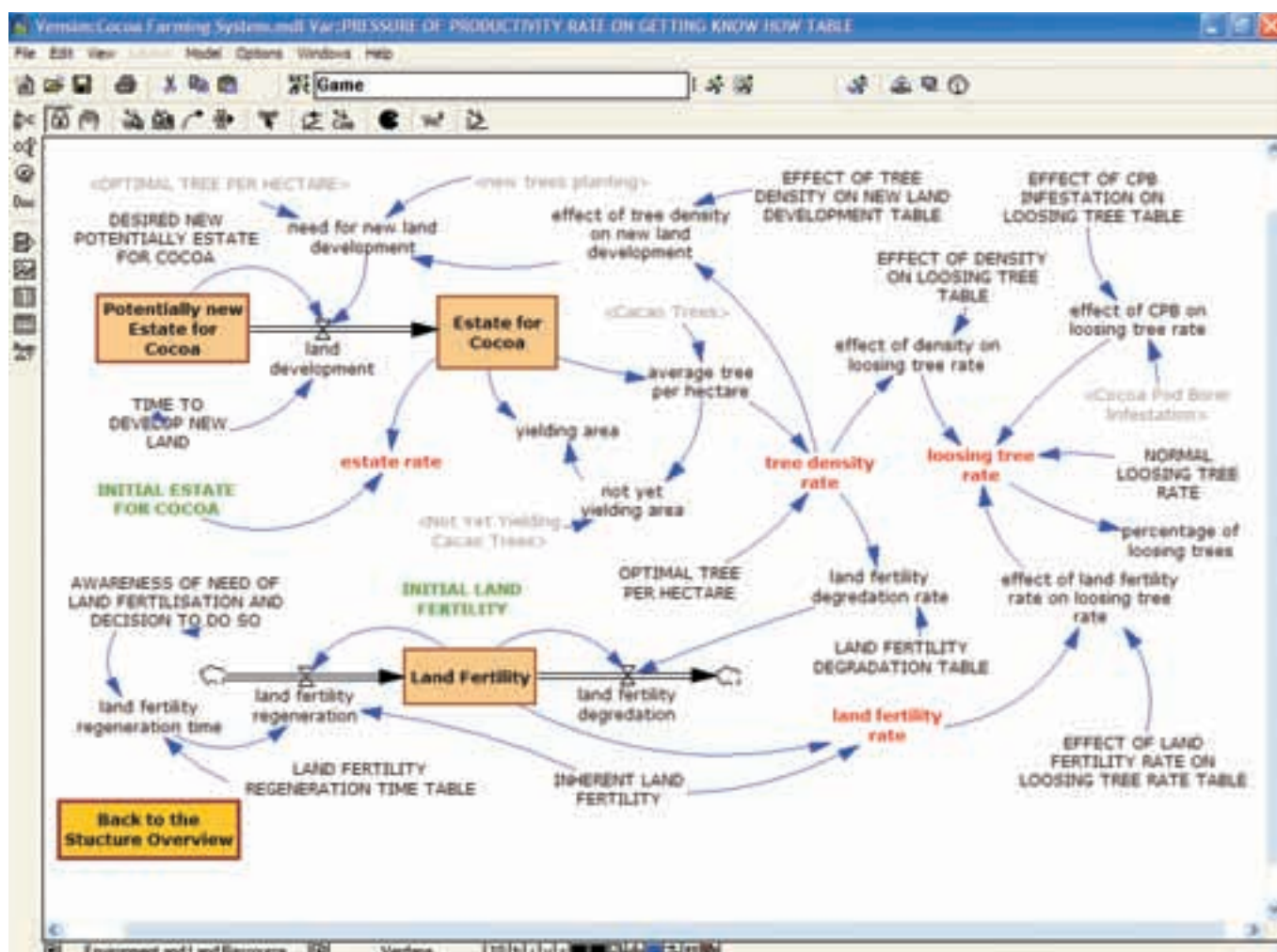
58. In practice, however, its potential complexity and rigour seem to have deterred some from using the log-frame as an implementation tool – even though it has certainly led to improvements in design in some cases. A simpler approach, that still required programme designers and implementers to be explicit about what they were expecting to see, may be more appropriate for wide application.

59. The causal model should provide the ‘backbone’ of any measurement of impacts, giving clear pointers as to what to measure, and when; it is, however, only valid, in the linear way outlined above, for discrete elements within an overall system (such as a value chain, or the private sector as a whole). Some have taken the argument further, to try to model an entire value chain; the illustration below was prepared using modelling software called Vensim, to model the cocoa value chain in Indonesia<sup>25</sup>. It was presented during the Chiang Mai Seminar by the author, who is currently working with Swisscontact.

<sup>23</sup> <http://www.technoserve.org/documents/TNS2006AR.pdf>

<sup>24</sup> FT, Dalberg and UN Global Compact [www.ft.com/reports/philanthropy2007](http://www.ft.com/reports/philanthropy2007) although note that “core business and advocacy partnerships are perceived to have higher impact than those focused on philanthropy”.

<sup>25</sup> Borer, Manfred, 2006. Application of System Dynamics in the International Development Cooperation. Analysis of decision taking among Cocoa Farmers in Flores. Thesis Business Process Manager, University of Applied Sciences, North-western Switzerland.



60. The important point about this model is not so much its predictive capability, but more its value in enabling staff to think through the logic of their interventions. The design of this kind of model requires those involved to be more explicit about what they see as the key influences during implementation, and their relative importance in achieving the desired outcomes. It does also allow the effects of important delays to be included in the thinking; in this particular case, for example, there is a delay between the planting of fresh cacao trees and the first harvest of cocoa, which needs to be taken into account when measuring impacts.

61. Once staff are more explicit about the causal model (or models) implicit in their work, then it is possible to validate this model, and measure the results, in more objective ways. Those ways are explored in more detail, in the following Section.

## Experimental and Quasi-Experimental Designs

62. The most rigorous approach to measurement is to establish a group of people or enterprises that are identical in every way to the treatment group, except that they do not receive the treatment (the 'control group'). For the sake of measurement, the samples of both treatment and control groups should be selected at random, hence the term Randomised Controlled Trials (RCTs).

63. RCTs are, from a theoretical point of view, the 'gold standard' or 'default option'; they represent a fully experimental design. In other words, given a substantial measurement budget and adequate expertise, all impacts should be measured through RCTs<sup>26</sup>. A truly randomised approach, with a convincing control group and sufficiently large

<sup>26</sup> See <http://www.povertyactionlab.com/research/rand.php> for a good description of the arguments for randomised trials, and for a rather extensive reading list on the subject.

## Management Training for Women Entrepreneurs in Vietnam

A relatively thorough example of the application of a causal model to measurement of impacts, on a limited budget, was conducted in Vietnam in 2002<sup>27</sup>. The study looked for changes in management practices, 4-6 months after delivery of a short training course in management skills – relative to a baseline and to a control group. It then looked for changes in business performance, for example through innovation, changes in sales and employment, and growth in income.

Having established changes in business performance, the study then investigated the developmental outcomes that could be resulting from these changes; since the trainees were women entrepreneurs, for example, the impact on gender relations was examined. The survey found a significant increase in incomes of the women, but the increase in employment was not significant, partly because the sample size was too small.

Admittedly, the research suffered from some significant shortcomings; for example, questionnaires were administered by mail, and the results were calculated only for those recipients who responded.

Nonetheless, the report shows that it is possible to develop a plausible story about likely impacts, on a rather limited budget, by carefully thinking through the implicit causal model.

samples, avoids the need for a baseline; statistical analysis can show, beyond reasonable doubt, that the benefits were due to the treatment, or support, provided.

64. There are, however, many reasons why it can be practically impossible to establish a control group of high quality; these reasons are discussed in the following Chapter. In such cases, one alternative is a 'quasi-experimental' approach, in which a control group is constructed to be as like the treatment group as possible – and then a random sample is taken from each group for measurement. In this approach, measurements are taken before ('baseline') and after treatment. This is therefore sometimes known as a 'differences of differences' method.

65. There are a number of sophisticated statistical methods that can be applied, to enhance the accuracy of this sort of approach, but they will not be explored in detail here. Further information is available from, for example, the Poverty Action Lab at MIT ([www.povertyactionlab.com](http://www.povertyactionlab.com)).

## Proxies

66. Finally, there is another promising avenue for measuring results at less cost than the direct measurement of the desired impacts: the measurement of proxies. Direct measurement of something like economic activity can be very costly, but there are indicators that might reasonably be expected to be closely correlated with economic activity – but which are much cheaper to measure; this might include, for example, traffic through the local bus park or electricity consumption locally.

67. These are known as proxies, and normally need to be validated – in other words, to be demonstrated to correlate closely with the indicator for which they are intended to be a proxy. This requires more rigorous and costly measurement, ideally using RCTs; such validation does not, however, need to be carried out every time.

<sup>27</sup> Voeten, 2002. Management Training Effects on Women Entrepreneurs who own and manage MSEs. Vietnam/Maastricht School of Management. [www.value-chains.org/dyn/valuechains/docs/250/TWMSE2%20working%20paper%203.pdf](http://www.value-chains.org/dyn/valuechains/docs/250/TWMSE2%20working%20paper%203.pdf)

## E. A Discussion of the Methodological Options

*The previous Chapter outlined the main tools available to people measuring impacts; it argued above all that implementers need to be clear and explicit about the 'causal model' on which their programme is based. The tools available to validate that model are discussed below.*

### Randomised Controlled Trials (RCTs)

68. Some academics argue that randomised controlled trials (RCTs) are the 'default option' – David Storey of Warwick University suggests that there is no middle ground between randomised trials and "asking a man in a pub". Strictly speaking, he is probably correct; the reality on the ground is, however, more nuanced, and RCTs suffer from a number of constraints and limitations.

69. A major constraint in practice is that RCTs are relatively expensive and technically demanding to do well; some argue that they are not even always necessary, since the attribution or causality may already be rather clear. The impacts of Alternative Dispute Resolution (ADR) in the Balkans illustrate this assertion well, in the sense that the causality is rather evident. It can be argued with reasonable plausibility that ADR has reduced business costs by over €42m<sup>28</sup>.

70. Another example is provided by the impacts of cell-phones for poor fishermen in India<sup>29</sup>; apparently, prices paid to the fishermen stabilised and became more predictable. In addition, reductions in wastage reduced the final price to the consumer; it seems clear that all of these effects can reasonably be attributed to the advent of cell-phones, without a pressing need to conduct RCTs.

71. Most importantly, RCTs can be problematic in interventions aiming to develop whole markets, and to build on commercial dynamics. Indeed, many of the examples used to illustrate the merits of RCT relate to delivery of public services; this is somewhat easier than PSD to control, and not subject to spontaneous replication or market development. But it is difficult to construct RCTs that take into account spontaneous replication outside the treatment group (and potentially into the control group); many interventions to develop markets are working hard to encourage such replications.

72. Also, recipients of the benefits of market development programmes are self-selecting, meaning that selection bias is very difficult to avoid. A further factor is that market development interventions are often several steps removed from the actual target group; the intermediaries are likely to be in the private sector, with cost structures close to the target group, and therefore to be poorly placed for major data-gathering and reporting exercises.

73. In addition, RCTs suffer from a number of other potential drawbacks:

- control groups are often 'contaminated' with treatment from other programmes and agencies
- some things cannot be randomised (e.g. exchange rate policy) and others may be hard to quantify (e.g. empowerment).
- there are examples of RCTs where errors were made in what to measure, assumptions of causality, etc.; the methodology may not be applied correctly (while still giving the illusion of precision).

<sup>28</sup> [www.businessenvironment.org/dyn/be/docs/141/Bentham.pdf](http://www.businessenvironment.org/dyn/be/docs/141/Bentham.pdf)

<sup>29</sup> <http://southasia.oneworld.net/article/view/128332/1>

74. There are also concerns about the Hawthorne effect: those randomly denied support (if they are aware of the situation) may either give up, or try harder, in order to prove the programme staff wrong. Either way, the measurement of impacts of the treatment would be skewed by the altered behaviour, and expertise is required to minimise such effects.

75. Some practitioners object to the randomisation of treatment, on the basis that it would not be ethical to refuse (or postpone) support to some people who were eligible for it, just for the sake of the experiment. However, it would presumably be even less ethical to provide assistance to businesses without being clear about the likely effects and impacts of that assistance. Randomisation in this case is probably the best way to assess those impacts.

76. The comments, above, about the cost of RCTs imply a question about the appropriate proportion of a programme's budget that should be allocated to measuring the impact; while this question is often asked, there is regrettably no single answer. There may be cases where an expensive study is fully justified, for example in validating the impacts of a new approach, or in validating a proxy. In general, however, there is a feeling that measurement costs should not greatly exceed 10% of the overall programme costs.

77. Given also that relatively few programmes anyway seem to have the resources to conduct full RCTs, what are the other options for measuring impacts?

## Alternatives to Experimental Designs

78. Many practitioners favour a qualitative approach, exploring in detail what has actually happened as a result of the intervention or treatment. This exploration is necessarily somewhat open-ended, probably identifying both positive and negative outcomes. There are many case studies written along these lines, and in practice, such an

approach will always be needed – at the minimum, to complement, and to make sense of, any quantitative measurements.

79. A core question is who is involved in preparation of such case studies; often, they are prepared by the programme staff themselves, which means that they reflect an intimate knowledge of the programme, but may not be perceived as impartial or objective. Case studies prepared by outsiders may be more credible, especially if the authors have not been commissioned and paid by the programme itself.

80. Some favour participatory approaches, whereby the opinions of those affected by the programme are key to the evaluation. In theory, and particularly from an ideological point of view, this is not just desirable, but necessary – to be able to interpret the data, and as an input to the design of future programmes. But participatory approaches on their own are not a substitute for – or morally superior to – more quantified approaches, as they do not allow for any objective comparison beyond the programme itself. They also give little indication of progress in alleviating poverty, or in improving people's lives.

81. Nonetheless, league tables can be generated from such perceptions, in some circumstances; in business environment reform, for example, Transparency International rank the level of corruption in a country according to perceptions of respondents. Similarly, the Business Environment and Enterprise Performance Survey (BEEPS) of the World Bank and EBRD measures perceptions of business people about the business environment, and generates a ranking based on that.

82. Such an approach would be in contrast, for example, to the Cost of Doing Business survey, which aims to quantify parameters such as the number of days to register a business, according to a standard methodology.

83. Another example of a subjective ranking, already referred to above, was produced recently by the Financial Times, Dalberg and the UN Global



Compact<sup>30</sup>, who ranked non-profit organisations according to the perceptions of businesses about how effective they were. Four criteria were applied: accountability, adaptability, communication and execution. The top organisations globally were voted as shown in the table below.

84. Such approaches are important, and are mentioned here, partly because they often attract extensive media coverage; technocratic discussions about evaluation methodologies, in contrast, are far from the public eye, with the associated risks and benefits. But subjective rankings raise a number of important methodological issues; for example, it is not very clear from the survey reported above which businesses and which non-profit partners were included (the list does not seem to be comprehensive).

85. Also, such rankings do not necessarily tell us much about the relative cost-effectiveness of a programme; if, for example, it was making very generous hand-outs, one would expect the programme's recipients to have a very positive

perception – even if the impacts in the longer term are negligible, or even negative.

