

60084

SP DISCUSSION PAPER

NO. 1107



## Results Readiness in Social Protection & Labor Operations

Laura Rawlings, Maddalena Honorati,  
Gloria Rubio and Julie Van Domelen

February 2011

# **RESULTS READINESS IN SOCIAL PROTECTION AND LABOR OPERATIONS**

**Social Protection and Labor, Human Development Network (HDNSP)**

**February 2011**

## ABBREVIATIONS AND ACRONYMS

AFR	-	Africa Region
AIN-C	-	Estrategia de Atención Integral en Nutrición a Nivel Comunitario
CBO	-	Community-Based Organization
CCT	-	Conditional Cash Transfer
CDD	-	Community Driven Development
CPS	-	Country Partnership Strategy
DEC	-	Development Economics Vice Presidency
DIME	-	Development Impact Evaluation Initiative
DPL	-	Development Policy Lending
EAP	-	East Asia and the Pacific Region
ECA	-	Europe and Central Asia Region
ERL	-	Emergency Recovery Loan
FSRDC	-	Le Fonds Social de la République Démocratique du Congo
ICR	-	Implementation Completion and Results Report
IDA	-	International Development Association
IDI	-	Institutional Development Indicators
ILO	-	International Labour Organization
ISR	-	Implementation Status and Results Report
KPI	-	Key Performance Indicator
LAMIS	-	Local Authority Management Information System
LCR	-	Latin America and Caribbean Region
MIS	-	Management Information System
MNA	-	Middle East and North Africa Region
M&E	-	Monitoring and Evaluation
NGO	-	Non-Governmental Organization
OPCS	-	Operations Policy and Country Services
PAD	-	Project Appraisal Document
PDO	-	Project Development Objective
PRSC	-	Poverty Reduction Support Credit
QER	-	Quality Enhancement Review
QUIBB	-	Household Core Welfare Indicators Survey
RBIL	-	Result-Based Investment Lending
SAL	-	Structural Adjustment Loan
SAR	-	South Asia Region
SFD	-	Social Fund for Development
SMART	-	Specific, Measurable, Attributable, Realistic, Targeted
SP	-	Social Protection
SSIU	-	Sector Strategy Implementation Update
SSN	-	Social Safety Net
TA	-	Technical Assistance
TOR	-	Terms of Reference
TTL	-	Task Team Leader

# RESULTS READINESS IN SOCIAL PROTECTION AND LABOR OPERATIONS

## CONTENTS

Summary .....	iv
I. Context, Objectives and Methodology .....	1
II. Project Development Objectives .....	3
III. Key Performance Indicators.....	4
IV. Monitoring and Evaluation Arrangements .....	7
V. Use of Results within the World Bank .....	10
VI. Key Themes and Recommendations by Lending Instrument .....	11
VII. Conclusions and Recommendations .....	14
Annex 1. Methodology and Coverage.....	
Annex 2. Project Worksheet for Portfolio Review of PADs, ISRs.....	
Annex 3. Output Tables from Portfolio Review of PADs, ISRs.....	
Annex 4. Basic Vocabulary for Results Based management and Evaluation.....	
Annex 5. Good Practice WB project Results Frameworks.....	

*This work was led by Laura Rawlings (TTL) and Maddalena Honorati (co-TTL) with inputs from Ravindra Cherukupalli, Prem Garg, Gloria Rubio and Julie Van Domelen, under the overall guidance of Arup Banerjee (Director, HDNSP) and Bassam Ramadan (Manager, HDNSP). Excellent assistance was provided by Francine A. Pagsibigan and Dung Thi Ngoc Tran. We would like to thank the peer reviewers, Aline Coudouel, Benjamin Loevinsohn and Gail Richardson, as well colleagues for their valuable comments, including Harold Alderman, Rita Almeida, Colin Andrews, Ariel Fiszbein, Margaret Grosh, Jody Kusek, Phillippe Leite, Kathy Lindert and Lucian Pop. The views reflected in this report are those of the authors.*

## RESULTS READINESS IN SOCIAL PROTECTION AND LABOR OPERATIONS

### SUMMARY

***The main focus of the Social Protection and Labor portfolio is on strengthening client's institutional capacity in the design and implementation of programs, but projects are not well equipped to track progress in this area.*** Correspondingly, there is a need to strengthen approaches to measuring and monitoring a “missing middle” of service delivery, precisely those areas for which counterpart institutions are responsible during the course of a project. In particular, better measures of the primary functions of social protection and labor agencies are needed, such as identifying and enrolling beneficiaries, targeting, payment systems, fraud and error control, performance monitoring of service delivery providers, responsiveness to citizens, transparency, efficiency, management information systems and monitoring and evaluation systems.

***New World Bank initiatives - particularly standard core indicators by sector and the introduction of results- based investment lending - call for substantial improvements in the use of monitoring and evaluation (M&E). Present practice in the use of M&E, notably during project implementation, is weak:*** close to 40 percent of indicators lack baseline data; 12 percent of outcome indicators (and 19 percent of intermediate indicators) listed in the project appraisal documents (PADs) are not listed in the implementation status and results report (ISRs); and only 21 percent of projects have regularly updated key performance indicators (KPIs). Development policy lending operations (DPLs) are particularly weak and require a tailored approach to M&E. Overall, project supported M&E could be improved through the better use of ISRs and greater specificity, commitments and follow-through regarding the implementation of M&E arrangements in client countries, notably concerning data sources, frequency and responsibility of reporting, and expected use of the indicators by stakeholders.

***Strengthening monitoring and evaluation arrangements in client countries must be a priority.*** There is little attention to this issue in current practice, but is it the area with potentially the greatest payoffs. Strengthened sectoral M&E systems are a foundation for improved performance in the sector overall. Social protection would particularly benefit from improved sectoral results-based management given the diversity of programs and policies that comprise the sector. This would be a powerful tool for moving from a often ad-hoc project approach to one supporting social protection programs and systems.

***Impact evaluations are included in about half of projects and should continue to be used selectively and strategically,*** particularly when the program is innovative, replicable and/ or scalable to reach a broader set of beneficiaries, addresses a knowledge gap and is likely to have a substantial policy impact. Structuring evaluations around core themes with common outcome measures is fundamental to building a global knowledge base on development effectiveness.

# RESULTS READINESS IN SOCIAL PROTECTION AND LABOR OPERATIONS

## I. CONTEXT, OBJECTIVES AND METHODOLOGY

1. **The Bank is increasingly focused on promoting results-based management, both as a corporate strategy and as an element of support to client countries.** Accordingly, there has been a growing focus on the outcomes and impacts of Bank operations and on strengthening the links between these operations and country development goals. Strengthening project level monitoring and evaluation (M&E) can be a platform for expanding knowledge on development effectiveness, improving broader sectoral M&E in client countries, supporting public sector management and making public programs and more accountable to citizens and international partners.

2. **Upcoming Bank policies will increase the need for effective M&E in projects and promote opportunities to strengthen clients' results based management systems.** The current discussions taking place to develop a new results-based investment lending (RBIL) instrument as well as the Operations Policy and Country Services (OPCS) mandate to develop standardized operational indicators by sector will put M&E at the center of operational policies. As the corporate expectation on results-based approaches intensifies, task teams are increasingly seeking technical guidance and examples of effective practice.

