INDEPENDENT Evaluation Office
United Nations Development Programme

وزارة التخطيط والمتابعة والإصلاح الإداري
Ministry of Planning, Monitoring and Administrative Reform

LEAVING NO ONE BEHIND: EVALUATION for 2030

2019 National Evaluation Capacities Conference

Theory-based evaluation in practice

Jos Vaessen, PhD
World Bank IEG

@undp_evaluation

#NECdev

IEG Independent Evaluation Group | WORLD BANK GROUP
World Bank • IFC • MIGA

Theory-based evaluation in practice

Session 1: Introduction

October 2019

Jos Vaessen, PhD

Learning objectives

Focus:

1. What is theory-based evaluation and why is it important?
2. What are useful principles for reconstructing a program theory?
3. How can we apply theory-based evaluation in practice?

Learning outcomes:

- After this course, participants have developed an initial (but sound) understanding of the role of theory in evaluation and how to apply theory-based evaluation in practice

Outline of the workshop

9.00-9.30 Session 1: Introduction

9.30-10.30 Session 2: Principles of TBE

10.30-11.00 Break

11.00-13.00 Session 3: Reconstructing a program theory (group exercise)

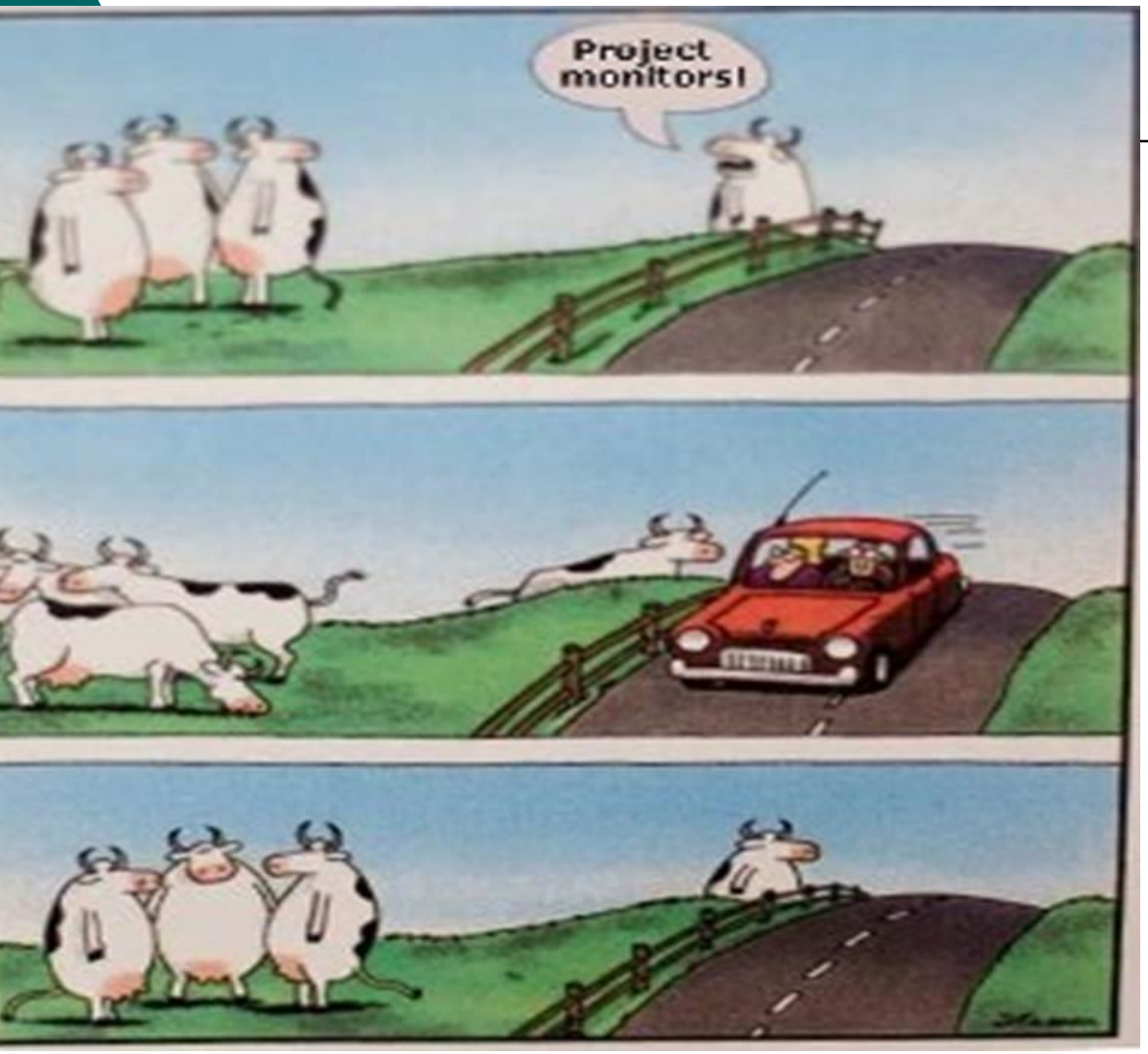
13.00-14.00 Lunch

14.00-15.30 Session 4: Reconstructing a program theory (group presentations and discussion)

15.30-16.00 Break

16.00-17.15 Session 5: Applications of TBE

17.15-17.30 Closing and evaluation



What kind of biases can you identify?