## Agency Rankings of Own Programmes

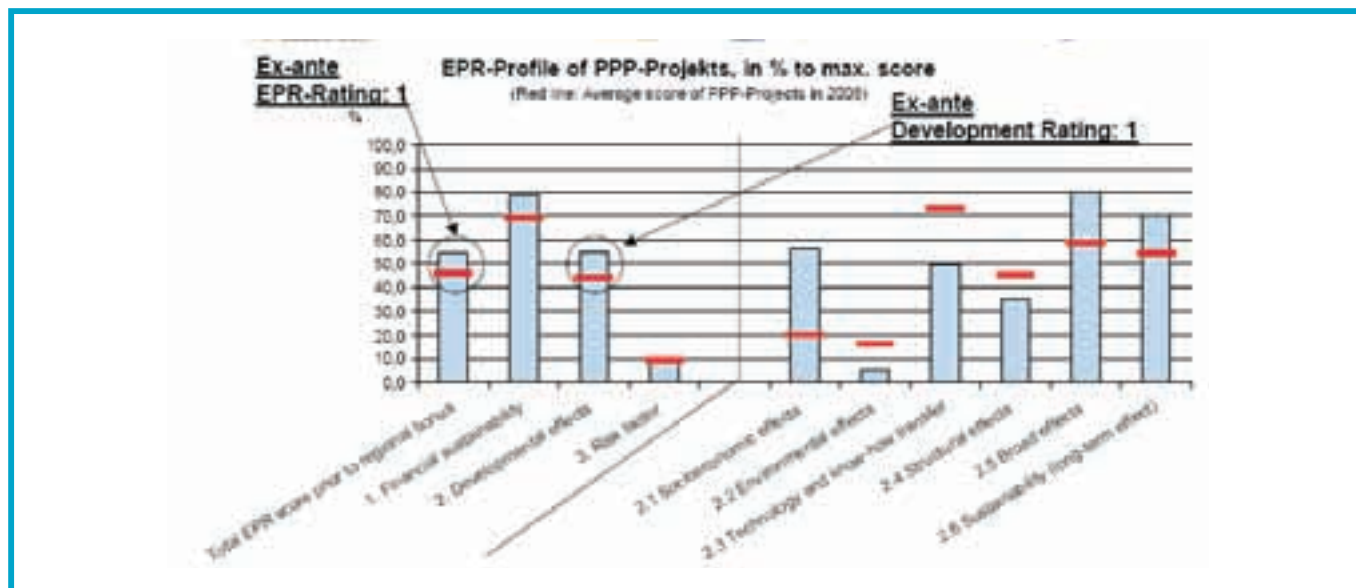
86. In the absence of agreement about methodologies, some agencies resort to a somewhat subjective overall rating of their programmes. Development banks in particular often rate their programmes by the perceived degree of success, giving a headline verdict – for example between 'Very successful' and 'Moderately successful'; the overall portfolio of programmes may achieve success rates of 80% or more. Given the contexts in which the programmes operate, this seems optimistic.

87. Indeed, success rates of commercial investments are often little more than 50%, as staff have to beat the market average<sup>31</sup>. In venture capital, success rates are usually 20% or less, with the big successes compensating for the under-performers.

Rank	Organisation	Country of HQ	Accountability	Adaptability	Communication	Execution
1	Lions Club International	USA	5.0	4.8	4.3	4.3
2	Environmental Defence	USA	4.7	4.3	4.9	4.4
3	WRI	USA	4.8	4.2	4.7	4.4
4	TechnoServe	USA	4.6	4.6	4.8	4.2
5	Rotary International	USA	4.6	4.5	4.3	4.4
6	GBC HIV/AIDS	USA	4.3	4.4	4.1	4.5
7	Conservation International	USA	4.4	4.5	4.3	4.2
8	WWF	Switzerland	4.4	4.2	4.3	4.4
9	CARE	USA	4.6	4.0	4.4	4.3
10	Greenpeace	Netherlands	4.5	4.3	4.3	4.0

<sup>30</sup> FT, Dalberg and UN Global Compact [www.ft.com/reports/philanthropy2007](http://www.ft.com/reports/philanthropy2007)

<sup>31</sup> 50% of investment projects evaluated by WBG's IEG were successful at a real ERR of about 15%. [www.mmw4p.org/dyn/mmw4p/docs/585/Michelitsch.pdf](http://www.mmw4p.org/dyn/mmw4p/docs/585/Michelitsch.pdf)



And the social aspects of development programmes are probably even more challenging than the market-related ones. Credibility therefore becomes the core problem; self-assessments of success are not convincing; systematic judgements that include subjective elements are not convincing either.

88. One illustration of this is provided by the evaluations of the training voucher scheme of the World Bank in Kenya. Initially, the Bank Performance in this scheme was rated as Highly Satisfactory in the ICR Review, but a subsequent and more detailed PPAR evaluation<sup>32</sup> by the Operations Evaluation Department (now the IEG) in 2005 down-graded that rating to Unsatisfactory – a major difference of opinion.

## Building Systems for Management

89. Most evaluation methodologies present a list of indicators that can or should be measured, either leaving the evaluator to choose which ones to use, or suggesting a suite of indicators which, taken together, give a good picture of the extent of the achievements. Various organisations are experimenting with – or have already adopted – interfaces that allow staff to use a range of performance-related data. DEG, for example

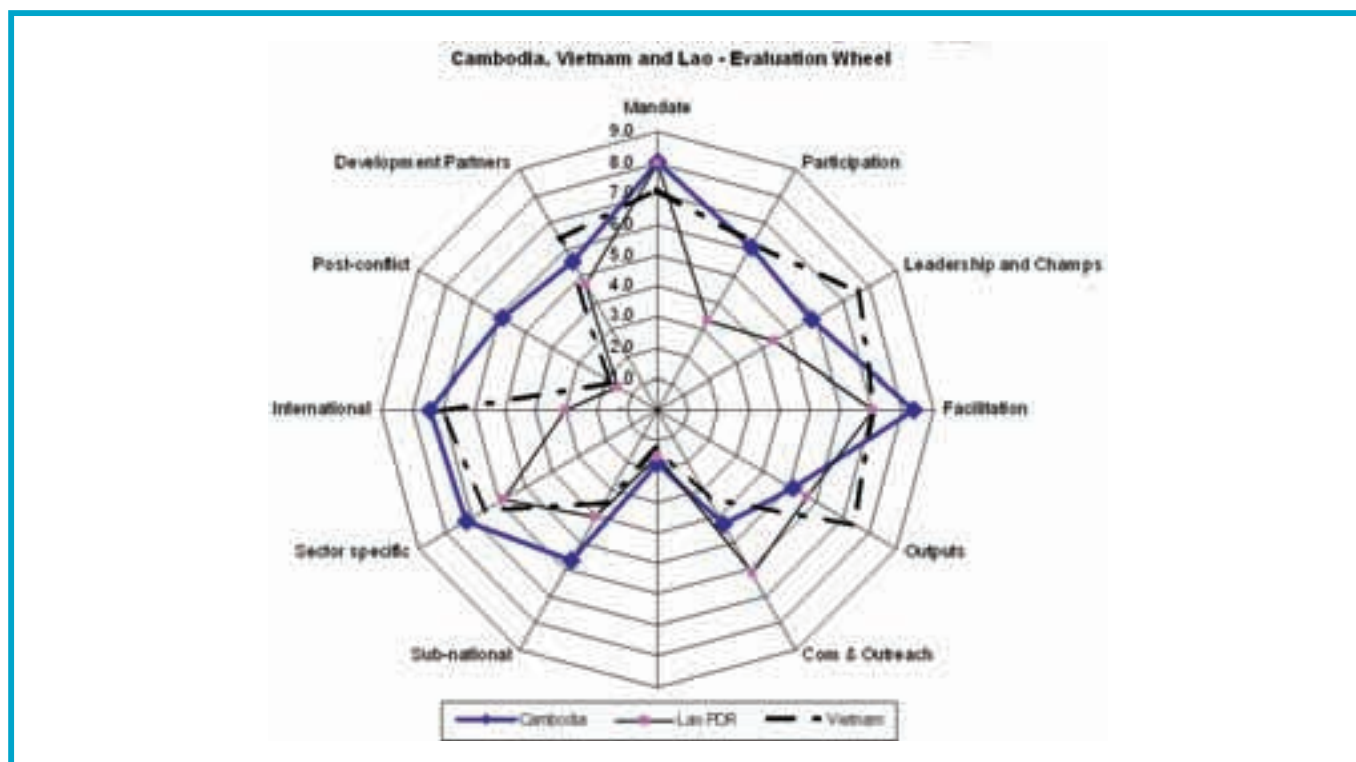
(see illustration), has adopted such an approach to indicate at a glance the relative health of its portfolio, according to various attributes.

90. The Acumen Fund is also working with private-sector partners such as Google.org to develop a system for social investors. While initially designed for use by Acumen, it may also in the future be possible for several social investors to access the same system. This would enable them to coordinate their support for individual social entrepreneurs, for example, and to share common metrics – as part of a wider conversation between social investors around performance evaluation and impact assessment.

91. At the time of writing, Acumen has not yet decided which developmental indicators to measure across the whole organisation, beyond the business and operational ones. They might, for example, relate the 'increase in number of lives impacted' to 'increase in net income', to see whether the outreach is increasing over time.

92. Both of these systems require staff to quantify approximately the likely development-related outcomes of each set of activities; the resulting scores allow managers to generate up-to-date visual displays of the overall state of the current agency portfolio. Above all, they are relatively easy to use, and are therefore likely to be used in practice.

<sup>32</sup> World Bank, 2005. Project Performance Assessment Report: Kenya MSE Training and Technology Project. [www.mmw4p.org/dyn/mmw4p/docs/178/Kenya%20PPAR%20final.pdf](http://www.mmw4p.org/dyn/mmw4p/docs/178/Kenya%20PPAR%20final.pdf)



93. This approach is also realistic, in the sense that reality is multi-faceted, and limiting measurement to a few numbers is always going to strip out important detail. One example of this approach is currently being applied to measure the success of Public Private Dialogue (PPD), using an evaluation 'wheel' (see illustration above); this graphical approach enables comparison of achievements across programmes and countries<sup>33</sup>.

94. The methodology does also report on quantified outcomes, for example in terms of private-sector savings (\$310m), and the implicit return

on investment (\$291 per dollar invested) through PPD in the Mekong region. An additional, important aspect of this methodology is that it considers how PPD programmes and countries evolve over time, and the phases through which they pass in that evolution – a dimension that is often ignored.

95. All of the approaches mentioned in this Chapter are briefly summarised in the following Table; in practice, of course, they are not mutually exclusive, but likely to be complementary, and best used in combination, depending on the circumstances and budget.

Approach	Advantages	Disadvantages
Experimental designs	Credible	Demanding, in cost and expertise
Rankings of perceptions	Attract media attention	Dependent on methodology; subjective
Agency rankings	Simple internal tool	Susceptible to internal influences
Systems with graphical interface	Easy to read, therefore likely to be used	Not in themselves rigorous or reliable

<sup>33</sup> [www.businessenvironment.org/dyn/be/docs/141/Herzberg.pdf](http://www.businessenvironment.org/dyn/be/docs/141/Herzberg.pdf) See also [www.PublicPrivateDialogue.org](http://www.PublicPrivateDialogue.org)

## F. Indicators ('What to Measure')

*The previous Chapter outlined the main methods in use to measure impacts – the 'how to measure' question; it argues that there is no one method that will always be most appropriate, and a combination is required. This Chapter focuses on the choice of indicators, or 'what to measure'.*

96. The choice of what to measure, or 'indicator', is perhaps the most important in the development field, since it implicitly defines the priorities of the intervention. There is, therefore, an immediate split between some agencies that focus on poverty alleviation directly, and other agencies that give priority to overall PSD.

### Poverty Alleviation or Economic Growth?

97. The indicators will be different; in the former case, agencies are likely to want to measure the number of people lifted out of poverty, for example in relation to the Millennium Development Goals. The latter agencies are more likely to measure indicators like private-sector savings, or the volume of investment that they have leveraged from the private sector – on the assumption that these will lead to pro-poor PSD.

98. This Reader will tend to focus more on the former: the extent to which development activities are alleviating poverty over time. This is partly because it is more difficult, methodologically, and partly because it probably addresses the urgent agenda outlined in previous Chapters – for example around the need of tax-payers in donor countries to understand more about what is being achieved with 'their' money.

99. Even with a focus on poverty alleviation, there are several divergences of view. Governments in developing countries, for example, will often place high priority on the generation of employment opportunities, since half of their population are typically children who are growing up and who will soon be looking for jobs. Social stability depends, in this line of thinking, on the private sector generating job opportunities for many of these young adults, quickly.

100. The measurement of jobs is, however, complex in the poorer countries, since employment is often highly seasonal (in agricultural economies) and is usually informal (without an employment contract, for example). It is also often part-time; finally, it may be unpaid, or paid in kind. The employee may be performing the job because of the lack of alternatives, or in the hope of accruing goodwill and social capital – including reciprocal support in the future.

101. Labour statisticians tend to count people as employed if they are not unemployed – in other words, if they have worked for at least one hour during the reference period (typically one week); that is not particularly helpful for practitioners aiming to generate and measure additional employment opportunities. Given the part-time nature of many of those opportunities, they are often added together, to constitute Full-Time Equivalent (FTE) jobs – in broad terms, two half-time jobs would make one FTE job, for example<sup>34</sup>.

<sup>34</sup> See [http://en.wikipedia.org/wiki/Full-time\\_equivalent](http://en.wikipedia.org/wiki/Full-time_equivalent) for more information.

## Measuring Programme Costs

102. The above discussion is essentially about the benefits of an intervention, however defined; ultimately, they should be related to the costs, in order to arrive at a cost-benefit ratio. Although the costs should in principle be an easier figure to obtain, it is not often published.

103. Ultimately, it would be helpful to reach agreement about which costs should be included; how should programme management costs, for example, be pro-rated across a wide range of activities? Should managers keep timesheets to enable this figure to be calculated? Should the costs to the government, the beneficiaries and others be factored in?

104. These questions have not been addressed in any systematic way, although some of the studies outlined later in this Reader do consider them. Therefore, the following Chapter focuses in more detail on indicators of poverty, with particular reference to the Millennium Development Goals (MDGs).



Photo: Charles Bodwell

## G. PSD and the MDGs

*The previous Chapter argued that the choice of 'what to measure' depends on the core priority for the programme; many donors are increasingly feeling that they will need to have a response to the Millennium Development Goals<sup>35</sup> as the deadline of 2015 approaches. Since PSD potentially cuts across many MDGs, they are discussed in more detail, below.*

105. The most important MDG for PSD practitioners is probably MDG 1, to "reduce by half the proportion of people living on less than a dollar a day". Importantly, few agencies are currently measuring this indicator in their PSD programmes, for a range of reasons. One is that it is difficult for practitioners in the field to implement the 1993 level of purchasing power parity (PPP) concept of the dollar in the MDG. Another is that many people live in communities that straddle this income level, in a state of flux just above or just below it; capturing the precise transition in a credible way is tricky.

106. Some therefore prefer to measure those living on less than \$2 per day, using a dollar measurement that can be taken from field measurements (rather than PPP calculations). However, such complexities again may not offer the 'sound-bite' that journalists and others are looking for, when reporting on the achievements of development agencies. Many agencies are therefore actively exploring what the term 'poverty' means, depending on the local context – and in ways which can be measured.

107. USAID has been mandated by the US Congress to ensure that at least 50% of all USAID micro-enterprise funds benefit the very poor. To provide a check on whether this mandate is being met, Congress has more recently instructed USAID

to develop low-cost methods to identify the poorest households, and to require that USAID-supported micro-enterprise programs use these methods to assess how many of their beneficiaries are very poor<sup>36</sup>.