3. The Social Protection and Labor (SP+L) sector is building on a comparatively strong base with respect to approaches to monitoring and evaluation. Many SP+L projects have been path-breaking in fostering better monitoring and evaluation frameworks and more effective use of results. The SP+L sector has comparatively good ratings for project level M&E, and has been the focus of two meta-analyses of the development effectiveness of particular programs, covering social funds published in 2001 and conditional cash transfers (CCT) published in 2009, based on robust experimental and quasi-experimental impact evaluations.

4. **This "Results Readiness" work assessed progress to date on results-based management in the SP+L portfolio and generated operationally relevant knowledge on how to strengthen M&E.** This review of results frameworks in the Bank's SP+L portfolio seeks to answer the question "How well are SP+L projects performing with respect to their focus on results?" Specifically, the review takes stock of the status and quality of M&E in the SP+L portfolio, including both investment and policy-based lending, and then develops guidance for task teams based on the review. The work identifies trends, strengths and weaknesses, and good practice M&E approaches and indicators to incorporate a better results focus in project design and implementation.

5. **Phase I of this process, reflected in these findings, consisted of a rapid desk review of the current SP+L portfolio.** The cohort consists of all 70 SP+L mapped projects approved in the fiscal years 2005-2009 period, with an additional sample of 7 DPLs with a substantial safety nets component mapped to other sector boards, for a total of 77 operations (see Table 1 below and Annex 1). The primary focus is on design and implementation issues using Project Appraisal Documents (PADs), Implementation Supervision Reports (ISRs) and Implementation Completion and Results Reports (ICRs),

with analysis of final results limited by the small number of ICRs (13) that have been produced so far in this relatively young cohort.

**Table 1. Project Cohort in “Results Readiness” Work**

	Investment Lending		DPLs	Total
	Non-emergency	Emergency		
<b>Total</b>	<b>43</b>	<b>13</b>	<b>21</b>	<b>77</b>
Labor Markets	5	0	4	9
Service Delivery	5	1	6	12
Pensions	2	0	1	3
Social Funds	7	5	1	13
Social Safety Nets	24	7	9	40

6. The products of this Phase I work are a digestible set of resources for task teams and include:
- a) This Summary Report and a full length Powerpoint on main findings and recommendations;
  - b) Technical Guidance Notes for Task Teams in the areas of:
    - Safety Nets
    - Labor Market Programs
    - Service Delivery
    - Social Funds<sup>1</sup>
  - c) Data on Results in SP+L Operations based on the development and application of a standard instrument used to collect performance data across the SP+L portfolio (see Annexes 2-3)

7. The assessment followed the project cycle through both design (assessing PADs) and implementation (assessing ICRs and available ISRs), with reviews of PDOs, outcome and intermediate outcome KPIs, M&E arrangements and the use of M&E within the World Bank. In response to requests made at the Concept Note review, there was a special emphasis on DPLs, impact evaluations, and indicators of institutional development. Performance ratings are assessed on a range from 1 (highly satisfactory) to 6 (highly unsatisfactory), with a 1 or 2 denoting *strong performance*, 3 indicating *room for improvement* and 4-6 *not meeting expectations*. Table 2 summarizes the main results of the performance assessment.

**Table 2. Percent of SP+L Projects with Strong Performance on Results**

	Quality of PDO formulation	Quality of outcome indicators	Quality of intermediate outcome indicators	KPIs updated regularly in ISRs	M&E arrangements at design	M&E at implementation
<b>% of projects with strong performance</b>						
<b>Total</b>	<b>44</b>	<b>43</b>	<b>38</b>	<b>21</b>	<b>51</b>	<b>22</b>
Labor Markets	67	33	22	11	44	44

<sup>1</sup> There are only three pension projects mapped to the SP+L sector from FY05-09. Therefore, while pensions are included in the statistical analysis of the SP+L portfolio, the cohort is insufficient to provide technical guidance in this thematic area. All Technical Guidance Notes are available at [www.worldbank.org/sp](http://www.worldbank.org/sp).

Service Delivery	8	50	17	25	67	17
Social Funds	62	38	62	31	54	38
Social Safety Nets	40	43	40	18	48	13

## II. PROJECT DEVELOPMENT OBJECTIVES

8. One of the cornerstones of good project design is a clear statement of project development objectives (PDOs) to convey what the project intends to achieve and a set of accompanying performance indicators to monitor implementation progress.

9. **This review mapped the SP+L portfolio into 10 main areas** which have been used throughout this analysis, namely:

- **Access to public services and infrastructure**, including schools, health centers, roads and other local public goods
- **Access to economic opportunities** such as microcredit, active labor market programs, skills development
- **Temporary income support**, including unconditional transfers, public works and temporary employment, CCTs, wage subsidies
- **Human development outcomes**, including education, health and nutrition, and HIV status
- **Reaching specific vulnerable and targeted groups**
- **Community empowerment**
- **Local government capacity building**
- **National/ sectoral government capacity building**
- **Quality of services**
- **Fiscal objectives and economic efficiency**

10. **The SP+L portfolio is strongly focused on improving the performance of SP+L institutions; building national capacity is an objective in the majority of projects in the SP+L portfolio** (60 percent of projects). Social funds are an exception to this, where the focus is instead on community-level capacity building (69 percent of PDOs). Reaching vulnerable groups (55 percent) and improving the quality of service delivery (40 percent) are also frequent objectives. PDOs often encompassed more than one objective, with an average of three per project across this SP+L cohort's PDOs.

11. **Although the quality of PDOs is generally strong, they are often stated too broadly, promising high level outcomes beyond the reach of the project (especially in DPLs), with weak specification of target beneficiaries.** PDOs often focus on means rather than ends (on what is going to be done instead of what is going to be transformed) or simply restate project activities. Another weakness is the lack of specificity of the target population or of a project's intended scope/ scale. Finally, framing PDOs for institutional development issues is a particular challenge, as they should be linked to the changes in behavior or status expected from these reformed institutions but also be measurable and attributable to the project.

12. In general, PDO specification would be improved by:



a) Clear, well defined and realistic PDO statements, not overly broad or ambitious. A good practice example is:

- *“To improve the quality of life of people with disabilities and their families in areas affected by the October 8, 2005 earthquake, in Pakistan by ensuring better mobility, improved physical and mental health, increased participation in social and economic life.”*

- Pakistan Earthquake Disability Project

b) Focus on the outcome that the project can directly influence more than *how* the projects goals are to be achieved. A good practice example is:

- *“To improve Honduras’ social safety net for children and youth. This would be achieved by (i) improving nutritional and basic health status of young children through expanding the AIN-C program, and (ii) increasing the employability of disadvantaged youth through piloting a First Employment program.’*

- Honduras Nutrition and Social Protection Project

c) Adequate specification of targeted beneficiaries. A good practice example is:

- *“To increase the effectiveness of Argentina's income transfer programs for the unemployed and families with children, by improving selected design features and the transparency and accountability of the "Family allowances" and "Seguro" program, and by transferring beneficiaries from other, less effective schemes and programs, to the Family allowances and Seguro programs.”*

- Argentina Basic Protection Project

### III. KEY PERFORMANCE INDICATORS

13. **The use of human development indicators reflects a welcome focus on outcomes, but there is a ‘missing middle’ measuring service delivery that needs strengthening and is fundamentally linked to the portfolio’s primary focus on strengthening institutional capacity.** KPIs, defined during project preparation, tracked during implementation and used as a key input to end of project ratings are the main measurement instruments to monitor project results. On average, there were six outcome measures and nine intermediate outcome measures per SP+L project. The most frequent outcome measures were those related to human development outcomes, despite the fact that this was not as frequently found among PDOs. One possible explanation for this is that even though SP+L projects may focus their objectives on institutional development or service delivery, as is the case of many of the safety net and service delivery projects, in the end the concern is on making tangible improvements to people’s lives. Reaching vulnerable groups and improved quality of service delivery were also widely represented among outcome KPIs. The most frequent intermediate outcome indicators measured

institutional development, underlining the more output orientation of these short-term measures. Service delivery and quality of services were also common intermediate outcome indicators.<sup>2</sup>

14. **Overall, the quality of KPIs at project design was strong.** They are generally SMART (specific, measurable, attributable, realistic, and targeted) with two-thirds of outcome indicators clearly linked to PDOs. Stronger linkages are in the area of access to public services and human development outcomes resulting from greater utilization of basic services. KPIs measuring community empowerment, local and national government capacity building perform were less well specified.