IEG | WORLD BANK GROUP
Independent Evaluation Group | World Bank • IFC • MIGA

Theory-based evaluation in practice

Session 2: Principles of TBE

October 2019

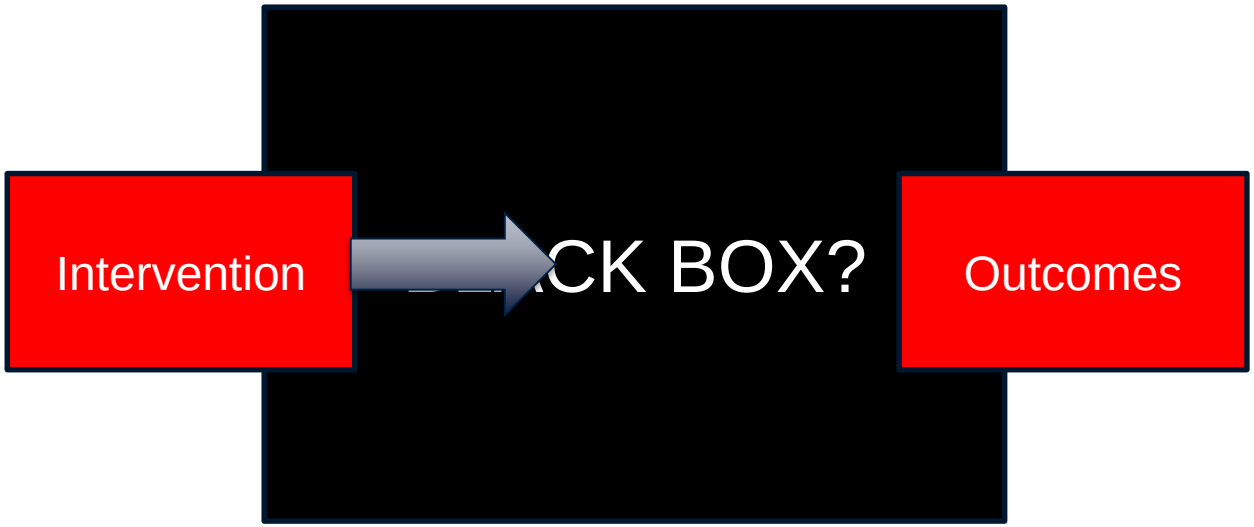
Jos Vaessen, PhD

Purposes of TBE

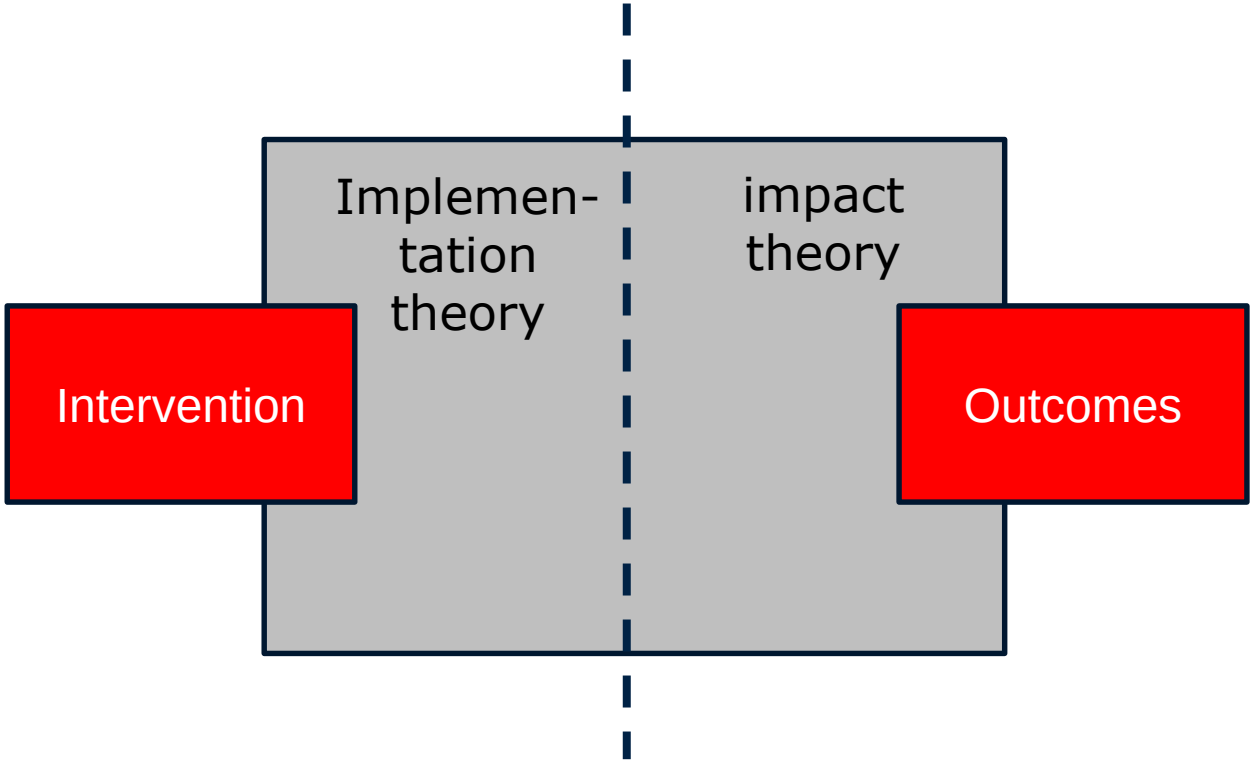
- **Understanding why** interventions do or do not work (implementation versus theory failure)
- **Generating a consensus** on what the intervention is intended to achieve and how (formative use)
- Program theory as an overall **sense-making framework**
- Using program theory as **a basis for data collection and analysis** or M&E system
- Dealing with **causality**


I THINK YOU SHOULD BE MORE SPECIFIC HERE IN STEP TWO
THEN A MIRACLE OCCURS ...

How and under what conditions are interventions expected to contribute to processes of change and outcomes?



Suchman (1967), Patton (1997) Weiss (1998), Rossi et al. (2004)