108. After exploring some generic methods to identify the very poor, it became clear that the characteristics of households living in extreme poverty vary from one country to another. More focused tools are therefore being developed, to take account of local context.

109. Similarly, CGAP, Grameen and Ford are proposing a tool called the Progress out of Poverty Index (PPI), to be used over time to determine improvements in client economic levels and their ultimate graduation out of poverty. These would be country-specific; in the example of the Philippines, they would include the materials used in house construction, the type of toilet, ownership of a gas stove, children in school and number of televisions owned. The aim is that such baskets of indicators can be used as proxies for income levels in the household<sup>37</sup>.

110. The Performance Measurement Framework (PMF) was included in the Donor Committee's Guiding Principles on Business Development Services, primarily in the form of a list of indicators<sup>38</sup>. These indicators were essentially proxies for ultimate impact; in other words, they were based on an assumption that increased purchase of business services (for example) was positively correlated with business growth, increases in incomes and/or employment etc. In practice, validation of these proxies would require substantial investment, one which was never actually made.

<sup>35</sup> <http://www.un.org/millenniumgoals>

<sup>36</sup> See [www.povertytools.org](http://www.povertytools.org) for more information.

<sup>37</sup> CGAP Focus Note number 41, May 2007. Beyond Good Intentions: Measuring the Social Performance of Microfinance Institutions. [www.cgap.org/portal/binary/com.epicentric.contentmanagement.servlet.ContentDeliveryServlet/Documents/FocusNote\\_41.pdf](http://www.cgap.org/portal/binary/com.epicentric.contentmanagement.servlet.ContentDeliveryServlet/Documents/FocusNote_41.pdf)

<sup>38</sup> <http://www.enterprise-development.org/groups/group.asp?groupid=3>

111. Arguably, though, the challenges of measurement in PSD are likely to become more relevant in future, in other developmental disciplines, like education and health; increasingly, policy makers are realising that the private sector is already playing a central role in the provision of many basic services. The winning essay in the 2006 FT/IFC competition, for example, makes a very convincing case for the major role played by the private sector in many countries, in the provision of education<sup>39</sup>:

112. “Recent research has found a large majority of schoolchildren in selected poor urban and peri-urban areas of India and Sub-Saharan Africa using private schools, while in rural India, half of all schoolchildren are privately enrolled. Even in impoverished rural China large numbers of private schools exist off the official radar.

113. “The research showed that private schools for the poor are superior to government [schools:] schoolteachers are more committed, the provision of important inputs better, and education outcomes better even after controlling for background variables. All this is accomplished for a fraction of the per-pupil teacher cost of government schools.”

114. Others note a comparable situation in health care; rural dispensaries are often the first point of contact for people in poverty seeking primary health care. Even when people arrive at government facilities, the situation may not be very different from that in the private sector; “the problem with the [government] primary health centres [in Uganda] is that you have to give money to each and every health worker that attends you . . . but in private clinics or hospitals, you just pay once’. Another person said, ‘the treatment you get in the primary health centre is in relation to the amount of money you pay’ and they ‘are the same as the private clinics because in both places, you need to pay’.”<sup>40</sup>

115. Yet others argue that the private sector can help to meet the MDGs in areas such as rural sanitation<sup>41</sup>; however, such observations remain controversial for the time being, and mainstream opinion perceives publicly-funded provision of health and education services to be the key to meeting the MDGs in those sectors. Therefore, this Reader will continue to focus on the development of value chains and service markets.

<sup>39</sup> [www.ifc.org/ifcext/economics.nsf/AttachmentsByTitle/Competition\\_booklet2006/\\$FILE/2006essaybook.pdf](http://www.ifc.org/ifcext/economics.nsf/AttachmentsByTitle/Competition_booklet2006/$FILE/2006essaybook.pdf) or [www.ifc.org/competition](http://www.ifc.org/competition)

<sup>40</sup> The quality of care by private practitioners for sexually transmitted diseases in Uganda, by Walker, Muyinda, Foster, Kengeya-Kayondo and Whitworth. Health Policy and Planning, Oxford University Press, 2001.

<sup>41</sup> See for example “Harnessing Market Power for Rural Sanitation”. Water and Sanitation Program Field Note, World Bank / IDE, February 2005. <http://www.mmw4p.org/dyn/mmw4p/docs/419/VN-IDE-masons.pdf>

## H. Other Measurement Issues

*All of the previous Chapters focus on the methodological challenges of measuring impacts; the area is, however, multi-faceted, and there are aspects of the process that cannot be addressed, for example, only by hiring expert consultants. These aspects are touched on below.*

### The time factor

116. Development practitioners are famously in a hurry; they typically have 'windows' of 2-3 years in which to prove both their programmes and themselves. Yet many worthwhile changes probably take longer than that; for example, it is estimated that the adoption of ox-plough technology took about 70 years, in the Machakos District of Kenya<sup>42</sup>. Similarly, "it took nearly 100 years from the days of Henry VII for Britain to catch up with the Low Countries in woollen manufacturing."<sup>43</sup>

117. At the other end of the technology spectrum, "it took Nokia 17 years to earn any profit from its electronics subsidiary, which is now the biggest mobile phone company in the world". Similarly, "it took Toyota more than 30 years of protection and subsidies to become competitive in the international car market, even at the lower end of it. It was a good 60 years before it became one of the world's top car makers."<sup>44</sup>



Photo: Charles Bodwell

118. Another aspect of this dimension is that effects may appear negative in the short term; for example, enterprises may rationalise and shed labour in the short term, in order to be more competitive in the medium term. A programme which had as its ultimate performance indicator the number of jobs created would therefore appear to have failed in the short term, even though the remaining jobs (and future jobs to be created) may be more secure and better paid.

119. The difference in timescales presents a major challenge, particularly for practitioners who are aiming to develop whole markets, systems and economies. The concept of developing markets in a systemic way is premised on greater impacts in the long term, through sustainable change. Arguably, the only way to 'square the circle' is to invest time in projecting the anticipated impacts in the coming years, well beyond the life of the programme.

<sup>42</sup> Mortimore, Michael, and Kate Wellard, January 1991. Environmental change and dryland management in Machakos District, Kenya, 1930-1990. ODI Working Paper no. 57, London.

<sup>43</sup> Chang, Ha-Joon, 2007. Bad Samaritans: Rich Nations, Poor Policies and the Threat to the Developing World. Random House, London.

<sup>44</sup> Op. Cit.



## Human and Institutional Factors

120. Development is implemented by people, whose behaviour is moulded by incentives, abilities and beliefs; these are, however, rarely considered in discussions about the measurement of results. Gary Woller, in a recent USAID discussion forum<sup>45</sup> on this subject, noted:

121. "There are indeed strong disincentives to conduct impact assessments. There is the free rider problem ('why should I assess my program, let so-and-so assess theirs?'), a public goods problem ('why should I invest my money in impact assessment when others benefit from it, let someone else invest in it'), a career problem ('what if I do an assessment and the results are not favourable, how will that affect my career?'), a cost problem ('gee, impact assessment sure costs a lot, better to save the money and channel it to program beneficiaries'), and so forth."

122. At a more personal level, people who are good at getting results are often impatient with the idea of rigorously measuring them. They think of themselves as action-oriented, rather than as statisticians; the inertia to carry out rigorous measurement can be a function of personality as much as anything else.

123. There is also the issue of organisational structure, and the incentives that are implicit in that structure. Paul Gertler of UC Berkeley, for example, has noted that individual programme designers are rarely also the implementers; indeed, many donors forbid consultants responsible for programme design from participating in the implementation. Should the programme manager be penalised for the poor results achieved by faithfully implementing a poor programme design?

124. Sendhil Mullainathan of Harvard University notes that often the agency task manager is responsible overall for the design, implementation

and evaluation; this implies incentives to only evaluate programmes with positive outcomes, or those that the agency wants to stop. It also implies possible penalties for innovation and experimentation; in this perspective, the challenge is not so much to do with evaluation methodologies, but more about the way in which many agencies are structured.

125. Indeed, more broadly, field-based staff of many agencies wonder whether their HQ-based colleagues really understand the local situation, with all its complexity and challenges. Pressure from HQs to measure performance can therefore be interpreted as a way for senior managers to control and steer field-based staff, in ways which would only be acceptable if the methodology used was also acceptable. There is, therefore, often considerable resistance by field staff to measuring results.

126. As an aside, development interventions are universally based on the idea that there is 'a problem', to which the intervention is intended to be 'the solution'. Within this framework, it is difficult to consider opportunities taken or passed up; this probably contributes to the 'heaviness' in most programme evaluations, since creativity and added value may not be much appreciated unless they are clearly within the pre-determined framework. This makes it more difficult for field staff to perceive evaluations as potential learning experiences.

127. Intended to be slightly humorous, the list on the following page is often referred to by Danida's Evaluation Department; it illustrates nicely the resistance often found at the field level to evaluations.

128. If all of these objections can be overcome, the funds can be found, and the evaluation put in motion, the standard procedure in most agencies is to hire an external consultant (or team of consultants) to come in and conduct the evaluation.

<sup>45</sup> Impact Assessment: An Online Speaker's Corner Discussion, September 2006. [http://www.microlinks.org/ev.php?ID=12730\\_201&ID2=DO\\_COMMUNITY](http://www.microlinks.org/ev.php?ID=12730_201&ID2=DO_COMMUNITY)

## Standard Arguments to Avoid Evaluation

As part of our continuing efforts to facilitate the work of busy managers, we submit herewith a selection of some of the most common responses to proposals that an evaluation be conducted. Please note that none of these arguments will be accepted by the Evaluation Department.

1. Our project/programme is different.
2. It will cost too much.
3. We don't have the time.
4. The project (activity) is too small.
5. It wasn't in the implementation plan.
6. We've never done it before.
7. The government (institution) won't like it.
8. Give me the money.
9. We don't have the authority/responsibility.
10. There is no need for an evaluation.
11. It's an ivory tower exercise.
12. Let's get back to reality.
13. It's not our problem.
14. Why change it; it's working all right.
15. We're not ready for it yet.
16. It isn't in the budget.
17. The desk officer/TSA Advisor/Govt. representative /CTA/Counterpart is new/has recently been changed.
18. The desk officer/TSA Advisor/CTA/Counterpart has left.
19. The Director/CTA/Counterpart has not yet been appointed.
20. The counterpart staff is still in training/ on fellowships.
21. We/re doing all right without one.
22. It's never been tried before.
23. There must be an ulterior motive.
24. Who's trying to teach me my job?
25. That may work in any other organisation/ region/country/technical field but it will never work here.
26. I'm not convinced it will help.
27. "They" just want to "get us".
28. Think of the disruption it will cause.
29. It can't be done objectively.
30. It's too much trouble to change.
31. We've always done it this way.
32. We did what we said we'll do.
33. We executed what was in the Programme Document.
34. We have already been evaluated.
35. We don't have any problems.
36. There's been a change in the government.
37. The financial crisis put us behind schedule.
38. We were just audited.
39. The Ambassador says it's one of his/her best projects.
40. It's a pilot project.
41. It's a model project.
42. It's an experimental project.
43. The project is too young.
44. The project is almost over.
45. Construction has not yet been completed.
46. The equipment has not yet arrived.
47. The equipment has not yet been installed.
48. Legal status has not yet been provided.
49. We can't find the original workplan.
50. I wasn't the responsible officer when the project started.
51. The government is satisfied with the project.
52. The government hasn't yet supplied the inputs.
53. The project is not evaluable.
54. We don't have the data.
55. The project design is too vague.
56. It's a local holiday.
57. It's the rainy season.
58. Let the Auditor General do it.
59. Outsiders won't understand the complexities.
60. We evaluate all the time ourselves.
61. I'm due for home leave.
62. We are already planning the next phase.
63. Our colleagues in the SWAP won't like it.
64. We are preparing a PCR.

While this has the advantage of probably being somewhat impartial and objective, it often means that the consultants have to learn rapidly about all aspects of the programme.

129. Their Terms of Reference typically require them to assess the programme against the original project document and log-frame (if there was one) – although sometimes the consultants are also asked to assess the programme against current understanding of best practice (which may have changed substantially since the project document was written).

130. In these circumstances, the consultants will often conclude that, without baseline data, they cannot provide a meaningful assessment of the results. However, they may also feel that they need to find significant shortcomings in the programme, in order to justify their fee – again leading many practitioners to dread any kind of evaluation.

131. A recent World Bank report on ‘How to Build M&E Systems to Support Better Government’<sup>46</sup> talks of “carrots, sticks, and sermons. Many of these incentives have been used to help institutionalise M&E in developed and developing country governments. Carrots provide positive encouragement and rewards for conducting M&E and utilizing the findings. They include, for example, public recognition or financial incentives to ministries that conduct M&E.

132. “Sticks include prods or penalties for ministries or individual civil servants who fail to take performance and M&E seriously—such as financial penalties for ministries that fail to implement agreed-on evaluation recommendations. Finally, sermons include high-level statements of endorsement and advocacy concerning the importance of M&E. They also include efforts to raise awareness of M&E and to explain to government officials what’s in it for them.”

<sup>46</sup> See [www.worldbank.org/ieg/ecd](http://www.worldbank.org/ieg/ecd) for more information. Table 11.1 (pages 63-64) in the report gives many examples of each of these three categories; most of the examples would be rather relevant for development agencies too.

# I. The Benefits of Approximate Measures

*The previous Chapters have explored many of the complexities and layers of the topic of measuring results; it is time now to consider possible solutions, and ways to avoid paralysis in the face of such daunting challenges.*

133. The most plausible conclusion of this discussion is that ‘better is the enemy of good’; the cost and effort to achieve rigour seem so daunting that practitioners finally measure nothing. Sendhil Mullainathan notes that all numbers created the “illusion of precision”, whereas they are often highly subjective – or are based on assumptions that are subjective.

134. Subjective discussions are not in themselves a problem, as long as people realise that it is possible to have quantitative, subjective discussions that can be approximately correct. For example, if the approximate impact of one programme is apparently 100 times the approximate impact of another one, then it is probably better; such differences are already apparent in some cases. Simple calculations of the cost per person reached, for example, might already permit the screening out of some programme designs.

135. If given indicators could be measured in ways that could be benchmarked and compared, this would inevitably shape the incentives for, and behaviour of, agency staff. Some argue that this would encourage too narrow a focus on achieving specific targets; since the world is so complex, this would probably lead to counter-productive responses.

136. Indeed, mainstream discussions tend to include thoughts of a menu of indicators, from which development practitioners can choose at will, according to their circumstances and priorities. While more theoretically sound, this approach suffers from a number of disadvantages.

137. One is that it exposes agency staff to a long list of objectives and priorities; some of these overlap, but others may involve trade-offs. At present, there is no way to choose between objectives, with the result that programme choices tend to be either based on personal preferences or on perceived opportunities arising.

138. On the other hand, there would be three major advantages to agreeing on what to measure, if that agreement could be around a very small number of indicators; no amount of qualitative or descriptive evaluation will address any of these three opportunities (and increasingly, urgent priorities):

## **Bulking Up Programme Achievements Across the Whole Agency**

139. By using a few, comparable indicators, development agencies could add together the estimated impacts of all of their PSD interventions, to give their managers some headline numbers about the impacts they are achieving.

140. Parliaments in donor countries are increasingly asking what their development agencies have achieved recently – and are getting ‘headline’ answers that they can use, in the fields of education and health. Indicators like numbers of girls educated in school, and numbers of children immunised, while essentially outputs, conjure up images of pro-poor outcomes that parliamentarians and voters can identify with.