15. However, there was a notable tendency to confuse output and outcome indicators. There was also a lack of specificity in measures (indicators like “*improved quality*” or “*adequate capacity*” are not well specified), the utilization of indicators that are not sensitive to program performance (indicators such as “percentage of beneficiaries satisfied with program implementation” or “level of satisfaction with payments”) and failure to adequately specify the target population (widespread use of generic terms like “among the poor”, or “vulnerable groups”).

16. Two-thirds of all KPIs were able to specify baseline data, but the identification of KPI data sources as well as clear responsibility for their collection was not adequately specified on average for more than half of KPIs.

17. **However, KPIs are not regularly updated during implementation, undermining their effectiveness as project management tools.** During implementation, there was a tendency to trim down the number of KPIs actually tracked, with 12 percent of outcome indicators and 19 percent of intermediate outcome indicators listed in PADs not being included in ISRs. Only 21 percent of projects consistently and regularly updated KPIs during implementation. 19 percent of projects substantially restructured their KPIs, with the highest incidence among social funds projects (54 percent) largely reflecting incorporation of standard IDA 14 indicators. Poor performance and lack of reporting of KPIs in ISRs during implementation is mainly due to: (a) inadequate selection of KPIs at project design, (b) unclear specification at project design of the source, frequency and responsibility for measurement and reporting of KPIs, (c) data collection delays, for example, delays in establishing MIS and procurement delays in contracting external consultancies to carry out surveys as well as delays in indicators derived from national and sectoral sources external to the project.

18. **Institutional development indicators pose the greatest challenge in designing and reporting KPIs.** The SP+L portfolio’s focus on institutional objectives calls for the development of good measures of performance in this area. The performance ratings for institutional development indicators were consistently lower, including SMART ratings, linkages with PDOs, existence of a baseline, and adequate specification of data sources. The main weaknesses are related to specificity and measurability. Activity statements, like “implement management information systems (MIS)”, are often used as indicators instead of measurable variables. A stronger example would be percent of beneficiaries registered in the MIS or percentage of local offices producing monthly reports using automated MIS data.

19. Recommendations to improve KPIs include:

At project design:

---

<sup>2</sup> The Technical Guidance Notes for Safety Nets, Labor Market Programs, Service Delivery and Social Funds developed as part of this “Results Readiness” work include a list of indicators and typical data sources used in SP+L projects.

- a) KPIs should be identified along the stages of the results chain: Output measures help track project implementation, explain outcomes and should be explicitly included and measured. Using the term “intermediate outcome” can be confusing - There needs to be a differentiation between outputs and short-term service delivery outcomes. There is a need for clarity and measurability around a “missing middle” in service delivery;
- b) KPIs should include an appropriate balance of both output and outcome measures: *Output* indicators should rely on accessible, high frequency data (usually administrative data). *Outcome* indicators usually rely on less frequently available data such as population data, impact evaluation surveys, citizen scorecards, tracer studies or beneficiary assessments;
- c) Before choosing indicators, it is good practice to perform an analysis of what data are available and what data can be collected given resources and feasibility constraints;
- d) Need to determine at which level (individual, household, community, facility) indicators should be measured. Consider disaggregation by gender, income deciles, age, and geographical area;
- e) Set baselines and target values at the design stage;
- f) The quality of institutional development indicators needs to be strengthened;
- g) Indicators used in impact evaluations require a framework that allows for determining net program effect (a counterfactual)

<b>Example of Weaker and Stronger KPIs</b>	
<b>Weaker</b>	<b>Stronger</b>
<ul style="list-style-type: none"> <li>• Increase in access to health centers</li> <li>• Increase in access to education</li> <li>• Contribution to short-term employment and income generation</li> <li>• Program resource allocation favors poorer participating communes and villages</li> <li>• Ensuring that communities have benefited from social mobilization and facilitation using highly participatory methodologies and that community based organizations are well represented</li> </ul>	<ul style="list-style-type: none"> <li>• Increased medical consultations in FSRDC built or rehabilitated health centers</li> <li>• Increased enrollment rate in basic education by 7 percent for boys and girls in the areas the SFD intervened centers</li> <li>• 40-50 percent of microfinance savers/borrowers in poor participating communities confirming an improvement in their household living standard</li> <li>• Distribution of financing of subprojects consistent with regional targeting criteria as estimated by the QUIBB 2006 Survey</li> <li>• Over 50 percent of participating households record increased levels of trust and cooperation among stakeholders at the community level</li> </ul>

The use of KPIs during implementation could be strengthened by:

- a) Selectivity in identifying the mix of KPIs during project design so that they can be tracked;
- b) Ensuring that KPIs in PADs are listed in the ISRs;
- c) Identifying the frequencies at which KPIs are measured as part of the ISR in order to update indicators accordingly;

- d) PDO ratings in ISRs should be well-substantiated by KPI reporting. In cases where KPIs were not consistently reported it is difficult to justify PDO ratings;
- e) Simplifying data collection processes, notably by being clear on data requirements, setting realistic time frames, and specifying responsibilities for data collection, reporting and review.

#### IV. MONITORING AND EVALUATION ARRANGEMENTS

20. **M&E arrangements are made up of M&E instruments and institutional arrangements. Both are generally well specified though there is room for improvement in: knowledge sharing across types of projects; strengthening client capacity for implementing M&E; and developing suites of instruments and arrangements tailored to measuring and supporting institutional capacity development and service delivery.**

21. The instruments or component parts of an M&E framework should match both the project and country context. M&E components in the SP+L projects reviewed consisted of MIS (in 71 percent of projects); impact evaluations (51 percent), though the robustness is not always evident; qualitative approaches through beneficiary assessments (42 percent) and participatory M&E (29 percent); operational assessment through process evaluation (36 percent); spot checks (22 percent) or technical/operational audits (23 percent). There is an interesting assortment of “other” M&E arrangements (39 percent) include facilities surveys, sectoral information systems, national household surveys, audits that are worth exploring more systematically, especially as many of these seem particularly well suited to supporting institutional capacity development.