Theory failure vs. implementation failure

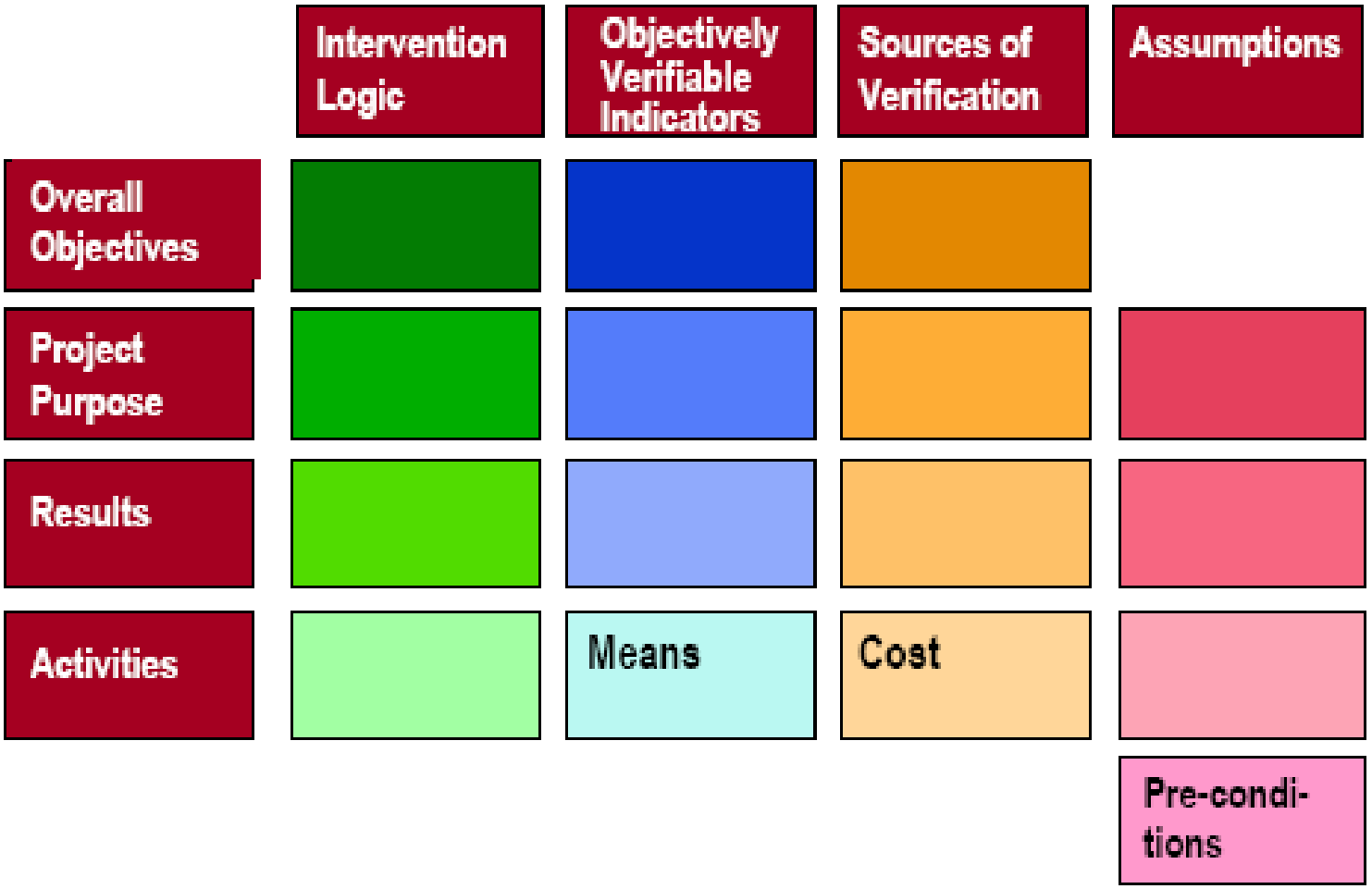
Definition

"[Program theory] is a set of hypotheses upon which people build their program plans" (Weiss, 1998:55).

"[TBE] consists of an explicit theory or model of how the program causes the intended or observed outcomes and an evaluation that is at least partly guided by this model" (Rogers et al., 2000:5).