141. The field of Private Sector Development (PSD), meanwhile, is at best reporting on outputs that do not particularly excite the electorate: ‘number of days to register a business’, or ‘increase in sales’, do not necessarily paint a picture of poor people finding new opportunities.

142. The Paris Declaration stresses that the priorities established by partner governments should dictate what donors fund; one implication of this is that every country programme is designed according to a different logic; the 'causal model' is different, and in practice, therefore, the indicators used are also different – since there is no consensus at present in the PSD community about what indicators to use.

143. Bilateral donors in particular tend to work on a government-to-government basis, with country programmes being shaped by dialogue and local demand. This trend is amplified when each programme is evaluated against the original programme document (as in most agencies), without much reference to any agency-wide framework or set of indicators.

144. The IFC has been leading some of the thinking in this area<sup>47</sup>. It is already able to report, for example, that its investments last year benefited 2.4m patients, 320,000 students, 12m power customers and 9m water customers. It is also introducing a Development Outcomes Tracking System (DOTS), anticipating that it will guide future investments.

145. In Corporate Governance, for example, the organisation is now tracking the number of legislative acts adopted as a result of advisory services provided (25 last year). Other data from 2006 include investment enabled (\$1,366m) and the leverage achieved (\$68 per dollar of funding for services)<sup>48</sup>.

146. These numbers are particularly interesting, as they enable the IFC to present its achievements in a clear and brief format; on the other hand, they do raise significant questions of attribution – particularly in a world where joint programmes and inter-agency collaboration are (hopefully) becoming more common.

## More Rational Funding Choices

147. If there were agreement around a very few indicators of impact, anticipated impacts could be estimated *ex-ante* in ways that were comparable across programmes. Funding agencies could then ultimately make more rational funding choices between different programmatic opportunities, depending for example on local conditions.

## Comparison Between Programmes and Agencies

148. A more disruptive aspect of adopting a few indicators on an inter-agency basis would be that comparison between programmes and agencies might then become possible. Michael Klein, for example, currently the joint IFC-World Bank vice president for PSD, has called for more competition within the aid industry, based on the measurement of results; others, though, have expressed concern about the transaction cost of this approach<sup>49</sup>. Also, this is likely to lead to competition for key local staff, further weakening local government; "declines in bureaucratic quality are associated with higher donor fragmentation"<sup>50</sup>.

149. More importantly, the idea that results can be measured and compared across agencies, countries and time implies a mechanistic model: 'given equal conditions, the same inputs should lead to the same results'. In practice, of course, no two sets of conditions are equal; even if circumstances within a given country could be controlled, it still operates in a global business environment that is constantly changing in very significant ways.

<sup>47</sup> Programme and selected presentations from recent IFC workshops on this theme can be found at [www.value-chains.org/dyn/valuechains/bdssearch.details?p\\_phase\\_id=585&p\\_lang=en&p\\_phase\\_type\\_id=4](http://www.value-chains.org/dyn/valuechains/bdssearch.details?p_phase_id=585&p_lang=en&p_phase_type_id=4) and [www.businessenvironment.org/dyn/be/besearch.details?p\\_phase\\_id=141&p\\_lang=en&p\\_phase\\_type\\_id=4](http://www.businessenvironment.org/dyn/be/besearch.details?p_phase_id=141&p_lang=en&p_phase_type_id=4)

<sup>48</sup> See the following presentation for more examples, also for different disciplines. [www.businessenvironment.org/dyn/be/docs/141/Batra.pdf](http://www.businessenvironment.org/dyn/be/docs/141/Batra.pdf)

<sup>49</sup> See for example [http://psdblog.worldbank.org/psdblog/aid\\_effectiveness/index.html](http://psdblog.worldbank.org/psdblog/aid_effectiveness/index.html)

<sup>50</sup> Knack and Rahman, 2004. Donor Fragmentation and Bureaucratic Quality in Aid Recipients. World Bank Policy Research Working Paper 3186, Washington DC. [www.econ.nyu.edu/cvstarr/conferences/ForeignAid/papers/Knack.pdf](http://www.econ.nyu.edu/cvstarr/conferences/ForeignAid/papers/Knack.pdf)

150. The idea that results can be compared is also a function of the desire of most development agencies to avoid risk; there is a constant search for ways to reduce risk, and benchmarking of performance will probably encourage that tendency; a programme with a high chance of a modest impact will generally be preferred, over a programme with a modest chance of a very high impact.

151. Indeed, PSD is rather more ambitious than some other development interventions, in that it aims to enhance an entire system ('the private sector'); it is therefore more complex, dynamic and multi-dimensional than, for example, programmes to increase distribution of vaccines. This does not, however, absolve practitioners from demonstrating results.

## Early Adopters of Core Indicators

152. TechnoServe is using 7 core indicators to measure progress against programme forecasts and goals internally; these indicators include measures of business success, namely sales and net profits, as well as indicators that proxy socio-economic impact on the poor, namely value of purchases from rural producers, number of producers, wages paid and number of employees.

153. In addition, it carries out focused assessments of projects from time to time, to triangulate on other impact measures and to document important anecdotal evidence of the broader socio-economic impact of its work.

154. Interestingly, it also measures Return on TechnoServe Investment (ROTI), which is the incremental gross revenues generated by its SME clients from project commencement to three years after client engagement, divided by the full cost of TechnoServe assistance. This gives a measure of cost-effectiveness, to measure progress against long-term objectives. In other words, it provides a broad measure of whether the performance of the organisation is improving, over time.

155. A major reason for choosing this indicator was that it was likely to be auditable over time, in a way that left a paper trail, and was replicable without relying unduly on assumptions or local conditions. In this context, it was felt that employment was too difficult to measure in a way that was consistent and reliable; issues such as seasonality and possible child labour also posed methodological challenges.

156. Also, it was felt that increased revenues were a more reliable proxy indicator of the total benefits to the poor, since employment is only one channel through which they would lead to benefits; purchases from other businesses, taxes paid and profits re-invested were also likely to be making important contributions. While the measurements will never be exact, they are likely to be approximately right.

157. Attribution issues remain, and documenting revenues is also challenging in many situations even if TechnoServe often works with lead firms in value chains, meaning that they do have records of some kind. Nonetheless, market development effects (such as spontaneous replication of innovations) mean that tracking increases in revenues is a major task in itself.

158. Indeed, and more generally, the potential for comparison of performance highlights the methodological pitfalls; if development professionals are going to accept any kind of ranking, they are clearly going to need to believe in the methodology being used. Nonetheless, Technoserve already find the ROTI indicator useful, apparently, as part of the client/project selection process (which must also include justification of how the intervention will benefit the rural poor).

159. Importantly, the ROTI allows staff of TechnoServe to orient and manage their efforts around the organisation's mission, which is to help "entrepreneurial men and women in poor rural areas of the developing world to build businesses that create income, opportunity and economic growth for their families, their communities and their countries". However, it is not yet used in any direct form as an externally-published measure of achievement.

160. Helvetas is also moving towards a uniform set of key indicators for all their value chain projects, starting with a measure of the change in farm household income. This is currently focusing on the increase in business incomes as a result of interventions (as with Technoserve) and does not yet make any estimate of displacement or deadweight. The number of households reached is also being measured, disaggregated by sex. Ultimately, these measures will be related to total project costs per farm household involved in the chain.

161. Some practical examples of recent impact assessments that are in the public domain are explored in the following Chapters. There are others that are not yet fully in the public domain, because of concerns about rigour; indeed, all of the measurements reported below imply a number of methodological challenges.

162. Again, though, it is better to be approximately right, than completely mysterious about outcomes; in the absence of any reporting of results, critics and sceptics will always say: "maybe you cannot measure it because it isn't there..." In that light, therefore, any reporting of results will be an improvement on the current situation, notwithstanding the methodological challenges.

## J. Reporting on Scale



Photo: Gavin Anderson

*Approximate measures are therefore better than no measures at all; what can be measured approximately by most practitioners? Where to make a start?*

163. Development agencies are sensitive to the suggestion that their interventions may be affecting only a lucky few; meanwhile, millions remain in poverty. Some programmes are therefore considering above all how many people they are reaching; for example, a donor-funded initiative to support Uganda radio stations in launching programmes for small business estimated that there were 7 million listeners; the programmes were reaching at least two thirds of the poorest members of society<sup>51</sup>.

164. Similarly, the Rainforest Alliance<sup>52</sup> (RA), a US-based NGO, issues a seal of approval, which “makes it easy for consumers to know they are

buying a product that has been grown or made sustainably. The companies who sell these products also make sure their customers know they are good neighbours in their communities, and that they take care of their workers and protect the environment.” Specifically, RA reports the following, rather impressive outcomes<sup>53</sup>:

- **Forest Products:** RA has certified 3% of the working forests (100m acres) in 50+ countries, working with IKEA, Gibson USA, B&Q, Domtar, Klabin, Tembec, Grupo Nueva

- **Agriculture:** RA has certified 1.3% of the world’s coffee, and 15% of the world’s bananas; in total, it has certified nearly 10,000 farms, growing coffee, citrus, bananas, pineapple, cacao, flowers, macadamia nuts, ferns, passion fruit, plantains, in collaboration with Kraft, Chiquita, Caribou Coffee, Whole Foods Markets

165. Following up on this, the targets that RA is now setting are ambitious; for example, the organisation is working with UNDP / GEF to certify 10% of the world’s coffee, to protect biodiversity on coffee farms; similarly, it is collaborating with Kraft to certify 100% of their sales under existing trademarks in Western Europe and the US (12,000 tonnes in 2006, benefiting 27,000 families).

166. Similarly, the Shell Foundation says that its “work is showing results:

<sup>51</sup> MSE Radio Programme Listener Survey, by FIT-SEMA, September 2004  
[www.value-chains.org/dyn/valuechains/docs/381/RadioProgListeners9-04.pdf](http://www.value-chains.org/dyn/valuechains/docs/381/RadioProgListeners9-04.pdf)

<sup>52</sup> [www.rainforest-alliance.org](http://www.rainforest-alliance.org)

<sup>53</sup> [www.enterprise-development.org/resources/download.asp?id=412](http://www.enterprise-development.org/resources/download.asp?id=412)



- 263,000 people in Mexico City travelling more sustainably every day
- \$100m committed to small and medium sized enterprises in Africa
- 2,000 jobs created and 14,000 incomes improved in India, Eastern Europe and Africa
- 1 million women and children around the world benefiting from cooking stoves that drastically reduce health-damaging indoor smoke pollution.”

167. It also gives as its “vision for the next five years:

- 1,000,000 people travelling more sustainably every day
- \$300m committed to small and medium sized enterprises in Africa
- 5,000 jobs created and 30,000 incomes improved in India, Eastern Europe and Africa
- 100 million people around the world benefiting from cooking stoves that drastically reduce health-damaging indoor smoke”<sup>54</sup>

168. These indicators reflect a mixture of increased inputs (“funds committed”) un-quantified impacts (“more sustainably”, “incomes improved”), and quantified impacts (“jobs created”). They paint a picture of vigorous action – and importantly, they point to a starting point for measurement, that most practitioners can see as feasible. Even if they do not measure much, almost anyone involved in a programme can estimate how many people are likely to be affected directly by it; intuitively, a programme that reaches millions of people, even superficially, is likely to be more interesting than one that reaches just a few hundred, in depth.

169. Nonetheless, the extent of the scale reported above, while a critical first step for impact measurement, does not necessarily say much about the changes that it made in people’s lives, nor does it form the basis for a common measurement system. The following Chapters outline some assessments of actual impacts, in the development of value chains and service markets.



Photo: Gavin Anderson

<sup>54</sup> Shell Foundation, 2007. Down to Business: New Solutions to Old Problems. [www.shellfoundation.org](http://www.shellfoundation.org)

## K. Impact Assessments for Technical Assistance in Value Chain Development

Photo: Charles Bodwell



*Several programmes have made serious efforts to measure their achievements – and to relate them to the cost of achieving them. Most of these have published their findings only recently, and some are not yet even published. The following Chapters summarise some of the most interesting examples – particularly for those aiming to up-grade whole market systems.*

### SEDF in Bangladesh

170. The IFC's South Asia Enterprise Development Facility (SEDF) has been working since 2003 in the sectors of ready-made garments, light engineering

and agri-business – under its Sector Development and Business Services (SDBS) programme. A major element in the strategy was to engage directly with market actors at the top of the value chain, where possible; instruments included technical assistance and cost-sharing for access to advisory services.

171. SEDF also aimed to address any constraint that was more broadly limiting enterprise growth in the chosen sectors, including for example constraints around access to finance, and the business environment. In addition, it considered how competitive Bangladeshi industries could be in global markets, when designing interventions<sup>55</sup>; the Table below illustrates this, for the Ready-Made Garments sector.

172. SDBS introduced a results-based performance monitoring system, based on a logical framework that linked activities to specific sector-level changes; a customised management information system was introduced as part of this. Action plans were linked to the log-frame, and staff incentives were built in. This system enabled staff to publish a review of the results achieved<sup>56</sup>, which also provided much more detail about the individual activities.

Measure	SEDF assisted	Bangladesh baseline	Sri Lanka	China	Vietnam	Cambodia	Global
Labour turnover pa, %	11	15	6-8	6	7	7	5-6
Rejection rate, %	8	10	3	0.5	3	5	1
Plant availability, %	44	30	65	80	70	60	75
Labour cost, \$/hour	0.37	0.30	0.46	0.50	0.30	0.40	
ROI, % (knit)	18	15	18	25	20	18	
Women as supervisors, %	1	0.5	40	70	60	65	40

<sup>55</sup> [http://www.value-chains.org/dyn/valuechains/docs/497/Adhikary\\_Monitoring.pdf](http://www.value-chains.org/dyn/valuechains/docs/497/Adhikary_Monitoring.pdf)

<sup>56</sup> Adhikary and McVay, 2006. Market Development in Practice: Sector Development and Business Strategy. IFC SEDF, Dhaka. [www.value-chains.org/dyn/valuechains/bdssearch.details?p\\_phase\\_id=560&p\\_lang=en&p\\_phase\\_type\\_id=5](http://www.value-chains.org/dyn/valuechains/bdssearch.details?p_phase_id=560&p_lang=en&p_phase_type_id=5)

173. The review reported that “assisted clients have generated 16,239 jobs, of which the majority (65%) were for women. This includes 318 new supervisory jobs for women. Based on a total programme cost of \$6.08m to date, these figures represent a programme cost of \$374 per job generated by an assisted client. In total, assisted firms employ 126,003 people, of whom 62% are women.”

174. While these numbers represent, to some extent, actual impacts, some of the donors supporting this work chose in the Foreword to the review to highlight instead that the programme “has achieved a leverage ratio of around 1:19 times our combined investment, a respectable figure given the challenges of working in Bangladesh”. In this sense, therefore, they chose a proxy that could be measured with some accuracy, over a direct impact where the accuracy could be called more into question.

175. Either way, SEDF reports that it was able to achieve good results, despite concerns about the overall competitiveness of local industry, by addressing a wide range of constraints to growth; furthermore, it addressed them in a diverse and flexible way. The results therefore serve to validate the overall strategy, more than any one intervention or input.