**Table 3. Use of M&E Components by Thematic Area (%)**

	MIS	Process Evaluation	Beneficiary Assessment	Spot Check	Impact Evaluation	Technical/Operational Audits	Participatory M&E	Expenditure Tracking	Other
<b>Total</b>	<b>71</b>	<b>36</b>	<b>42</b>	<b>22</b>	<b>52</b>	<b>23</b>	<b>29</b>	<b>9</b>	<b>39</b>
Labor Markets	56	0	44	33	78	0	11	11	33
Service Delivery	42	25	8	8	33	0	8	8	67
Pensions	67	0	0	0	0	0	0	0	0
Social Funds	100	77	77	8	54	46	54	15	23
Social Safety Nets	75	38	43	30	55	30	33	8	40

22. Interestingly, standard “packages” of M&E tools are present in particular types of operations, with variations across project types reflecting their specific operational needs as well as information sharing within particular regions and project communities. For example, most social funds have an MIS, process evaluation, and beneficiary assessments. Labor market operations focus mainly on MIS/beneficiary registries and impact evaluations. There are some regional patterns – Africa region (AFR) and Latin America and Caribbean region (LCR) use the largest variety of M&E components. There is

undoubtedly room for improved knowledge sharing across regions and types of projects as well as the need to develop data collection instruments and institutional arrangements that are well tailored to measuring and supporting institutional capacity development and results-based management.

23. Only 24 percent of projects report specifically on M&E costs. These ranged from \$400,000 to \$10 million with a median value of \$2.4 million.

24. **Impact evaluation is a specific and important component of many M&E frameworks, with fifty one percent of projects planning to apply an impact evaluation.** However, only 27 percent of projects provided sufficient detailed methodological information to consider them rigorous impact evaluations (e.g. those that specify a counterfactual). More than a third of projects in the Latin America and Caribbean region (LCR) and South Asia region (SAR) intend to measure impact. Most of these projects also include early plans for baseline and follow up data collection. Among rigorous evaluations that specify the use of counterfactuals, randomization is the most frequent methodology used to generate a counterfactual, mainly used in CCT safety net projects. Only 20 percent of the rigorous impact evaluations reported costs, with an average of US\$724,000 per evaluation.

25. Despite relatively good performance in M&E arrangements, areas for improvement include:

- a) Only 34 percent of PADs had a strong assessment of client M&E capacity;
- b) Plans to strengthen client capacity are missing or need improvement in almost half of projects;
- c) Less than 40 percent of projects had a clear and comprehensive definition of M&E roles and responsibilities;
- d) Integration with sectoral or national M&E systems is more common among service delivery and labor markets projects.

26. **M&E systems were considerably weaker at operationalization** (45 percent strong design versus 27 percent strong implementation). KPI updating, as discussed earlier, was not systematic. There were delays in data collection plans (including both KPI monitoring and evaluation): only 20 percent projects under implementation have carried out data collection as intended. This problem was particularly pronounced in safety net and service delivery projects, most likely because of strong focus on human development outcomes which are often longer-term and less frequently measured. There was also a frequent disconnect between progress in implementing M&E activities and task teams' reported ISR M&E rating. There was overall lack of preparation and overly-ambitious expectations about client capacity and time-frame needed to carry out M&E. Readiness on the ground was limited. Unless they were repeater operations, few projects were ready to hit the ground running. Contracting of M&E staff took a good part of the first year of implementation, as did drafting M&E manuals. Typical problems were delays in establishing MISs capable of reporting on KPIs as well as delays in contracting the various external studies including developing terms of reference (TOR) and procurement process.

27. There were several good practice examples of M&E arrangements identified:<sup>3</sup>

- The Jamaica Conditional Cash Transfer Project provides a comprehensive and well described system and appropriate mix of M&E approaches, including management information systems for routine monitoring, a process evaluation for informing project design, internal audits and

---

<sup>3</sup> Each of these good practice examples is described in more detail in the corresponding Technical Guidance Note

spot checks to validate implementation, piggy-backing on an annual, national household survey to review targeting and coverage, and an impact evaluation to assess outcomes.

- The Malawi 3<sup>rd</sup> Social Action Fund Project is an example of a very effective approach to reinforcing a national system from a project platform as well as appropriate use of multifaceted evaluation instruments. The M&E system uses: (a) well defined Results Framework with clearly defined goals, objectives, outputs, and activities with corresponding indicators, means of verification and key assumptions; (b) a strategy for project processes, information requirements, tools and methodologies for data collection, analysis and reporting; (c) a plan with clear roles and responsibilities for indicators tracking with respect to data gathering and reporting; (d) a computerized Local Authority Management Information System (LAMIS) to cater to local level information needs; (e) internal and external periodic assessment and evaluations in the form of annual tracking studies, baseline studies, engendered community score cards, mid-term evaluations, ex-post evaluations, and impact evaluations; and (f) participatory community monitoring and accountability approaches and systems using Citizen Report Cards and Community Score Cards.
- The Argentina Head of Households Transition Project's M&E plan builds on the experience of the Government of Argentina and the World Bank during the implementation of previous programs. Important lessons are drawn from the experience of *Jefes*, notably the importance of transparency and accountability in the administration of benefits. Specifically, social accountability was enhanced through: (i) regular monthly cross-checks control of the registry database with other databases by an independent agency; (ii) greater use of public information and participatory mechanisms through workshops in order to inform and promote participation of local organizations); (iii) more supervision missions and media monitoring.

28. Key recommendations for improving M & E arrangements are:

- a) Project design should involve a thorough consideration of technical, institutional and political aspects of M&E and the costs, actions and technical assistance needed to implement the proposed M&E system. This should be summarized in an M&E Manual drafted as part of project preparation;
- b) Set a realistic timeframe for M&E activities, including contracting processes;
- c) Simplify data collection processes;
- d) Be realistic about links with national and sectoral systems, especially at initial stages and in certain country contexts and type of operations;
- e) The mix of evaluation components reflect project information needs at a specific point in time. Simplify M&E start up and consider system evolution as operation consolidates and new operational or impact issues arise;
- f) Use impact evaluations selectively, most usefully when a project is innovative, replicable/ scalable, addresses a knowledge gap, and is likely to have substantial policy impact. Impact evaluations must be closely coordinated with project implementation. Operational program rules will frame how to identify the comparison group needed for the counterfactual;
- g) During implementation, project supervision needs to pay close attention to how project implementation affects planned M&E activities, in particular: the timing of baselines, delays in

project activities that affect reporting on results, contracting of external evaluations and surveys, and whether KPIs are being generated as planned through M&E instruments.

## V. USE OF RESULTS WITHIN THE WORLD BANK

29. **The use of results has two main purposes: to “prove” (accountability) and to “improve” (management).** Within the World Bank, project level results are reported through two standard instruments: (a) ISRs, the main documents that track implementation advance and progress towards PDO achievement; and (b) ICRs which provide a final evaluation of achievement of PDOs at project closing.

30. **Results monitoring is *not* being used effectively in the World Bank to track project implementation.** Only one quarter of projects showed strong performance in using KPIs to measure progress and inform decisions. First, there is a disconnect between KPI findings and PDO ratings as PDO ratings in ISRs are not well-substantiated by KPI reporting. Second, whereas the M&E rating in the ISR provides the opportunity to flag issues, it is rarely used. Even when ratings or indicators change, justifications for the change are rarely discussed. Finally, the section on “Quality of Information on Outcomes” is not commonly used. Overall, the ISR has limitations as a monitoring tool by virtue of both design (for example, the format doesn’t lend itself to identifying the expected frequency of KPI reporting) and use.