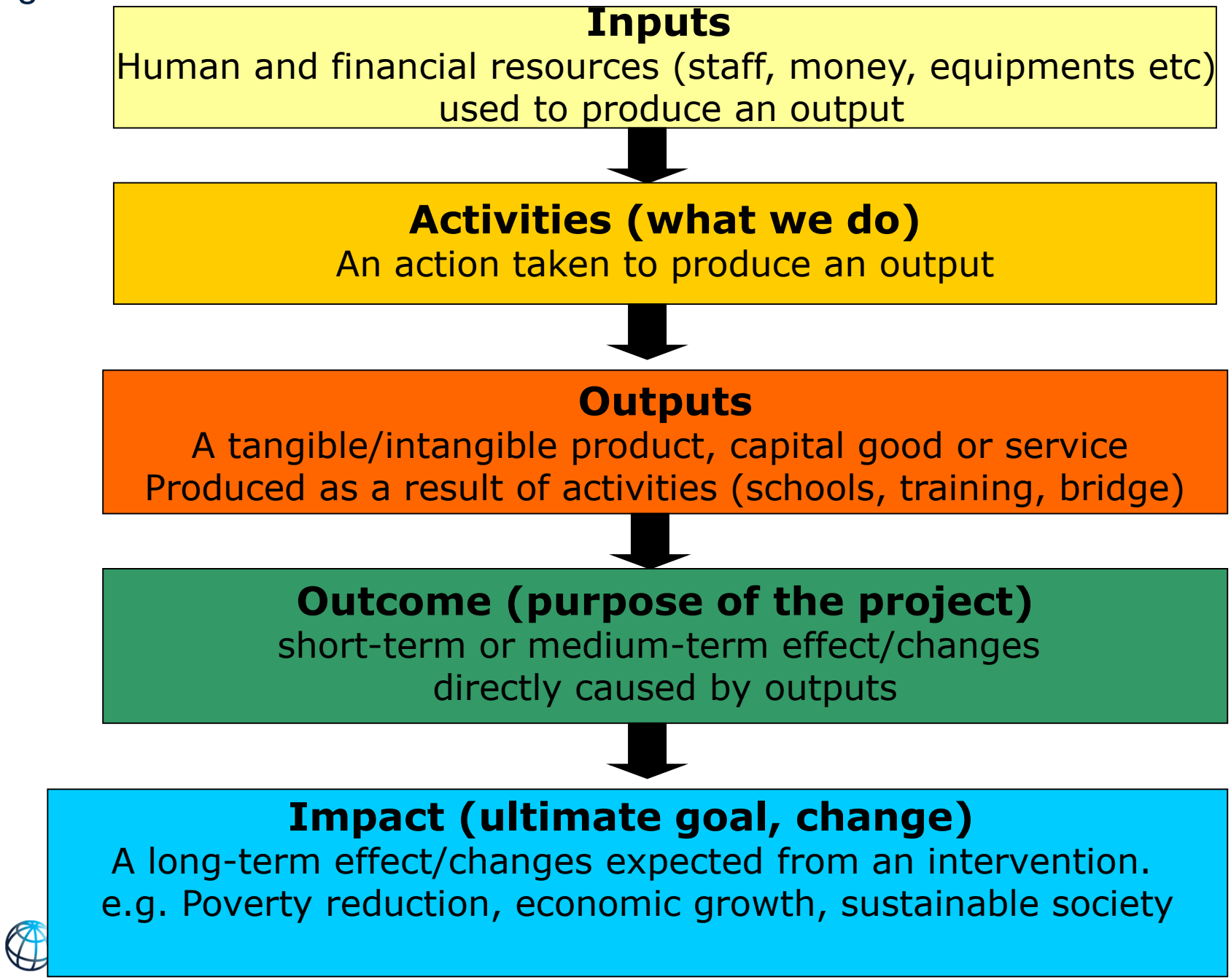
Program theory cannot be simply 'observed' but most be reconstructed