Photo: Charles Bodwell

## The Prosperity Initiative and Bamboo in the Mekong

176. In 2006, a detailed feasibility study was carried out by Oxfam Hong Kong, working with the IFC's Mekong Private Development Facility (MPDF); the study benchmarked the potential of the bamboo sector in Vietnam, Laos and Cambodia against the successful experience of China<sup>57</sup>. This study was able to evaluate the pro-poor impact potential at a regional scale, based on projections of future global demand, and the competitiveness of the local bamboo industry; it cost about \$300,000.

177. The study found that pro-poor impacts could be created at a regional scale if the three bamboo sub-sectors could become efficient, and produce a range of products that were competitive in a growing global market. Each of the three sub-sectors has a particular structure, with different potentials to target poverty. Industrial bamboo (flooring, furniture, charcoal, blinds, mats, panels, chopsticks etc) has a high potential impact on farmers, due to growing demand for raw material. Handicrafts impact the labour market (women in particular), and bamboo shoots again benefit farmers – although in smaller numbers.

178. By looking across the three sub-sectors and targeting a range of collaborative interventions in the region, the project team estimated that a \$40m investment by donors into a coordinated strategy to develop bamboo across Vietnam, Laos and Cambodia over the coming decade could generate 800,000 jobs at a cost of \$50 per job. Through this, it could lift 1million people out of poverty; importantly, this approach is based on measurable poverty impact goals, bottom-line indicators against which the overall performance of the programme could in principle be evaluated in future.

179. At the time of writing, PI (the new organisation which has formed to take this work forward along with IFC) was gearing up to implement the programme on a multi-agency basis, as the Mekong

<sup>57</sup> [http://www.value-chains.org/dyn/valuechains/bdssearch.details?p\\_phase\\_id=597&p\\_lang=en&p\\_phase\\_type\\_id=2](http://www.value-chains.org/dyn/valuechains/bdssearch.details?p_phase_id=597&p_lang=en&p_phase_type_id=2)



Bamboo Consortium (MBC). Driven by the knowledge and linkages generated in the sector feasibility work, IFC and PI have been working with businesses and partners to develop a demonstration supply chain in Thanh Hoa province, Vietnam.

180. A range of activities have been developed, for example in support of small factories that dramatically reduce wastage of bamboo, by pre-processing the culms close to the farms; these factories have been established mostly by local, small-scale investors. By late 2006, they were employing more than 500 staff, 88% of whom came from ethnic minorities, and 85% of whom were women. Employees were earning \$37.50 per month on average, which compared well with the average household income locally of about \$17 pm.

181. The impacts on farmers as a result of the increased efficiency have, however, been even more significant. According to local authorities – and the farmers and traders themselves – the emergence of the pre-processing factories was pivotal in creating the conditions for rapid expansion of demand for bamboo. Specifically, between 2004 and 2006, demand increased by 23%, and the price increased by 32%. Farmers are now planting bamboo, and moving out of less sustainable practices in the mountainous areas. About 8,500 people of the 125,000 people who actually were directly benefiting from income from the bamboo sector (4.8% of the total population) were found to have moved out of

poverty in the previous 2 years. On the other hand, of households without any bamboo income, over 3,000 people had fallen back into poverty.

182. In summary, bamboo was shown to be the only economic sector positively correlated with movement out of poverty; the 30% of the population who were persisting with other sectors and not making the change to bamboo were slipping backwards into poverty. Plans to replicate this type of experience are now being rolled out by MBC and its partners in the region, using the clear targets set using the feasibility analysis. There is great potential for growth; costs of raw bamboo and labour are still much lower in Vietnam than in China, although the Chinese experience suggests that both will rise substantially if the programme is successful – with corresponding impacts on poverty.

## Cluster Initiatives in Sri Lanka

183. USAID in Sri Lanka commissioned an assessment of its various cluster initiatives there, and the summary report was published in 2003<sup>58</sup>; it subsequently achieved honourable mention in USAID's Impact Assessment Contest on building credible impact information on PSD programmes<sup>59</sup>. The programme included a portfolio of activities with a wide range of sectors, such as rubber, precious stones, ceramics, tourism, coir, spices and tea.

<sup>58</sup> Bolnick, 2003. The Economic Impact of Cluster Initiatives: The Competitiveness Initiative Project in Sri Lanka. Nathan Associates Inc. for USAID Colombo. [www.nathaninc.com/?bid=422&PubID=949](http://www.nathaninc.com/?bid=422&PubID=949)

<sup>59</sup> [http://www.microlinks.org/ev\\_en.php?ID=12529\\_201&ID2=DO\\_TOPIC](http://www.microlinks.org/ev_en.php?ID=12529_201&ID2=DO_TOPIC)

184. The assessment looked particularly at the cost-benefit, calculating that the 8 activities selected had a present value of net additional income of \$156m. Factoring in uncertainty and concerns about deadweight, this was reduced to \$69m. This was apparently achieved at a cost of about \$7m, giving a benefit to cost ratio of about 10:1 (not taking into account costs contributed by entrepreneurs themselves – which would reduce the ratio to 8.6).

185. The author stresses that this is a conservative estimate. Also, there was a large difference between the performances of the different activities examined; two generated income of more than \$25m each, while at the other end of the spectrum, one activity had a negative yield. Although job creation was mentioned explicitly in connection with the rubber sector (“the programme will create jobs for 77,000 people”), it was not measured in this assessment.

186. The author concludes: “even though the results reported here are reasonably good, there is an astonishing paucity of data on the economic impact of cluster initiatives. As a result, the cluster approach has yet to meet the burden of proof as an effective use of development assistance. To remedy the lack of information and resolve the arguments about the role of competitiveness projects, it is essential to ensure that monitoring and evaluation of the economic impact of cluster initiatives is part and parcel of every competitiveness project”. This might sound familiar to the reader.

## Katalyst in Bangladesh

187. Katalyst<sup>60</sup> is a multi-donor programme working to develop value chains and service markets in Bangladesh, through an extensive portfolio of activities in many different sectors. Between 2002 and 2007, the programme staff calculate that they have generated at least 183,000 additional jobs (full-time equivalent), at a cost of about \$100 per job. The great majority of these jobs will be for people living in poverty; half of the Bangladeshi population currently lives below the international poverty line.

188. Attribution has been estimated for each set of activities individually: once impact has been measured, it is reduced in line with staff estimates of how much of that impact can reasonably be ‘claimed’ by Katalyst, rather than being a result – for example – of investment by the private sector. This usually involves reducing the measured impact by 60-90%, to calculate the impacts reported above. The estimates are therefore rather conservative, to take into account the many other influences on the target group.

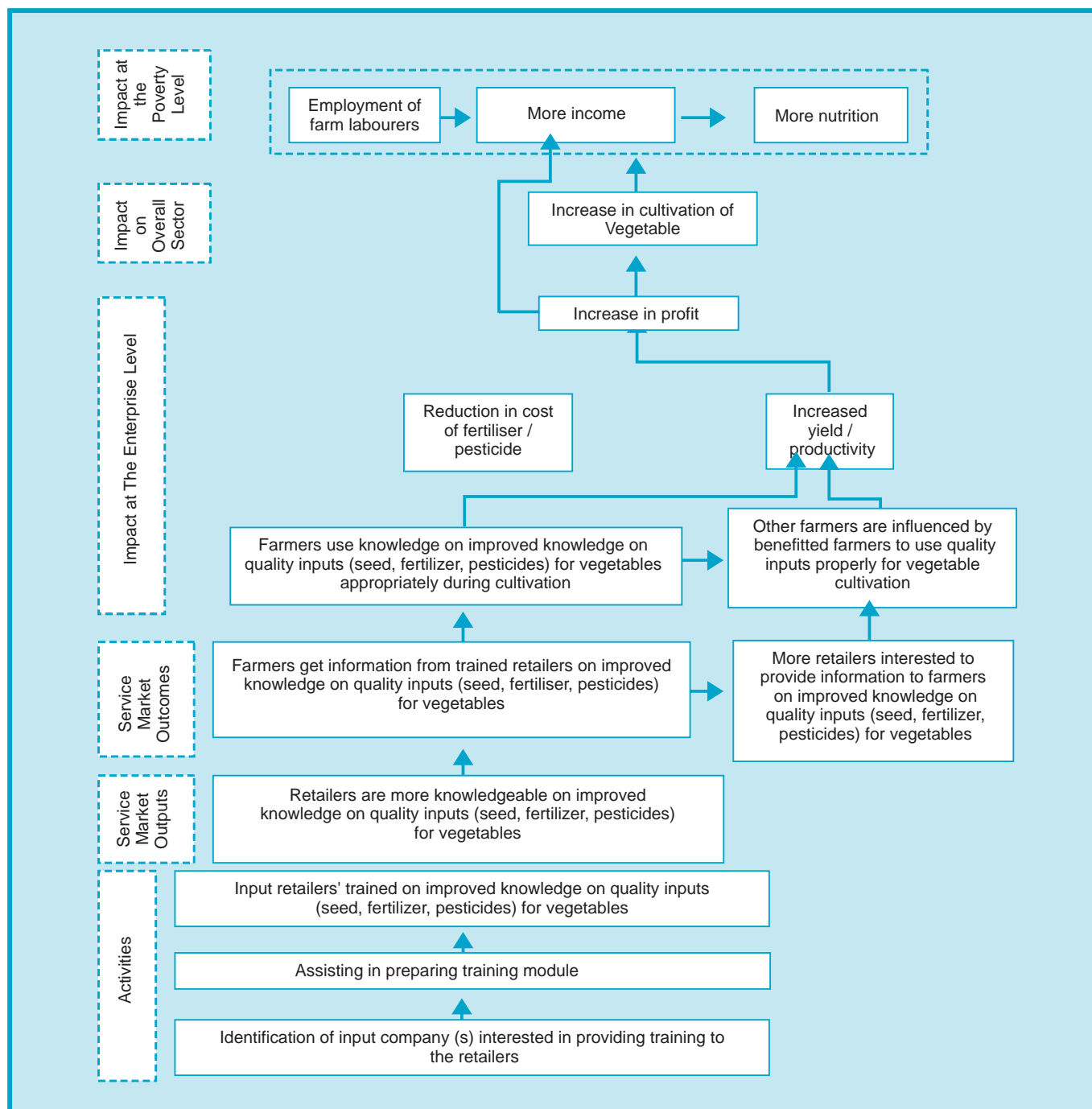
189. The calculation of impacts does include the estimated effects of ‘crowding in’ of new entrants to the sector (through a demonstration effect), and also offsets estimated displacement effects. It does not include possible indirect impacts, through the forward and backward linkages stimulated by the direct impacts. Increased maize output, for example, is likely to stimulate more poultry production, leading to increased employment also in that sector. Although likely to be very substantial in some sectors, these indirect impacts are not included, as they proved too difficult to estimate with any accuracy. In summary, the numbers are likely, therefore, to be rather conservative.

190. As part of its portfolio of activities, Katalyst has partnered with leading international suppliers of agricultural inputs, to provide training to over 1,300 retailers supplying vegetable farmers with inputs. The training enabled the retailers to give smallholder farmers better advice on how to use inputs, such as chemicals and seeds, more effectively. This would be provided as an ‘embedded’ service – as part of the sales transaction, rather than being paid for separately. The causal model is illustrated below.

191. After one growing season, Katalyst surveyed 500 vegetable farmers. About half of the farmers surveyed had been purchasing their inputs from retailers who had taken the training; the other half were purchasing their inputs elsewhere. The two samples were selected at random, to the extent possible.

192. The research found that farmers with larger plots (over 1.5 acres) who were buying from trained retailers were 50% more productive than those

<sup>60</sup> www.katalystbd.com



buying from other retailers, partly because they were spending less on inputs. Earnings per acre were also 58% higher. Farmers with very small landholdings (less than 0.5 acres) who were buying from trained retailers also performed better than those who were not; while the differences were statistically significant, they were however smaller than for the larger farms – the profitability, for example, was only 6% better.

193. Nonetheless, the programme had achieved impressive scale; one year after training retailers,

about 239,000 vegetable farmers had benefited directly from the improved advice. Katalyst also looked at the market-wide effects, finding that vegetable farmers who were not buying from trained retailers were nonetheless copying the improved practices of those who were buying from trained retailers. Also, retailers who had not been trained were beginning to copy those who had been.

194. As a result of these effects, about 246,000 vegetable farmers were likely to benefit indirectly from the retailer training, within three years. This



Photos: Katalyst

suggested a total outreach of about 485,000 vegetable farmers, at a cost to Katalyst of \$100 per retailer – suggesting a return on investment (extra income to farmers / investment by Katalyst and companies) of about 1:9 after three years (including the indirect effects).

195. Further research is now needed, to understand these findings better, as they raise a number of questions. For example, were the farmers buying from trained retailers already more discerning? Questions of possible bias, related to self-selection of the treatment groups, are very often present in programmes that are developing whole markets. And how will the retailers respond to the implied loss of sales, as the larger farmers stop buying inputs that they did not really need? Finally, might further research reveal ways in which retailers could target the smallest farms more effectively?

## Water Pumps and Other 'Appropriate Technologies'

196. 'Appropriate technology' (AT) interventions are an example of a type of intervention that was popular in the 1980s. It became clear, however, that many of these projects were not adequately related to market structures and local priorities; the devices being worked on were either unaffordable or were failing to address a critical need (or both). Over time,

however, a few AT interventions achieved scale of outreach that was truly impressive<sup>61</sup>.

197. For example, KickStart (previously ApproTEC) is feted as a cutting-edge example of social entrepreneurship by the World Economic Forum, Time Magazine, ABC, Deutsche Welle and the Wall Street Journal. Furthermore, KickStart reports that its work (particularly with treadle water pumps for improving irrigation for smallholders) has led to the creation of 50,000 new businesses, which generate \$52m per year in profits and wages – contributing new revenues equivalent to 0.6% of Kenya's GDP<sup>62</sup>. This is the sort of very large scale impact that development agencies dream of.

198. Attribution in such cases is relatively straightforward, in the sense that the benefits can be closely linked to the intervention; similarly, deadweight and displacement are less problematic than, for example, interventions like management training (discussed below).

199. The success and recent recognition of KickStart and others raises the question about how donor support for different approaches grows and shrinks; it may not be based only on the evidence available at the time, or even later. In particular, do donors reward proven performance, when it is finally demonstrated – or are the negative experiences so traumatic and institutionalised that they are reluctant to look back?

<sup>61</sup> Indeed, substantial impacts have been documented, but did not alter the overall trend; see for example Poverty Alleviation as a Business, by Urs Heierli, May 2000, <http://www.intercooperation.ch/sed/product/heierli/main.html>

<sup>62</sup> <http://kickstart.org>

## L. Impact Assessments for Management Training

*Some reported impacts of management and entrepreneurship training are included here, as they are likely to be of particular interest to Seminar participants, and to students of Readers in previous years.*

200. Measuring the impacts of management and entrepreneurship training is notoriously difficult to do, partly because of the many other events and influences in the life of an entrepreneur. A small piece of advice picked up in a training session may ultimately make a huge difference to the business, but the entrepreneur may forget completely that she had picked it up in the training course. Nonetheless, many attempts have been made to measure the impacts, and some of the more interesting ones are outlined below.

### Management Training in India

201. Malcolm Harper describes the evaluation of the original achievement motivation training (AMT) experiments in India<sup>63</sup>, in his book with Gerry Finnegan entitled *Value for Money?*<sup>64</sup> "This evaluation calculated that the local training cost per new job created in trainees' businesses was Rs. 183 (or about \$25 in 1968 when the experiments took place), and that the trainees invested an average of \$100 of additional capital in their businesses for every \$5 of training cost."