31. Forty one percent of task teams consistently discussed M&E issues under the “Key Issues and Actions for Management” section in the ISR. Social fund and labor market task teams were more likely to identify and discuss issues related to results and M&E as part of project supervision than social service delivery and safety net projects. There may be a virtuous circle to this - projects with more attention to M&E issues by task teams tended to have higher ratings on M&E. Bank management comments on results and M&E is strong in 36 percent of projects, particularly when M&E issues are first flagged by task teams.

32. However, the review identified several cases of good practice<sup>4</sup>:

- The Ethiopia Productive Safety Net I Project stands out for the good linkages between PDO ratings and results reporting as well as for innovative results monitoring tools, results-oriented supervision by the Bank task team, and management engagement on results.
- The Benin National Community Driven Development (CDD) Project displayed an effective use of M&E ratings to flag issues. This shows a proactive approach by the task team, with M&E issues often picked up in Priority Action to alert management to evolving issues.
- The West Bank and Gaza Emergency Services Support Project III taps into lessons from previous phases the better use results to monitor progress toward project goals even within a very challenging country context.

33. Recommendations to improve the use of results by Bank management and task teams include:

- a) Ensure that PDO ratings are consistent with KPIs. PDO ratings justification should directly speak to KPI performance;

---

<sup>4</sup> Each of these good practices are described in detail in the Technical Guidance Notes

- b) Changes in KPIS should be formalized. There needs to be a process that is less onerous than a formal project restructuring that codifies changes in KPIS;
- c) Task teams should use the M&E rating proactively. Task teams can better track and highlight implementation of the M&E framework by adjusting ratings on M&E more responsively. Task teams should not face disincentives for being realistic about ratings and reporting problems;
- d) ISRs should eliminate the box to check for quality of information on final outcomes as it is not used. Rather, the KPI reporting table should require explanation for any KPI not updated;
- e) All outcome indicators should be available by the time of the ICR. The ISR can be better used to identify where KPIs will not be able to be gathered and provide for more formal restructuring. Which KPIs will be reported when should be made clear at the outset of the project.

## VI. KEY THEMES AND RECOMMENDATIONS BY LENDING INSTRUMENT

34. This “Results Readiness” assessment paid particular attention to the use of policy-based loans<sup>5</sup> and emergency recovery loans whose use is growing in response to the global economic crisis and rise in disasters. These approaches call for a treatment of M&E tailored to the nature of the support being provided. The introduction of results-based investment lending (RBIL) will also require a tailored and considerably strengthened approach to M&E.

35. **Development Policy Loans are considerably weaker in M&E than sector investment loans.** In terms of PDOs, the review found that PDOs are in general fully aligned with CAS results: 95 percent of SP+L DPLs’ PDOs focus on strengthening national institutional capacity building, a challenging area for M&E, but one that would benefit substantially from an improved approach to M&E. The next most frequent focus areas are improving quality of service delivery and improving human development outcomes. In general, DPLs’ PDOs are very broad, with outcomes not well specified and too long. There are recognized challenges with attribution: outcomes are short-to-medium term making it difficult to judge the extent to which the operation can be reasonably held accountable for any longer-term outcomes. However, this is not a good reason for not focusing on clear results chains and on establishing the logical chain of results that can be expected from a DPL and the contributions that these will have to longer-term, higher order sectoral, country-based outcomes.

36. Outcome KPIs were more frequent (on average seven indicators per project) than intermediate outcome indicators (on average 2.4 indicators per project) which are not explicitly required. Outcome indicators mainly focused on measuring human development outcomes whereas 44 percent of intermediate outcome indicators focus on measuring efficiency and fiscal objectives. Usually outcome indicators measure medium-term (about 3 years) objectives, while intermediate indicators measure short term activities or *actions* and should be collected on an annual basis. The same lesson drawn from the review of investment operations applies to DPLs: to be able to measure medium term goals, it is crucial to have a focus on the “middle part” of the results chain and on what can be measured in the short time line. Especially in the case of programmatic DPLs, intermediate indicators are critical to monitor program performance to achieve final objectives and need to be developed and revised each phase. KPIs performance in DPLs is generally worse than in investment operations in terms of SMART ratings and other aspects. The main weaknesses are related to the lack of specification of data

---

<sup>5</sup> Almost all policy based lending operations in our sample are *programmatic* DPLs (17 out of 21).



sources as well as data collection responsibilities and frequency which are vague, sometimes absent or not specified. KPIs are often found to be not attributable and the overall linkage to PDOs is weak. A recurrent constraint in developing a good result framework is the weak country statistical capacity which limits the set of available indicators<sup>6</sup>.

**Table 4: Project Development Objectives by lending instrument (% of projects)**

	Access to public services and infrastructure	Access to economic opportunities	Temporary income support	Human development outcomes	Reaching specific vulnerable groups	Community empowerment and capacity building	Local government institutional capacity building	National/sectoral government institutional capacity building	Quality of services/service delivery	Fiscal objectives/efficiency
DPLs	19	29	5	57	33	10	5	95	62	48
SILs	37	35	23	26	58	16	26	56	33	7
ERL	54	38	0	8	77	31	15	15	31	8
<b>Total</b>	<b>35</b>	<b>34</b>	<b>14</b>	<b>31</b>	<b>55</b>	<b>17</b>	<b>18</b>	<b>60</b>	<b>40</b>	<b>18</b>

37. **DPLs are potentially important vehicles for strengthening sectoral and national M&E capacity but this opportunity is under-utilized.** In terms of overall M&E arrangements, DPLs are expected to assess country implementation capacity as well M&E system quality. If client capacity is found to be weak, the DPL should identify what needs be done to strengthen national M&E systems. However, DPLs perform *less well* than investment operations in assessing M&E capacity and in developing plans to strengthen national M&E system. Often the implementation of M&E activities is supported by parallel technical assistance (TA). The most common M&E instrument is the use of national and sectoral surveys and other databases. One third of the projects plan impact evaluations though the robustness of the approaches which are often unclear. Several support a range of qualitative assessments (beneficiary assessments, participatory monitoring instruments M&E workshops) though less frequently than in investment operations.

38. In terms of the use of results, ISRs are not designed for DPLs or commonly used for DPL supervision. Instead, in programmatic DPLs –which constitute the bulk of SP+L DPLs-- the supervision phase coincides with the preparation phase for the new operation, with preparation substituting for the use of ISRs as a management and supervision tool.

39. Key issues and recommendations regarding DPLs are:

- a) More effective approaches to the use of the “Key Results” section in the policy matrix, ensuring that these results are measurable within the annual timeframes of DPL operations, including a greater attention to intermediate outcome indicators based on a clear results chain;
- b) More structure and guidance for the KPIs specified in the policy matrix. M&E quality would benefit from stricter requirements in terms of specifying data sources, data collection responsibilities, data collection frequencies, and M&E arrangements;

<sup>6</sup> The following are among the main challenges identified in the 2009 Development Policy Lending Retrospective: Flexibility, Customization and Results, OPCS October 2009: 1) country weak statistical capacity; 2) indicators to measure actions; 3) alignment between CAS, PSR and DPL objectives.

- c) Attribution and causality is a challenge because of the broad-based nature of policy reform and the national scale of objectives. Nonetheless, documentation should include a detailed results chain between actions supported by DPL operations and expected results;
- d) DPLs are an opportunity to strengthen national/ sectoral M&E systems: including capacity enhancement measures (TA or through other donors) is a good practice when strengthening M&E is a DPL objective;
- e) Improve result-oriented Bank supervision by creating an ISR format that better fits DPLs.