The Logical Framework Matrix

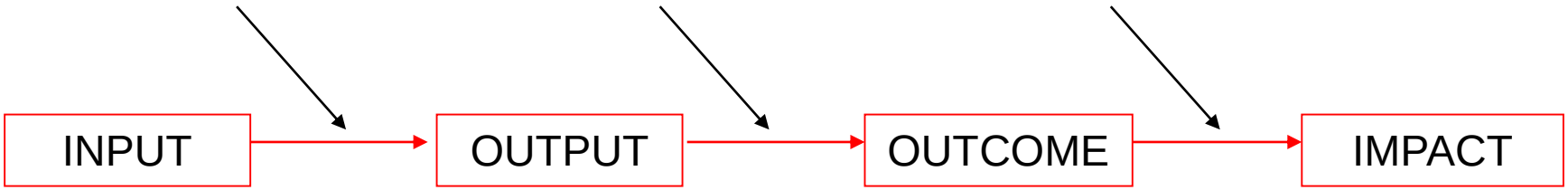


Source: EU PCM manual

Logic model / results chain

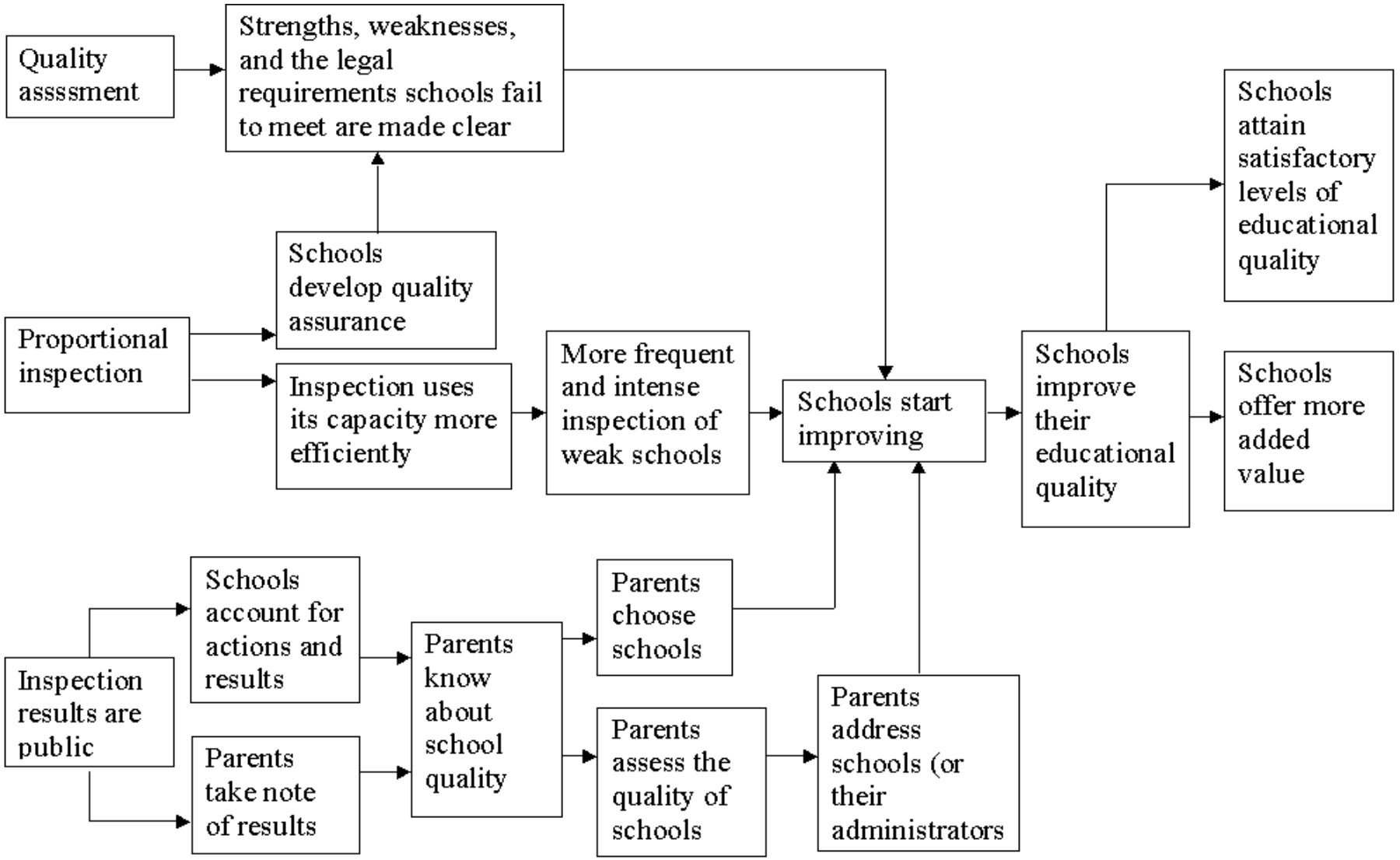


Simplified example causal chain: rural microfinance



INPUT	OUTPUT	OUTCOME	IMPACT
Organization	Mobilizing savings	Increased investments in education, housing, production, etc.	Increased income
Resources	Providing loans	Reduction in income volatility	'Empowerment
.....	Creating community groups	Enhanced self-respect clients	Reduced social exclusion
	.....	Increased participation of vulnerable groups in in socio-political arena	Institutional change
		.....	.....