202. "A further study of a 'Block Adoption Programme' in an Indian sub-district, carried out by the Industrial Development Bank of India (IDBI), showed that the average training cost per job created was about Rs. 4,600 (or \$150), and that it resulted in average annual earnings per trainee of Rs. 6,300 (or \$210)(Acharya, 1990)." While these numbers are interesting, it seems that no attempt was made to look at market-wide impacts, displacement etc.

### Training of Peruvian Micro-finance Clients

203. Researchers from Yale University<sup>65</sup> carried out a randomised control trial with women clients of FINCA, a micro-finance institution (MFI) in Peru that was also offering training to its clients in management skills. "Treatment groups received thirty to sixty minute entrepreneurship training sessions during their normal weekly or monthly banking meeting over a period of one to two years. Control groups remained as they were before, meeting at the same frequency but solely for making loan and savings payments."

204. Significant benefits were identified for the MFI; repayment and client retention were both found to have increased. Some benefits were also identified among the clients:

<sup>63</sup> McClelland and Winter, 1971. *Motivating Economic Achievement*. Free Press, New York.

<sup>64</sup> Harper and Finnegan, 1998. *Value for Money? Impact of Small Enterprise Development*. IT Publications, London.

<sup>65</sup> Karlan and Valdivia, 2006. *Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions*. Yale University. [aida.econ.yale.edu/karlan/papers/TeachingEntrepreneurship.pdf](http://aida.econ.yale.edu/karlan/papers/TeachingEntrepreneurship.pdf)



205. "Training participants demonstrated greater business knowledge [and] the greater knowledge translated into better business practices, though only in limited areas. The training increased the likelihood that individuals reinvested profits in their business ..., maintained sales records for their business ..., and maintained withdrawal records from their business". "The training has helped clients identify strategies to reduce the fluctuations in their sales."

206. "Larger effects were found for those that expressed less interest in training in a baseline survey. This has important implications for implementing similar market-based interventions with a goal of recovering costs." While the findings were interesting in terms of their general trends, the research was not sufficiently detailed or extensive to be able to detect increases in employment over time.

## M. Impact Assessments for Challenge Funds, Matching Grants and Public Private Partnerships (PPPs)

*Challenge funds, matching grants and PPPs are discussed in some detail in this Chapter, since they are outside the direct experience of many practitioners working in the field of value chain development. Nonetheless, they are popular with bilateral donors, and some are quite large. They also tend to focus on specific value chains.*

207. Many donors operate schemes that essentially make funds available to individual companies, in order to stimulate investment and innovation; the Canadian Council for Public Private Partnership ([www.pppcouncil.ca](http://www.pppcouncil.ca)) offers the following definition of PPPs: “a cooperative venture between the public and private sectors, built on the expertise of each partner, that best meets clearly defined public needs through the appropriate allocation of resources, risks and rewards.”

208. A Table summarising some of the larger schemes is included on the following page<sup>66</sup>. Such schemes are variously called challenge funds (DfID, AusAID), matching grants (World Bank and others) or Public Private Partnerships (GTZ); SDC prefers the term Public Private Development Partnerships (PPDPs) – to emphasise the developmental nature and intentions of such partnerships. The unique feature of such schemes is that the development agency provides direct assistance to an individual business, on the understanding that the interests of the agency and the business coincide – or at least overlap.

209. Often, such schemes encourage ‘lead’ companies to expand a value chain, or to diversify into a new one, by contributing (often 50%) to the cost of a pre-agreed activity. They therefore complement the technical assistance provided by many donor-funded initiatives, and the soft finance provided by specialist agencies like FMO and DEG, although the three approaches rarely work closely together, in practice.

210. While some of these programmes mention potential market-wide impacts of the individual partnerships, there have not been major efforts to try to stimulate such impacts. Also, none have yet made efforts to measure spontaneous replications. Anecdotally, the evidence is not clear; there are occasionally replications that happen, but these can seem almost accidental.

211. In such schemes, the ratio of operating or management costs to the value of the grants disbursed is widely considered as an important indicator of efficiency. A review of 10 matching grant schemes for the World Bank, for example, reported that this ranged between 18% and 57%<sup>67</sup>.

212. Such reporting encourages the idea that management costs are an ‘overhead’ to be minimised; this also minimises the extent to which programme managers can understand markets, build relationships and provide forms of assistance other than grants. Arguably, however, this is one

<sup>66</sup> For a more detailed description of such schemes, see “Engaging Business in Development: Results of an International Benchmarking Study” by GPPi, 2007 [www.gppi.net/fileadmin/gppi/Engaging\\_Business\\_Final\\_06222007.pdf](http://www.gppi.net/fileadmin/gppi/Engaging_Business_Final_06222007.pdf)

<sup>67</sup> Phillips, 2000. Implementing the Market Approach to Enterprise Support: A Comparative Evaluation of Matching Grant Schemes. ECSPF World Bank. <http://www.value-chains.org/dyn/valuechains/docs/311/BDStrainingDavidPhillips.pdf>

## Matching grants and Public Private Partnerships

Donor	CIDA	DANIDA	DANIDA	DFID	DGIS	DGIS	USAID
<b>Program name</b>	Industrial Coop. Prog. (Inc.)	Business to Business (B2B)	Public Private partnership (PPP)	Challenge fund (CF)	WSSD-Call for ideas	PSOM Cooperation with Emerging Mkts	Global Dev. Alliance
<b>Total budget</b>	\$49.3 mill (2005)	\$37mill (2006)	\$5.5 mill (2006)	\$148.9 mill	Need-based grant budget	\$70.3 mill.	Varies
<b>Number of projects</b>	150 (2005)	50 (2006)	20 (2006)	8 CFs: BLCF, AECF, PPP etc.	50 partnerships in 15 countries	45 projects completed, 98-05	Approximately 400, 99-05
<b>Ave. grant / range</b>	\$90,000 – 540,000	\$540,000 (start-ups \$88,000)	\$276,000	\$100,000-2m	\$275,000-1.4m	\$1m-2m	\$200,000-1m
<b>Indicators used</b>	Jobs created Additional tax income, wealth	Jobs created Forex earned Technology relevance Environmental improvements Sustainability Additionality	10 indicators reported by the partner companies Internal reporting Reporting against project-specific milestones	No. organisations linked No. corporations linked No. products, services created No. linkages created No. direct jobs created No. indirect jobs created	Depends on project	Depends on project	Depends on project
<b>Max. Contrib.</b>	75%	90%	60%	50%	Under review	50-100%	80%
<b>Project ideas come from</b>	Canadian companies, not publicly advertised.	DANIDA SME consultants and normally DANIDA country staff.	Program experiences, networks; road shows in Denmark	Advertisements	Dutch private companies, associations and Embassies	Dutch companies (mainly SMEs), road shows and publicity in the Netherlands	NGOs are typical source of ideas.
<b>Selection Process</b>	Rolling applications. 2 step selection: concept paper, full proposal. No tendering	Competitive tender, or identified by programme field coordinators in Embassies	Competitive tender, or identified by programme field coordinators in Embassies. Final decision made by Ambassador.	Calls for competitive proposals twice p.a. Concept note, then full proposal assessed by independent panel and technical experts	Occasional tenders, assessed against MDGs. Proposal assessed by experts, then joint planning. Call for Ideas: 3.5% approval rate	Tender process twice a year, with free training on proposal writing. Short-listed projects are visited, board approves; Embassy has veto right	NGOs submit concept paper, then full proposal, which are assessed by GDA staff and in-country officials.
<b>Duration</b>	1 – 2 years	Contact: up to 1 year Pilot: up to 2 years Project usually 3 years	Preparation: up to 1 year Project: up to 3 years	Average 3 years	No requirements set yet.	2 years, but exceptions are possible (e.g. Agriculture)	No limits
<b>Implement.</b>	1 HQ project manager per project and field officer in the embassy responsible for implementation	1 HQ fund coordinator with budget; implementation managed by Embassy staff	HQ manager supports Embassy staff to plan, implement with partners, who must report impacts annually for 3 years after completion	Consultants contracted to manage and evaluate, supervised by DFID and panel / board	Country projects implemented by Embassy. Regional partnerships managed by PPP Unit in The Hague	Dutch EVD manages partnership with Dutch recipient companies, according to agreed milestones	NGOs implement performance-based contract, under supervision of Mission Directors.
<b>Web</b>	<a href="http://www.acdi-cida.gc.ca/psd">http://www.acdi-cida.gc.ca/psd</a>	<a href="http://www.um.dk/en">www.um.dk/en</a>	<a href="http://www.um.dk/en">http://www.um.dk/en</a>	<a href="http://www.dfid.gov.uk">http://www.dfid.gov.uk</a> <a href="http://www.challengefunds.org">www.challengefunds.org</a>	<a href="http://www.minbuza.nl/en/home">http://www.minbuza.nl/en/home</a>	<a href="http://www.minbuza.nl/en/home">http://www.minbuza.nl/en/home</a>	<a href="http://www.usaid.gov/our_work/global_partnerships/gda/">http://www.usaid.gov/our_work/global_partnerships/gda/</a>

attraction of such schemes: that they can 'retail' relatively large sums with relatively low overhead. The idea that active management (and indeed partnership) can deliver equivalent results – and even be as cost-effective – needs to be demonstrated, and preferably also quantified, even if only approximately.

213. Another commonly-used measure with these types of programmes is the leverage ratio – the ratio of funds invested by the company to the funds invested by the development agency. If this is high, then the agency has been efficient in re-orienting large amounts of private sector funding to achieve its goals. In other words, the scale of what it has been able to achieve with a given amount of public funds has been substantially increased.

214. The issue of additionality (whether the grant made something happen that would not otherwise have happened) is discussed in some detail in "A Micro-econometric Evaluation of the Mauritius Technology Diffusion Scheme (TDS)", by Tyler Biggs for the World Bank in 1999<sup>68</sup>. The conclusion is that any such scheme needs to be able to show, quantitatively and qualitatively, that the public funds triggered events that would not otherwise have happened; in Mauritius, this was not the case.

215. Given that the events did happen, it is also important to relate the developmental impacts to the public funds contributed, and the rest of this Chapter

is dedicated to that question. If it could be resolved, then the leverage issue would be factored in to the final result, since the aim of such measurements would be to give an estimate of the overall effectiveness of the application of public funds.



Photo: Charles Bodwell

## Programme for Cooperation with Emerging Markets (PSOM), Netherlands

216. The Dutch Programme for Cooperation with Emerging Markets (PSOM), for example, makes funds available "to cost-share the initial financial risks that [Dutch or LDC] companies face when investing in the emerging markets of developing countries. PSOM aims to finance pilot investment projects that lead to follow-up commercial investments and / or a lasting trade relation between the Dutch and local companies."<sup>69</sup>

217. It was decided to expand the scheme in 2001, to a budget of €30m (\$41m) p.a., on the understanding that it was meeting explicitly developmental goals (in addition to the commercial ones of the participating companies): job creation, involvement of local SMEs, transfer of knowledge and concentration on poorer regions.

218. Indeed, a sophisticated tracking system was introduced, to try to measure these variables for every grant, and to move towards performance-based contracts – in addition to tracking the developmental impacts for five years after the grant had finished<sup>70</sup>. In 2004, the budget was further increased to €51m (\$70m) p.a. The average grant was about €825,000 (\$1.1m); project titles give some indication of the typical content, for example:

- Quality Management and Upgrading of the Thai Horticultural Chain
- Eco-Tech Pot-plant Propagation, Uganda
- Production and Processing of Fresh Frozen Vegetables, Indonesia
- Bakery School, Mozambique
- Organic production of Garlic, Sweet Corn and Celery, South Africa

219. An independent evaluation of the results in 2005 found the following cost per direct job created:

<sup>68</sup> [www.value-chains.org/dyn/valuechains/docs/165/BDStrainingTylerBiggs.pdf](http://www.value-chains.org/dyn/valuechains/docs/165/BDStrainingTylerBiggs.pdf)

<sup>69</sup> PSOM Evaluation 2005: Final Synthesis Report. ECORYS-NEI for the Netherlands Ministry of Foreign Affairs (version March 2006).

<sup>70</sup> The extent to which the scheme was tied to Dutch companies was reduced at the same time.

Sector	No. of projects	Cost per direct job created, US\$ <sup>71</sup>	Cost per total jobs created, US\$
Agriculture / agri-business	16	10,700	2,740 – 4,110
Industry / manufacturing	2	12,500	
Energy / environment	2	43,000	
Transport / infrastructure	1	192,200	
Tourism	1	62,100	
Overall	22	12,600	

220. The number of jobs created indirectly in agriculture was calculated by extrapolating from the results of two projects, in Indonesia and Thailand, that had contracted approximately 800 farmers each. However, more information about the classification of projects is not easily available, as the design of each depends on the individual circumstances and proposal.

221. The consultants concluded that the impacts were increasing with time after completion of the project, but did not have the data to quantify the increases. Indeed, these numbers are only very approximate, and do not take into account in a systematic way issues of attribution / additionality, displacement or deadweight. Also, the sample size for most of the sectors was not large enough to be able to deduce relative lessons about sectoral focus. Nonetheless, the numbers were calculated by external consultants, based on visits to 22 completed projects in the field, so do give some interesting pointers.

## DfID Challenge Funds

222. DfID's Business Linkage Challenge Fund (BLCF) committed a total of £16.6m (\$33.4m<sup>72</sup>) to 58 projects, implying an average grant size of £270,000 (\$542,700). Again, the project titles give a clue as to

the sort of intervention envisaged, including for example:

- Fair Trade Horticulture in the Gambia
- Pro-Poor Tourism in South Africa
- Positioning Mozambique's Fruit Industry in the Global Market
- Organic Herb Production Systems in Saint Lucia
- Smallholder Essential Oil Production in Zambia

223. A subsequent desk analysis<sup>73</sup> of the impacts measured by the fund managers of the first six rounds (£10.7m, or \$21.5m committed) indicated the following:

Direct jobs created / retained:	16,362
Number of skilled jobs created / retained:	3,146
Number of jobs created for women:	1,604
Number of skilled jobs created / retained for women:	1,196
Estimated number of indirect jobs created:	90,316
Estimated total jobs created / retained:	106,678

224. These figures imply \$1,300 per direct job created or retained, and \$200 per total jobs created or retained. In 2006, AusAID took the analysis further, dividing up the grants (and their reported impacts) according to the business motivation, to

<sup>71</sup> at €1.00 = \$1.37

<sup>72</sup> £1 = \$2.01

<sup>73</sup> Kiggundu, Baker, undated, "The Business Linkages Challenge Fund as a Private Sector Development Instrument". LSE Masters thesis. [www.businesslinkageschallengefund.org/components/download.aspx?id=1d2e4f44-2332-4e55-b9cf-759bc5204382](http://www.businesslinkageschallengefund.org/components/download.aspx?id=1d2e4f44-2332-4e55-b9cf-759bc5204382)

derive the numbers in the Table below. Note that these figures are based on calculations by AusAID, which were based on calculations by a Masters candidate, which were based on telephone interviews and other research by the BLCF managers. Apparently, no more detailed impact research was commissioned by DfID; as with PSOM, additional caveats apply, in the sense that no account has been taken of additionality, displacement or deadweight.

225. Again, though, as relative numbers, they are arguably sufficiently different to justify some reflection; for example, it would seem that employment creation or retention is relatively costly when creating new markets for new products – particularly relative to obtaining new inputs. This might be exaggerated in the Table, in the sense that new products in new markets may generate more employment in the long term; these numbers were generated in the immediate follow-up to grant implementation.