40. **Emergency Recovery Loans (ERLs) are designed in conditions have an important bearing on their results readiness and the use of M&E in these operations has been crafted appropriately: they represent good practice approaches to M&E even under difficult circumstances.** In emergency situations, national and local level institutions are usually weakened, including in their M&E systems. Household level shocks may have drastically changed wellbeing and access to services quickly, making establishment of baseline and targets difficult. The quick preparation time required to bring emergency responses into operation works against the adequate preparation of traditional M&E frameworks and longer-term institutional capacity building is not often an appropriate focus.

41. Despite these limitations, ERLs were found to have appropriate PDOs – 77 percent include a focus on reaching vulnerable populations, largely those affected by political and natural shocks. The next most frequent focus area was access to public services (54 percent), again closely related to restoring essential services in the wake of crises. KPIs were appropriately streamlined for ERLs, with an average of about half the outcome indicators than for DPLs and regular investment operations. KPIs (both outcome and intermediate indicators) were most frequently focused on measuring improved access to basic services.

42. Design of M&E arrangements was a challenge in ERLs, with lower quality ratings overall than for other kinds of projects, reflecting the rapid, crisis response focus of ERLs. The most frequent components were MIS systems, technical audits and beneficiary assessments. Appropriately, there were no instances of robust impact evaluations planned or carried out for ERLs. Selectivity has paid off in ERLs and the implementation of the more limited M&E frameworks showed performance similar to other investment operations. Despite some of the limitations in the M&E framework of ERLs, Bank task teams and management showed a use of results comparable to other investment operations.

43. The good practices identified for ERLs include:

- a) Well specified PDOs which capture the context of the emergency;
- b) Simple KPIs that can be tracked in challenging emergency environments. A focus on outputs and the restoration of services is probably more appropriate than any longer term impact on institutional capacity or household level outcomes;
- c) The flexible design of M&E approaches given the likely need to adjust to unanticipated needs due to the fragile context and rapid planning;
- d) M&E capacity (both in-house or externally contracted) may be very weak in these crisis contexts. Twinning with local non-governmental organizations (NGOs) or research teams and engaging communities in participatory monitoring where possible available is a good approach to strengthening weak systems and supporting accountability.

44. **The introduction of results-based investment lending (RBIL) presents an opportunity to substantially advance the results agenda within the Bank and client countries by moving from a project focus to a program focus, but the challenges of using RBIL effectively should not be underestimated.** The greatest opportunity presented by RBIL is the ability to shift project foci from inputs to outputs as well as to use projects as vehicles to strengthen country systems for results based management, procurement and financial management. However, we should not underestimate the change in culture and practice that will be needed within the Bank to make RBIL work. This “Results Readiness” work points to several of these areas that will be particularly salient for RBIL: (i) the shortcomings of ISRs in tracking key performance indicators and their limited use as a management tool; (ii) the challenge of identifying and tracking indicators that are meaningful for development outcomes, but for which the project can be held accountable for disbursement purposes under RBIL; and (iii) the difficulties in quantifying advances in institutional capacity building and improved service delivery which are likely to be at the heart of RBIL programs.

## VII. CONCLUSIONS AND RECOMMENDATIONS

45. This review concludes with ten main points:

**#1 -- The SP+L portfolio’s main focus is on strengthening client’s institutional capacity in the design and implementation of SP+L programs.** The three most common objectives in the current the SP+L portfolio are: (i) strengthening national government’s institutional capacity (60 percent of all projects list this as part of the PDOs), (ii) reaching specific vulnerable groups (55 percent of projects), and (iii) improving service delivery (40 percent of projects). It is challenging to craft PDOs that avoid both unrealistic higher level objectives and a narrow focus on means rather than ends, particularly given the current portfolio objectives and to match them with appropriate key performance indicators.

**# 2 -- There is a need to further develop indicators and approaches to measuring the “missing middle” of service delivery.** To better align measures of project performance with project development objectives, a series of project implementation and service delivery indicators, and related measurements tools, need to be further developed. These indicators should follow as a product of a well elaborated results chain for each project, from inputs to activities to outputs to outcomes, with a clear discussion of assumptions and the logic behind the sequence of results. A focus on thinking through, then tracking, indicators of service delivery, project implementation and capacity development would strengthen effective management at the client and Bank levels. With a clear logic chain established these indicators could be tracked to chart the project’s trajectory toward meeting outcomes and longer-term country based sectoral goals.

At present, there is a strong and welcome focus on measuring outcomes, often in human development, linked to the growing use of impact evaluations. However, relying on impact evaluations as the principal M&E approach next to MIS - as is the case in the current portfolio - reveals a degree of mis-alignment since process evaluations, spot checks, sectoral service delivery surveys and other types of M&E approaches are probably better suited to answering medium-term questions concerning institutional capacity development and service delivery. The measures for tracking the “missing middle” should be

developed across the portfolio as a complement to impact evaluations. These indicators are more easily measured, directly attributable to the project, and critically linkable in the results chain leading to improved welfare outcomes.<sup>7</sup> These types of indicators will be likely candidates to be used as measures against which to disburse funds in the new results based lending instrument under preparation.

**#3 -- Develop a better approach to measuring institutional capacity, particularly around the primary functions of SP+L agencies.** Given the portfolio's focus on institutional capacity development, measures of the performance of SP+L sector should be improved. Areas requiring clear measures include: identifying and enrolling beneficiaries, targeting, payment systems, fraud and error control, performance monitoring of service delivery providers, responsiveness to citizens, transparency, efficiency, MISs, performance M&E.

**# 4 -- Clarity is needed on the purpose of the standardized "core indicators" for SP+L being called for within the World Bank.** First, clarity is needed on the rationale for tracking these indicators. Different goals call for different instruments and measures: benchmarking, management, accountability, knowledge generation. Each of these goals would lead in a particular direction, so to avoid confusion the purpose of these indicators needs to be clearly established. Indicators should be selective, and if related to monitoring project implementation within the portfolio, sensitive to project type and stage of implementation to guide performance tracking. If accountability and portfolio performance are goals, then the measures should allow for aggregation across the SP+L portfolio.

The International Development Association (IDA) 15 and Africa results team efforts are a good start in performance measurement, but these measures are less applicable to more mature programs and systems found more commonly in middle income countries.

Implementation of this "core system" will be a key challenge, particularly given current weaknesses with ISRs. Feedback from the implementation of IDA indicators, the Africa efforts and the experience to date in other sectors should be used as guidance.

**# 5 -- Projects should be explicit regarding data sources, frequency and responsibility of reporting, and expected use of the indicators by stakeholders.** Good practice M&E in PADs includes a clear description of data sources, frequency and responsibility of reporting, and use of results for management and feedback. This kind of specificity facilitates the use of ISRs to track implementation and allows for closer linkages between KPI reporting, M&E ratings and PDO ratings.

During implementation, things rarely go as planned. Project supervision needs to pay close attention to how project implementation affects planned M&E activities. Conversely, good indicators of project implementation can be used to inform effective supervision.

**# 6 -- Though generally well formulated, KPIs are weaker in implementation.** There are problems with consistency across indicators from PAD to ISRs, limited use of baselines and infrequent updating of KPIs. Only 65 percent of outcome indicators and 64 percent of intermediate indicators have baselines

---

<sup>7</sup> Please see the "Technical Guidance Notes for Task Teams" developed as part of this Results Readiness exercise for lists of indicators and related sources of data, including service delivery and institutional development indicators.

identified; 12 percent of outcome indicators (and 19 percent of intermediate indicators) listed in the PAD are not listed in the ISR; and only 21 percent of projects have regularly updated KPIs. This hinders the ability of both Bank and client teams to use these data effectively as a management tool. Without active attention to tracking progress during implementation as part of regular supervision, with a reference to baselines and active use of ISRs, it is hard to conduct effective monitoring and management.

**# 7 -- Place more emphasis on supporting clients' M&E capacity. At project design, an M&E manual should be developed** encompassing the M&E framework, an identification of institutional arrangements, costing, and TA needed to implement the M&E system. Successful operationalization of an M&E system depends not only on well chosen KPIs and assessment tools, but also on developing the right incentives and capacity to manage their design, implementation and use. Ideally, project M&E systems should be integrated with sectoral and national systems, but this must be balanced with having simple, reliable performance data.

Good practice PADs include an explicit description of data sources, frequency and responsibility of reporting, and use of results for management and feedback. This kind of specificity facilitates the use of ISRs to track implementation and allows for closer linkages between KPI reporting, M&E ratings, and PDO ratings.

However, things rarely go as planned. Project supervision needs to pay close attention to how project implementation affects planned M&E activities. Conversely, good indicators of project implementation can be used to inform effective supervision. Delays in M&E readiness are often due to problems hiring staff, building MISs, and contracting studies and surveys.

**# 8 -- Impact evaluations should be used selectively and strategically.** There has been a notable rise in the use of impact evaluations: just over 50 percent of SP+L projects have impact evaluations planned, with 33 percent of projects having a rigorous impact evaluation planned with baseline and follow-up and the use of a counterfactual. The most common methodological approach is a randomized control design, followed by the use of matching and regression discontinuity design. There is also a thoughtful reference to past evidence on results in just over 40 percent of projects.

Impact evaluations which use counterfactuals to establish the causal impact of a program should be used strategically when the program is innovative, replicable and/ or scalable to reach a broader set of beneficiaries, addresses a knowledge gap and is likely to have a substantial policy impact. Impact evaluations can also often be used to explore implementation options among program treatment alternatives.

More broadly, the choice and combination of monitoring and evaluation approaches should reflect the policy question or management need at hand and task teams should not be pressured to necessarily include or exclude an approach based on the merits of the methodology. Future reviews of impact evaluations should assess whether these planned impact evaluations are carried out and summarize related findings.

**# 9 -- Use project ISRs more effectively.** Task teams need better incentives for using ISRs to identify problems and regularly update KPIs. At present, ISRs use as an effective management tool is compromised by limited reporting on indicators and lack of consistency between KPI and PDO ratings. The ISR format could be improved to better track results, including expected frequencies of KPIs.

**# 10 -- Diverse lending instruments need a tailored, strengthened approach to M&E. Project-based approaches to M&E have been tailored to traditional investment lending.** Development Policy Loan (DPLs) and Emergency Recovery Loans (ERLs) today --and Results-Based Investment Lending (RBIL) in the future-- need M&E approaches that are fully reflective of the different challenges and opportunities presented by these instruments. OPCS needs to partner with regional operational advisers and others to strengthen their support to teams to meet this challenge, developing clear guidelines, providing hands-on support, and identifying good practice examples.

Development policy loans (DPLs) present opportunities to strengthen national and sectoral M&E systems, especially programmatic DPLs, but these opportunities are not being fully exploited. DPLs are weaker in M&E than sector investment loans. Guidelines for DPLs require a clear assessment of M&E system quality and plans for strengthening weak systems, but actual practice is limited. Despite challenges in establishing causal links between DPL supported reforms and outcomes, DPLs should still have clear results chains and KPIs for both annual and medium-term performance. This review echoes the finding of the 2006 DPL retrospective, namely the need to strengthen results frameworks in DPLs by systematically including benchmark values, keeping program matrixes and M&E approaches simple and focusing more on M&E capacity building.

To support improved M&E in DPLs, DPLs would benefit from a better and standardized specification of M&E arrangements (e.g. an equivalent to Annex III for investment loan PADs), including data sources, responsibilities for data collection, frequency of KPI reporting and use of data by stakeholders. The ISR format also needs adaptation for DPLs.

Emergency recovery loans (ERLs) also require special treatment for M&E, particularly given the weakened state of government programs in the face of an emergency, including M&E capacity, and the need for a rapid response. Under these circumstances, PDOs should seek to capture the context of emergency and KPIs should keep it as simple as possible in order to make sure that they can be tracked in these more challenging environments. Often these will be focused on outputs and the restoration of services. Design of M&E systems should be flexible and may need to adjust to unanticipated needs due to the fragile context and rapid planning. Twinning with local NGOs or research teams and engaging communities in participatory monitoring were possible can strengthen M&E approaches.

The introduction of results-based investment loans (RBIL) present an opportunity to move from a project focus to a program focus and to shift much of task team's work from tracking inputs to measuring outputs and outcomes. This will necessitate a central focus on the identification, monitoring and evaluation of indicators that will be used for disbursement which can be measured accurately and for which the project can be held accountable. This will place monitoring and evaluation at the center of the project and these efforts will likely center on assessing program performance in service delivery.

However, this review has identified substantial existing challenges both with measuring the ‘missing middle’ of service delivery and with tracking project performance during implementation. Efforts will be needed to develop the next generation of instruments to accurately measure these types of outputs and shorter-term outcomes, including service delivery surveys of providers, audits of administrative data, sectoral public expenditure reviews, etc. and these tools will need to be coupled with changes in the practice and incentives around performance monitoring during project implementation.

## **Social Protection Discussion Paper Series Titles**

<b><u>No.</u></b>	<b><u>Title</u></b>
1107	Results Readiness in Social Protection & Labor Operations by Laura Rawlings, Maddalena Honorati, Gloria Rubio and Julie Van Domelen, February 2011
1106	Results Readiness in Social Protection & Labor Operations: Technical Guidance Notes for Social Service Delivery Projects by Julie Van Domelen, February 2011
1105	Results Readiness in Social Protection & Labor Operations: Technical Guidance Notes for Social Safety Nets Task Teams by Gloria Rubio, February 2011
1104	Results Readiness in Social Protection & Labor Operations: Technical Guidance Notes for Social Funds Task Teams by Julie Van Domelen, February 2011
1103	Results Readiness in Social Protection & Labor Operations: Technical Guidance Notes for Labor Markets Task Teams by Maddalena Honorati, February 2011
1102	Natural Disasters: What is the Role for Social Safety Nets? by Larissa Pelham, Edward Clay and Tim Braunholz, February 2011
1101	North-South Knowledge Sharing on Incentive-based Conditional Cash Transfer Programs by Lawrence Aber and Laura B. Rawlings, January 2011
1008	Social Policy, Perceptions and the Press: An Analysis of the Media's Treatment of Conditional Cash Transfers in Brazil by Kathy Lindert and Vanina Vincensini, December 2010 (online only)
1007	Bringing Financial Literacy and Education to Low and Middle Income Countries: The Need to Review, Adjust, and Extend Current Wisdom by Robert Holzmann, July 2010 (online only)
1006	Key Characteristics of Employment Regulation in the Middle East and North Africa by Diego F. Angel-Urdinola and Arvo Kuddo with support from Kimie Tanabe and May Wazzan, July 2010 (online only)