Business motivation	No. of grants	Ave. grant, \$	Ave. direct jobs created / retained	Ave. est. indirect jobs created	Total ave. est. jobs created / retained	Cost per direct job created / retained	Cost per total jobs created / retained
Reduction in supply costs	6	506,000	147	82	229	3,442	2,210
Obtain previously unobtainable resources	11	514,000	1,155	7,740	8,895	445	58
Create new markets for existing products	5	762,000	52	647	699	14,654	1,090
Create new markets for new products	3	1,000,000	22	167	189	45,455	5,291

Exchange rate £1.00 = US\$ 2.01

## N. Impact Assessments in Reform of the Business Environment

*Reform of the business environment has enjoyed great donor attention in recent years; what do we know about the impacts?*

226. Many development practitioners believe that excessive regulation ('red tape') is stifling enterprise growth, and that cuts in this regulation would have a very widespread impact, since they would benefit many businesses. The Doing Business<sup>74</sup> surveys of the World Bank, for example, have attracted widespread attention, by ranking countries according to the amount of red tape, in various areas of business operation. This has led many agencies to focus on this area, and improvements are being reported:

- the time required to get a permit in one city in the Philippines has been reduced from 17 days to 2 days<sup>75</sup>
- Kenya was found to have over 1,300 business licenses and fees imposed by 178 State bodies; 35 were eliminated in 2005. In the same year, 4,900 regulations were repealed in Ukraine<sup>76</sup>
- a new Business Registration Law in Serbia reduced registration time from 105 days to 5 days, and introduced a 'silence is consent' rule<sup>77</sup>

- the number of days to obtain an operating license in Lima was reduced from 60 to 1.6<sup>78</sup>
- a reform reduced registration time in Egypt from 366 days to 15 days<sup>79</sup>
- in Peru, the title registration process has been reduced from 7 years to 45 days, with the cost of title coming down from \$2,156 to \$49. This has led to 1.3m titles being registered between 1996 and 2002<sup>80</sup>

227. Intuitively, these achievements seem likely to help poor people to join the formal sector, and therefore to have pro-poor impacts. However, these impacts have not been measured in any detail, in formats that have been made publicly available; how many people have actually been lifted out of poverty as a result of these reforms? Questions also remain about how much regulation should be cut; there is consensus, for example, that enterprises should always conform with minimum labour standards (so less regulation is not always necessarily better than more regulation). The discussion about how much regulation is appropriate is still in progress.

<sup>74</sup> [www.DoingBusiness.org](http://www.DoingBusiness.org)

<sup>75</sup> Keppel, Binh and Spatz, 2006. Streamlining Business Registration and Licensing Procedures: Experiences from the Philippines and Vietnam. [www.businessenvironment.org/dyn/be/docs/117/Session1.1Paper1.1.1Keppel.pdf](http://www.businessenvironment.org/dyn/be/docs/117/Session1.1Paper1.1.1Keppel.pdf)

<sup>76</sup> Jacobs, Scott, 2006. The Regulatory Guillotine in Three Transition and Developing Countries. [www.businessenvironment.org/dyn/be/docs/81/Session4.2JacobsDoc.pdf](http://www.businessenvironment.org/dyn/be/docs/81/Session4.2JacobsDoc.pdf)

<sup>77</sup> Vignjevic, Jasmina, 2005. Engaging the Private Sector in Business Environment Reforms: Experience from Southeast Europe. [www.businessenvironment.org/dyn/be/docs/76/Session2.4VignjevicDoc.pdf](http://www.businessenvironment.org/dyn/be/docs/76/Session2.4VignjevicDoc.pdf)

<sup>78</sup> Haggerty et al, 2005. Cutting Red Tape: Simplifying Regulation at the Municipal Level in Latin America. [www.businessenvironment.org/dyn/be/docs/81/Session4.2HaggartyDoc.pdf](http://www.businessenvironment.org/dyn/be/docs/81/Session4.2HaggartyDoc.pdf)

<sup>79</sup> Omran and Waly, 2005. Donors Business Environment Reform Interventions in Egypt. [www.businessenvironment.org/dyn/be/docs/75/Session2.3Omran-WaliDoc.pdf](http://www.businessenvironment.org/dyn/be/docs/75/Session2.3Omran-WaliDoc.pdf)

<sup>80</sup> Muir, Russell and Xiaofang Shen, 2005. Land Markets: Improving Access to Land and Buildings by Investors. [www.businessenvironment.org/dyn/be/docs/77/Session3.1Shen-MuirDoc.pdf](http://www.businessenvironment.org/dyn/be/docs/77/Session3.1Shen-MuirDoc.pdf)

228. Nonetheless, advocates of cutting red tape stress that countries with less red tape grow faster; they have calculated that, if a country reformed its regulations sufficiently to move from the bottom quartile to the top one in the Doing Business rankings, its growth rate could be expected to increase by 2.2 percentage points<sup>81</sup>.

The causality of this assertion remains somewhat open, however; the place of business environment reform in value chain development also remains rather unexplored, for the time being<sup>82</sup>, with many agencies focusing instead on cutting red tape for all kinds of business.

---

<sup>81</sup> Doing Business 2007, by Caralee McLiesh, 2006.  
[www.businessenvironment.org/dyn/be/docs/132/Day1PlenaryPresentationMcLiesh.pdf](http://www.businessenvironment.org/dyn/be/docs/132/Day1PlenaryPresentationMcLiesh.pdf)

<sup>82</sup> Even though McKinsey's Institute and others assert that perhaps a majority of critical business environment constraints are sector-specific.



## O. Summary of Results to Date

*This Chapter provides some overview and final observations about the various, specific impact assessments reported in the previous Chapters.*

229. Briefly, and as a further technical resume of some of the information presented above, the following Table summarises which aspects of measurement are addressed most explicitly by each of the studies outlined above.

230. The single indicator reported most often is probably the Return on Investment – relating the returns to the private sector resulting from the original investment of development funds. This is broadly similar to the private sector’s own Internal Rate of Return (IRR): a tool for testing whether a future investment meets the company’s minimum standards, in terms of future profitability. One major advantage of its application in this setting is that it is broadly possible to measure both the size of the investment and the additional revenues for (or costs saved by) the target group as a result.

231. As a management tool, therefore, it has much to commend it – as Technoserve and others have noted. As a means of reporting impacts, however, it is not very helpful, because it does not indicate who has earned the returns – or what their pro-poor impact is likely to be. For example, the main

beneficiaries might be very large companies, which should not be expecting to profit at the tax-payers’ expense. But it may nonetheless be possible to construct a case for the intervention, on the basis that:

- the company would not have made the change, investment or innovation if the public funds had not been contributed; and
- the poor were benefiting very substantially (and preferably quantifiably) through increased markets for their produce or labour, improved services, increased tax revenues to government, etc.

232. If it were possible to relate the ‘returns’ on the investment, for example through increased employment and incomes, to numbers of people lifted out of poverty, that would be enormously helpful to donor agencies seeking to showcase their achievements domestically. The tax-paying public needs a clear picture of people emerging out of poverty on a large scale, to continue to support such investments.

233. Measurement of the ‘leverage’ has a similar problem as IRR, in the sense that knowing the amount of investment by the private sector, relative to the amount invested by the public sector, tells us little

	Bench- marking globally	Integrated M and E system	Capturing extent of 'crowding in'	Estimating upstream / downstream effects	Disaggrega- tion of data by sex	Focus on attribution and displace- ment
SEDF Bangladesh	X	X				
Prosperity Initiative, Mekong	X		X	X	X	
Katalyst Bangladesh			X			X
Challenge Funds				X	X	

Photo: Charles Bodwell



about the developmental impacts of the total investment. Indeed, although more leverage is generally implied to be better, there are limits; if the leverage is very high, the observer may question whether the public funds contributed made any difference,

either to the decision of the private sector to invest, or to the developmental impacts of that investment.

234. Alternatively, the measurement of jobs created may be more promising, since it is reasonable to assume, at least in some cases, that the great majority of the jobs created were for people who were living around the poverty line. Further research would be needed to quantify this assumption, and the definition of a 'job' remains problematic; but such a yardstick would be of particular interest to the governments of developing countries.

235. The following Table summarises some of the data presented earlier in the text, and includes a few additional statistics from the first Donor Committee conference on BDS<sup>83</sup>.

236. Clearly, the data presented above are relatively crude – for example in the sense that they are not all adjusted for 2007 values of the US dollar. The point is that there are orders of magnitude in the differences between the various numbers, so the relative values are likely to be significant – even if the margin for error is +/- 50%.

237. One exception to this may be in the multiplier to be used in calculating the indirect impacts, since this is much more difficult to define, even within a given industry. Besides, the possible multipliers can be large; Unilever Indonesia and Oxfam, for example, concluded that there were about 300,000 FTE jobs in the value chain of Unilever Indonesia, relative to the core workforce of about 5,000 people<sup>84</sup> – suggesting a multiplier of about 60.

Intervention	Cost/job	Notes
SEDF value chain development in Bangladesh	374	
Mekong Bamboo Consortium	50	Projected
Katalyst value chain development in Bangladesh	100	
Entrepreneurship training in India	25	1968 value
Entrepreneurship training in India	150	1990 value
PSOM matching grant PPP programme	2,700-4,100	All jobs (est.)
DFID Business Linkage Challenge Fund (n.b. jobs created <i>or retained</i> )	1,300	Direct jobs only
	200	All jobs (est.)
ApproTEC pumps etc. in Kenya	340	1998 value
K-MAP mentorship in Kenya	400	1998 estimate
AMKA marketing for SMEs in Tanzania	760	1998 estimate

<sup>83</sup> Tanburn, 1998. BDS: How sustainable can they really be? Report on the Harare Conference, October 1998. Committee of Donor Agencies for SED/DFID/ILO. <http://www.enterprise-development.org/events/event.asp?eventid=16>

<sup>84</sup> Clay, 2005. Exploring the Links Between International Business and Poverty Reduction: A Case Study of Unilever in Indonesia. Oxfam GB, Novib, Unilever. [http://www.cpi.cam.ac.uk/cms/pdf/unilever\\_indonesia\\_study.pdf](http://www.cpi.cam.ac.uk/cms/pdf/unilever_indonesia_study.pdf)

238. Similarly, Heineken<sup>85</sup> estimated that there were 6,780 people employed by suppliers, and distributors and retailers of beer produced by Sierra Leone Breweries, relative to 175 people employed directly in the brewery – a multiplier of 39.

239. These numbers illustrate how sensitive the overall calculation of jobs created is, to the multiplier

to be used; it would therefore be very helpful if a neutral agency could determine appropriate multipliers, and also provide some basic quality control for numbers being generated by each agency. This agency could also provide an important way for individual agencies to limit the political risk of adverse comparison, for example through aggregating the outcomes across agencies.

---

<sup>85</sup> Economic impact of Heineken in Sierra Leone. Amsterdam, November 2006. <http://www.bidnetwork.org/artefact-42202-en.html>

## P. The Future

*Pressures to measure and report on results are increasing, as new players join the discussion. This Reader aims to facilitate the process, leading to more serious efforts in this field.*

240. The measurement of results brings together many different aspects of development; the challenge is not just a technical one, but also has political, institutional and human dimensions. These are not often included in the debate, but do need to be addressed.

241. In particular, there are trade-offs to be made – for example between cost and rigour. Similarly, funds spent in measuring and communicating results will come from sources that otherwise might have been used to achieve yet more results. As yet, there has not been any in-depth discussion about these trade-offs, and where lines should be drawn; however, such discussion seems likely in the near future.

242. Pressure to report on results is growing; the increasing profile of development aid in the minds of the public, the approaching MDG deadline, and the arrival of new types of development organisation, are creating substantial pressures to paint a more convincing picture of the results being achieved. These results need to be communicated in ways that both reflect some degree of rigour (or at least honesty), and can be readily understood and appreciated by people who have never worked in development.

243. Some would argue that, the more you examine the issue of reporting on impacts, the more complex it becomes. Defining and measuring poverty, and establishing attribution, are just two of the aspects that are multi-dimensional, and that require a high level of expertise. But the tax-paying public do not appreciate these complexities, and see a lack of 'headline numbers' or sound-bites on results as a possible indication that they don't exist.

244. Practitioners are therefore in an uncomfortable position, of wanting to report results in ways that are convincing and credible, without over-simplifying a complex situation. Development agencies are often competing for funds, and there are therefore few incentives to publish information about results; either the results may appear to be poor, or they may appear so good as not to be credible. This has meant that practitioners would rather keep silent, than talk about their achievements in ways that are approximately right.

245. Beyond that, the lack of agreement around what to measure, and how to measure it, means that the observer has no way at present to tell whether any particular intervention, agency or approach was more or less effective than any other one (taking into account context-specific factors, of course). So there is no reasoned debate about which approach works best in which circumstances; instead, there is a rush from one paradigm to the next, often within the space of 3-4 years.

246. This Reader argues that, in such a situation, no-one gains – in particular, those living in poverty who would gain the most from effective development work must sometimes be bemused by the shifts in focus that they may observe. Above all, we owe it to them to work in a more intentional way towards some form of measurement of results, which would allow a more reasoned discussion, and some element of comparison and benchmarking.

247. Reflecting the growing interest in the topic generally, a consortium of donors is currently launching an International Initiative for Impact Evaluation (3IE); the consortium currently includes the Governments of Mexico, Uganda, UK, Netherlands and Canada, plus the African Development Bank and the Hewlett and Gates Foundations. The institutional home for this initiative has not yet been decided<sup>86</sup>.

<sup>86</sup> [www.cgdev.org/section/initiatives/\\_active/evalgap](http://www.cgdev.org/section/initiatives/_active/evalgap)

248. This type of initiative could provide the 'home' for standardising measurements that was referred to in the previous Section; specifically, the most important functions, which are needed on an inter-agency basis, could include:

- building agreement around key impact variables, such as which indicators to measure (jobs? people out of poverty?), and how to define them in ways that are practical, yet linked to theoretical models
- validating proxies that could be used to measure these key variables on an approximate basis, at low cost
- building agreement around which input costs should be included in the calculations (programme management? local contributions?)
- defining minimum standards for approximate measures of attribution, displacement, deadweight etc.
- research to generate reasonable multipliers, for example in the case of the likely indirect impacts upstream and downstream of the target group, in selected sectors (similarly for impacts on the local economy as a result of increased incomes, taxes etc.)
- development of additional, affordable methodologies to measure 'crowding in', 'copy-cats' and other spontaneous replications and market development effects that are the aim of much PSD work
- support to agencies to validate their own approximate calculations, and to give them greater credibility in the development community

- where needed, support to agencies to make comparative calculations more anonymous and aggregated, to limit potential reputational risk involved in being more transparent and open (at least in the short term)

249. It seems unlikely that a generic initiative such as 3IE can address the particular opportunities and challenges involved in measuring the results achieved in PSD. A more specialised initiative is therefore probably required – similar, perhaps, to the current work of CGAP in this area for microfinance institutions.

250. One idea, proposed by Jim Tomecko during the Chiang Mai Seminar, was the creation of a system for certifying the quality of measurements made. This would require the codification of 'good practice' in measurement, and a cadre of evaluators certified to sign off on the methodologies used; such a system could reduce transaction costs for programme managers, while giving donors and others a degree of assurance that the numbers being generated were credible.

251. Practitioners can contribute to this discussion, now, by proposing and piloting ways to report on their achievements, which allow observers to gain some idea of how effective they are being, and how impressive their achievements really are. Ultimately, bilateral donors can probably contribute the most to this debate, since they are not directly competing for funds or constituencies, and can therefore share their findings more openly.

# Annex A: New or updated entries in the last year on [www.Value-Chains.org](http://www.Value-Chains.org)

(currently the same site as [www.BDSKnowledge.org](http://www.BDSKnowledge.org))

*Note: For Readers unable to use the hyperlinks, using the Search function for the title should also work well.*

## Global Documents

- Global Value Chains in the Agrifood Sector, IDS, UNIDO 2006 (entered 25 Aug 2007)
- Upgrading in Global Value Chains, ILO 2004 (entered 25 Aug 2007)
- Enhancing the Role of SMEs in Global Value Chains, OECD Conference June 07 (entered 25 Aug 2007)
- The Future of Small Farms, IFPRI Conference 2005 (entered 20 Aug 2007)
- GTZ Value Chains Conference, Berlin, May 2007 (entered 31 Jul 2007)
- GTZ Conference: Local and Regional Economic Development in Asia, Vientiane, November 2006 (entered 30 Jun 2007)
- Analyzing the Contribution of Business Services to European Economic Growth, 2007 (entered 30 Jun 2007)
- Food Quality and Safety Standards - A Practitioners' Reference Book, GTZ, 2007 (entered 13 Jun 2007, last updated 25 Jun 2007)
- Upgrading to Compete - Book from IADB 2006 (entered 4 Jun 2007)
- Local Economic Development - Key Documents and Links (entered 14 May 2007)
- Business Linkages: Lessons, Opportunities and Challenges, 2007 (entered 11 May 2007, last updated 8 Jun 2007)
- Microfinance and BDS in Europe: A Guide to Good Practices, 2007 (entered 9 May 2007)
- Market Development in Crisis-Affected Environments, SEEP MDWG 2007 (English/French) (entered 5 May 2007, last updated 11 May 2007)
- Poverty Reduction through Small Enterprises, ILO 2006 (entered 2 May 2007)
- Eighth Annual Seminar on Developing Service Markets and Value Chains - Chiang Mai, 24-28 September 2007 (entered 17 Apr 2007, last updated 22 Aug 2007)
- GTZ Value Chain Info Newsletter (entered 23 Mar 2007)
- International Standards in Products, Production Methods and Services - Links and Key Documents, AGEG 2007 (entered 3 Mar 2007)
- Linking Farmers to Markets, FAO, 2007 (entered 3 Mar 2007, last updated 25 May 2007)
- Value Chain Development and Rural Employment: Round Table at the IFAD Governing Council, February 2007 (entered 27 Feb 2007, last updated 28 Feb 2007)

- First Regional Conference on Developing Service Markets and Value Chains, Jordan, February 2007 (English / Arabic) (entered 25 Feb 2007, last updated 19 Mar 2007)
- Making Markets Work for the Poor in Eastern and Southern Africa, ComMark Seminar, 12-14 March 2007 (entered 27 Jan 2007, last updated 17 May 2007)
- Shaping Value Chains for Development - Special Issue of the European Journal of Development Research, 2007 (entered 17 Jan 2007)
- Services Inputs and Firm Productivity in Sub-Saharan Africa: Evidence from Firm-Level Data, World Bank, 2006 (entered 17 Jan 2007)
- Donor Approaches to Supporting Pro-Poor Value Chains, Altenburg / Donor Committee, 2006 (entered 16 Jan 2007, last updated 2 Apr 2007)
- Workshop: Making value chains work for the poor, SDC, 2007 (entered 14 Jan 2007, last updated 27 Apr 2007)
- Lessons Learned on MSE Upgrading in Value Chains, ACDI/VOCA USAID AMAP 2007 (English/French) (entered 8 Jan 2007, last updated 15 Aug 2007)
- Investing in the Majority - SEEP Annual General Meeting 23-27 Oct, 2006 (entered 17 Dec 2006)
- Private Sector Development in Post Conflict and Peace Building Situations - Some Key Links and Documents (entered 7 Nov 2006, last updated 5 May 2007)
- Second National Conference of the BDS Donor Coordination Group, Kenya 2006 (entered 21 Oct 2006, last updated 18 Mar 2007)
- Proceedings of the Frontis Workshop on Agro-food Chains and Networks for Development, Wageningen, The Netherlands, 6-7 September 2004 (entered 13 Oct 2006)
- ILO Guide for Value Chain Analysis and Upgrading, 2006 (English/French) (entered 11 Oct 2006, last updated 28 Mar 2007)
- Micro & Small Enterprises: Unexplored Pathways to Growth, IRIS 2006 (entered 6 Oct 2006, last updated 13 Oct 2006)
- MSEs, Dynamic Economic Growth, & Poverty Reduction: A Review of the Conceptual & Empirical Effects of MSEs on Development , IRIS 06 (entered 6 Oct 2006)
- Rural and Agricultural Finance, USAID Resources (entered 28 Mar 2006, last updated 20 Aug 2007)
- Seventh Annual Seminar on Developing Service Markets and Value Chains - Chiang Mai, 18-22 September 2006 (English / French / Spanish) (entered 15 Mar 2006, last updated 2 Jun 2007)
- Strategic Alliances for Financial Services and Market Linkages in Rural Areas, SEEP 2005-7 (entered 16 Dec 2005, last updated 17 Sep 2006)
- Making Service Markets Work for the Poor: Some key links (entered 22 Nov 2005, last updated 20 Jul 2007)
- Linking Small Firms to Competitive Strategies - USAID Breakfast Seminar Series 2005-7 (entered 15 Nov 2005, last updated 20 Aug 2007)
- Calls for Papers / Participation (entered 11 Nov 2005, last updated 30 Jun 2007)
- Value chains: Some key links (entered 9 Sep 2005, last updated 25 Aug 2007)
- Les BDS: L'Actualité des services aux entreprises (entered 18 Feb 2005, last updated 30 Jun 2007)

- New opportunities related to value chains, service markets, MMW4P, etc. (entered 12 Aug 2004, last updated 31 Jul 2007)
- Top ten entries, in terms of visits in June 2007 (entered 21 Apr 2004, last updated 1 Jul 2007)
- Top ten entries, in terms of total visits to date (1 July 2007) (entered 21 Apr 2004, last updated 1 Jul 2007)
- Future events (entered 22 Dec 2003, last updated 23 Aug 2007)
- Who profits from linking BDS to financial services? Sievers and Vandenberg, 2004 (entered 25 Sep 2003, last updated 24 May 2007)

## Market Assessment

- Analyse de la filière Charbon de Bois au Sénégal, WRI 2006 (entered 20 Aug 2007, last updated 21 Aug 2007)
- Philippines Processed Banana Value Chain Analysis, SDCAsia 2006 (entered 18 Jun 2007)
- Artisanal Textiles value chain analysis, Senegal, 2006 (entered 18 May 2007, last updated 8 Jun 2007)
- Hibiscus (bissap) value chain analysis, Senegal, 2006 (French) (entered 17 May 2007, last updated 18 May 2007)
- Cashew value chain analysis, Senegal, 2006 (French) (entered 17 May 2007, last updated 18 May 2007)
- Mango value chain analysis, Senegal, 2006 (French) (entered 17 May 2007, last updated 18 May 2007)
- The Prosperity Initiative (research), 2007- (entered 5 May 2007, last updated 6 May 2007)
- Analysis of the Integration of MSEs into Value Chains, Tanzania, USAID AMAP, 2006 (entered 5 May 2007)
- Assessment of the Commercial Private Sector for Health Care Products in Bangladesh, AFE USAID 2006 (entered 5 May 2007)
- Facilitating Market Integration of the Upland Poor into Bamboo Value Chains, Viet Nam M4P (ADB, DFID) 2006 (entered 25 Apr 2007)
- Guide to Market Research for Agro-Processors, FAO 2003 English/Spanish (entered 22 Feb 2007, last updated 3 Mar 2007)
- Enterprise for Pro-Poor Growth Socio-Cultural Assessment, ILO, Sri Lanka 2006 (entered 7 Feb 2007)
- Madagascar Aromatic and Medicinal Plant Value Chain Analysis, ACDI/VOCA, IRG, USAID AMAP 2006 (entered 19 Jan 2007)
- Angola Enterprise Programme BDS Market Assessment, UNDP/Chevron/GoA, 2006 (entered 15 Dec 2006)
- Haitian Handicraft Value Chain Analysis, USAID AMAP, 2006 (entered 23 Nov 2006, last updated 14 Dec 2006)
- Rice Value Chains in Dien Bien Province Viet Nam, Agrifood Consulting, 2006 (entered 2 Nov 2006)
- Northeast Thailand Rice Value Chain Study, Agrifood Consulting Int., 2005 (entered 2 Nov 2006)
- Livestock sector study, USAID, Kenya, 2006 (entered 27 Oct 2006)
- Feasibility study on production of wines and jams, Zambia, 2004 (entered 13 Oct 2006)



- Studies for indigenous fruit markets in Tanzania, Zimbabwe and Zambia, 2005 (entered 13 Oct 2006)
- Feasibility study on production of indigenous fruit juice concentrate, Tanzania, 2004 (entered 13 Oct 2006)
- Feasibility study on production of fruit juice concentrate, Malawi, 2004 (entered 13 Oct 2006)
- Nature-Oriented Tourism in Ecuador: Applying the Value Chain and Nature, Wealth and Power Frameworks, ACDI/VOCA. IRG, USAID (entered 6 Oct 2006)
- Indonesia Cocoa Bean Value Chain Case Study, USAID AMAP, AFE 2006 (entered 6 Oct 2006)
- The Participation of the Poor in Supermarkets and other Distribution Value Chains, Viet Nam, M4P, 2005 (entered 2 Oct 2006)
- Mercy Corps, AED Assessment of the Hot Pepper Value Chain, Liberia, 2006 (entered 30 Sep 2006)
- Value Chain Analysis with a Financial Services Lens, EDA, India 06 (entered 17 Sep 2006)
- Participation of the Poor in the Tea Value Chain Vietnam, M4P (ADB, DFID) 2004 (entered 19 Jan 2006, last updated 2 Nov 2006)
- Value Chain Analysis (entered 16 Apr 2005, last updated 16 Aug 2007)
- GTZ and Swisscontact Viet Nam - BDS Market Assessment 2001 (entered 2 Sep 2003, last updated 10 Mar 2007)

## Project Design

- Philippines Banana AgriChain Competitiveness Enhancement (B-ACE) Design, SDCAsia 2007 (entered 18 Jun 2007)
- The Prosperity Initiative (design), 2007- (entered 5 May 2007, last updated 13 Jun 2007)
- Financial Integration, Economic Strengthening & Broad-Based Dissemination (FIELD-Support), USAID 2006-10 (entered 5 May 2007)
- AusAID Enterprise Challenge Fund for the Pacific and South-East Asia, 2006- (entered 3 Apr 2007, last updated 4 Apr 2007)
- Enterprise for Pro-Poor Growth, ILO, Sri Lanka 2005-8 (entered 31 Jan 2007, last updated 16 Apr 2007)
- Agricultural Sector Program Support, Mozambique, DANIDA 2006 (entered 17 Sep 2006)

## Implementation

- Enterprise for Pro-Poor Growth Progress Reports, ILO, Sri Lanka 2005-8 (entered 6 Feb 2007, last updated 16 Apr 2007)
- Market Development in Crisis-Affected Environments, SEEP AGM 2006 (entered 17 Dec 2006, last updated 1 Jan 2007)
- Economic Support and Income Generation Programming for HIV and AIDS Impacted Communities, SEEP AGM 2006 (entered 17 Dec 2006)
- Bringing Knowledge to Vegetable Farmers, KATALYST Bangladesh 05 (entered 17 Sep 2006)

- KATALYST - Implementation, Bangladesh 2002-2007 (entered 17 Sep 2006)
- The Thai German Programme for Enterprise Competitiveness, GTZ 2004-2012 (entered 16 Sep 2006, last updated 16 Aug 2007)
- Local Market Development, Helvetas/ICCO Kyrgyz Republic, 2006-7 (entered 28 Jun 2006, last updated 13 Apr 2007)
- Philippines SME Development and Sustainable Employment Promotion - Tourism Sector, GTZ, GFA (entered 11 Jun 2006, last updated 10 Jul 2007)
- Making Markets Work Better for the Poor (Implementation), Viet Nam, Laos, Cambodia (ADB, DFID) 2003-2007 (entered 17 May 2006, last updated 29 Mar 2007)
- Public Private Partnerships (PPP), GTZ, Viet Nam, 2006 (entered 7 May 2006, last updated 1 Oct 2006)
- Accelerating Shared Growth - Making Markets Work for the Poor in South Africa, ComMark 2006 (entered 5 May 2006, last updated 26 Dec 2006)
- Access to Contemporary Markets for Homebound Women Embroiderers in Pakistan, MEDA, ECDI 2003-2007 (entered 13 Nov 2005, last updated 9 Feb 2007)
- Kenya BDS Implementation, USAID / Emerging Markets Group, 2004-8 (entered 1 Mar 2004, last updated 31 Jul 2007)

## Impact Assessment

- Common Monitoring System for Proyectos de Integracion Productiva, IADB 2005 English/Spanish (entered 23 Aug 2007)
- ADB Evaluation Methods and Guidelines (entered 30 Jun 2007)
- IFC Results Measurement Events, 2007 (entered 29 Mar 2007, last updated 17 May 2007)
- Enterprise for Pro-Poor Growth Baseline Reports, ILO, Sri Lanka 2006 (entered 2 Feb 2007)
- Private Sector Development Impact Assessment Primer Series, USAID AMAP (entered 22 Jan 2007, last updated 11 Apr 2007)
- The Economic Impact of Cluster Initiatives under the Competitiveness Initiative Project, USAID 2003 (entered 2 Jan 2007)
- Assessing the Impact of the Micro and Small Enterprise Trade-led Growth Project of USAID/BRAZIL 2004-6 (entered 23 Nov 2006)
- Impact Assessment On-line Discussion, USAID microLINKS.org, Sept 06 (entered 14 Nov 2006)
- PROFIT Zambia Impact Assessment Baseline Research Design, USAID, DAI, 2006 English/French (entered 2 Nov 2006, last updated 27 Nov 2006)
- The Thai German Programme for Enterprise Competitiveness - Impact, GTZ 2005-6 (entered 16 Sep 2006, last updated 13 Oct 2006)
- Listener Survey, MSE Radio Programmes in Uganda, 2004 (entered 29 Jan 2005, last updated 20 Jul 2007)

## Final Documentation

- Making Markets Work for the Poor: Case Studies for SDC, 2007 (entered 9 Jul 2007, last updated 10 Jul 2007)
- Local/Regional Economic Development in South-Eastern Europe, GTZ 2006 entered 5 Jun 2007)
- Handmade Paper in Nepal, GTZ 2007 (entered 21 May 2007, last updated 22 May 2007)
- The Experience of IFC/SEDF with Sector Development and Business Service Strategy in Bangladesh, 2006 (entered 27 Nov 2006)
- RESTART - Help for Tsunami Impacted SMEs, GTZ, Thailand 05 (entered 12 Sep 2006)
- Exploring the Links between International Business and Poverty Reduction: Unilever in Indonesia, 2005 (entered 20 Dec 2005, last updated 20 Jan 2007)
- Value Chains and BDS Development: Linking Communities to Mainstream Markets in Mindanao, SDCAsia, 2002-6 (entered 18 Jan 2005, last updated 12 Jun 2007)



ISBN 978-92-9049-437-9

Design and printing: International Training Centre of the the ILO, Turin Italy