- 1005 Non-Public Provision of Active Labor Market Programs in Arab-Mediterranean Countries: An Inventory of Youth Programs  
by Diego F. Angel-Urdinola, Amina Semlali and Stefanie Brodmann, July 2010 (online only)
- 1004 The Investment in Job Training: Why Are SMEs Lagging So Much Behind?  
by Rita K. Almeida and Reyes Aterido, May 2010 (online only)
- 1003 Disability and International Cooperation and Development: A Review of Policies and Practices  
by Janet Lord, Aleksandra Posarac, Marco Nicoli, Karen Peffley, Charlotte McClain-Nhlapo and Mary Keogh, May 2010
- 1002 Toolkit on Tackling Error, Fraud and Corruption in Social Protection Programs  
by Christian van Stolk and Emil D. Tesliuc, March 2010 (online only)
- 1001 Labor Market Policy Research for Developing Countries: Recent Examples from the Literature - What do We Know and What should We Know?  
by Maria Laura Sanchez Puerta, January 2010 (online only)
- 0931 The Korean Case Study: Past Experience and New Trends in Training Policies  
by Young-Sun Ra and Kyung Woo Shim, December 2009 (online only)
- 0930 Migration Pressures and Immigration Policies: New Evidence on the Selection of Migrants  
by Johanna Avato, December 2009 (online only)
- 0929 Ex-Ante Methods to Assess the Impact of Social Insurance Policies on Labor Supply with an Application to Brazil  
by David A. Robalino, Eduardo Zylberstajn, Helio Zylberstajn and Luis Eduardo Afonso, December 2009 (online only)
- 0928 Rethinking Survivor Benefits  
by Estelle James, December 2009 (online only)
- 0927 How Much Do Latin American Pension Programs Promise to Pay Back?  
by Alvaro Forteza and Guzmán Ourens, December 2009 (online only)
- 0926 Work Histories and Pension Entitlements in Argentina, Chile and Uruguay  
by Alvaro Forteza, Ignacio Apella, Eduardo Fajnzylber, Carlos Grushka, Ianina Rossi and Graciela Sanroman, December 2009 (online only)
- 0925 Indexing Pensions  
by John Piggott and Renuka Sane, December 2009 (online only)

- 0924      Towards Comprehensive Training  
by Jean Fares and Olga Susana Puerto, November 2009
- 0923      Pre-Employment Skills Development Strategies in the OECD  
by Yoo Jeung Joy Nam, November 2009
- 0922      A Review of National Training Funds  
by Richard Johanson, November 2009
- 0921      Pre-Employment Vocational Education and Training in Korea  
by ChangKyun Chae and Jaeho Chung, November 2009
- 0920      Labor Laws in Eastern European and Central Asian Countries: Minimum Norms  
and Practices  
by Arvo Kuddo, November 2009 (online only)
- 0919      Openness and Technological Innovation in East Asia: Have They Increased the  
Demand for Skills?  
by Rita K. Almeida, October 2009 (online only)
- 0918      Employment Services and Active Labor Market Programs in Eastern European  
and Central Asian Countries  
by Arvo Kuddo, October 2009 (online only)
- 0917      Productivity Increases in SMEs: With Special Emphasis on In-Service Training  
of Workers in Korea  
by Kye Woo Lee, October 2009 (online only)
- 0916      Firing Cost and Firm Size: A Study of Sri Lanka's Severance Pay System  
by Babatunde Abidoeye, Peter F. Orazem and Milan Vodopivec, September 2009  
(online only)
- 0915      Personal Opinions about the Social Security System and Informal Employment:  
Evidence from Bulgaria  
by Valeria Perotti and Maria Laura Sánchez Puerta, September 2009
- 0914      Building a Targeting System for Bangladesh based on Proxy Means Testing  
by Iffath A. Sharif, August 2009 (online only)
- 0913      Savings for Unemployment in Good or Bad Times: Options for Developing  
Countries  
by David Robalino, Milan Vodopivec and András Bodor, August 2009 (online  
only)

- 0912 Social Protection for Migrants from the Pacific Islands in Australia and New Zealand  
by Geoff Woolford, May 2009 (online only)
- 0911 Human Trafficking, Modern Day Slavery, and Economic Exploitation  
by Johannes Koettl, May 2009
- 0910 Unemployment Insurance Savings Accounts in Latin America: Overview and Assessment  
by Ana M. Ferrer and W. Craig Riddell, June 2009 (online only)
- 0909 Definitions, Good Practices, and Global Estimates on the Status of Social Protection for International Migrants  
by Johanna Avato, Johannes Koettl, and Rachel Sabates-Wheeler, May 2009 (online only)
- 0908 Regional Overview of Social Protection for Non-Citizens in the Southern African Development Community (SADC)  
by Marius Olivier, May 2009 (online only)
- 0907 Introducing Unemployment Insurance to Developing Countries  
by Milan Vodopivec, May 2009 (online only)
- 0906 Social Protection for Refugees and Asylum Seekers in the Southern Africa Development Community (SADC)  
by Mpho Makhema, April 2009 (online only)
- 0905 How to Make Public Works Work: A Review of the Experiences  
by Carlo del Ninno, Kalanidhi Subbarao and Annamaria Milazzo, May 2009 (online only)
- 0904 Slavery and Human Trafficking: International Law and the Role of the World Bank  
by María Fernanda Perez Solla, April 2009 (online only)
- 0903 Pension Systems for the Informal Sector in Asia  
edited by Landis MacKellar, March 2009 (online only)
- 0902 Structural Educational Reform: Evidence from a Teacher's Displacement Program in Armenia  
by Arvo Kuddo, January 2009 (online only)
- 0901 Non-performance of the Severance Pay Program in Slovenia  
by Milan Vodopivec, Lilijana Madzar, Primož Dolenc, January 2009 (online only)

**To view Social Protection Discussion papers published prior to 2009, please visit [www.worldbank.org/sp](http://www.worldbank.org/sp).**



### **Abstract**

This review of results frameworks in the Bank's Social Protection & Labor portfolio seeks to answer the question: "How well are Social Protection & Labor projects performing with respect to their focus on results?" Specifically, the review assesses the status and quality of Monitoring & Evaluation (M&E) in the portfolio, including both investment and policy-based lending, and develops guidance for task teams based on the review. The review identifies trends, strengths and weaknesses, and good practice M&E approaches and indicators to incorporate a better results focus in project design and implementation.

**HUMAN DEVELOPMENT NETWORK**

#### **About this series...**

Social Protection Discussion Papers are published to communicate the results of The World Bank's work to the development community with the least possible delay. The typescript manuscript of this paper therefore has not been prepared in accordance with the procedures appropriate to formally edited texts. The findings, interpretations, and conclusions expressed herein are those of the author(s), and do not necessarily reflect the views of the International Bank for Reconstruction and Development / The World Bank and its affiliated organizations, or those of the Executive Directors of The World Bank or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work.

For free copies of this paper, please contact the Social Protection Advisory Service, The World Bank, 1818 H Street, N.W., Room G7-703, Washington, D.C. 20433 USA. Telephone: (202) 458-5267, Fax: (202) 614-0471, E-mail: [socialprotection@worldbank.org](mailto:socialprotection@worldbank.org) or visit the Social Protection website at [www.worldbank.org/sp](http://www.worldbank.org/sp).