Program theory: effects of school inspection



Program theory: training in organic agriculture

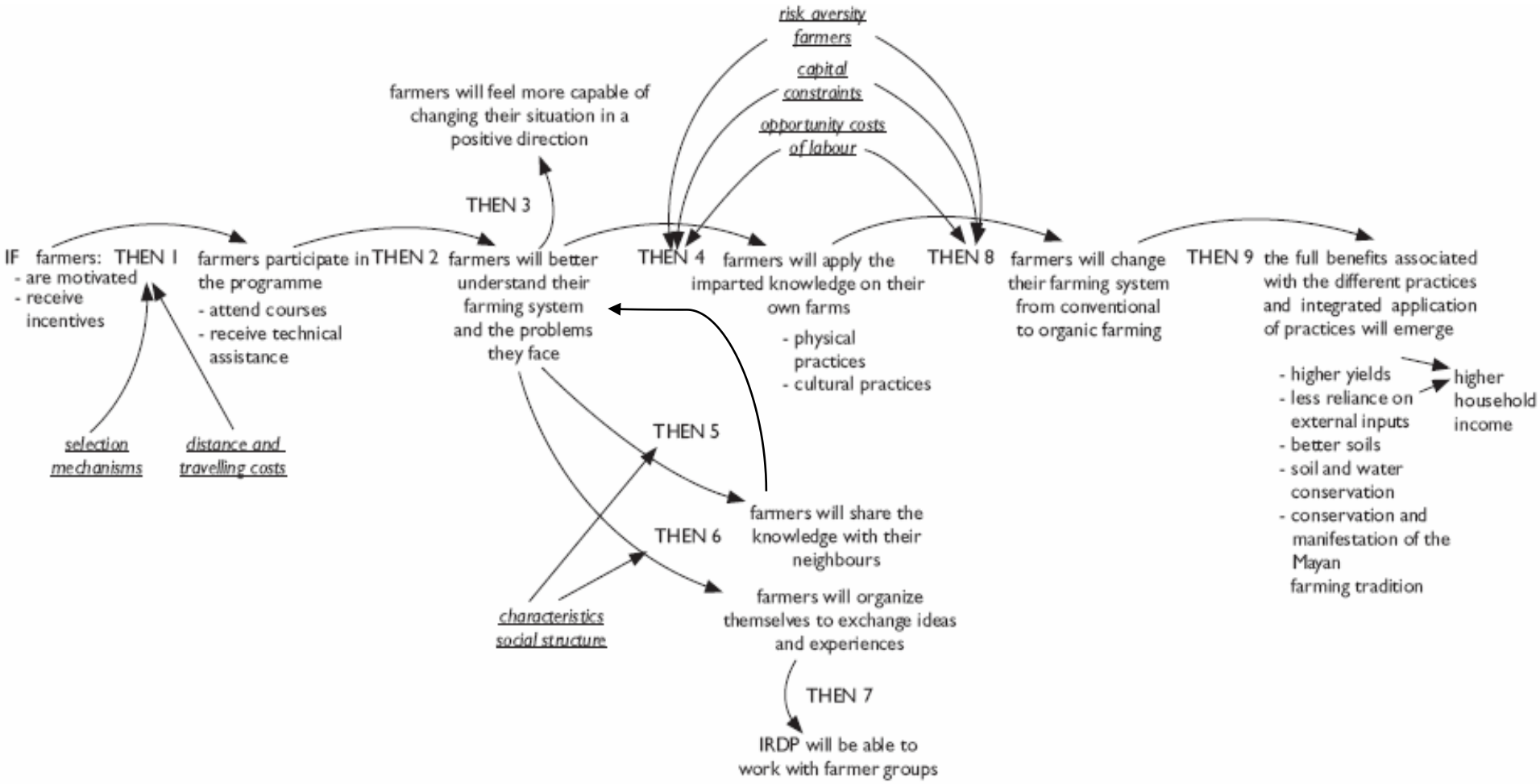
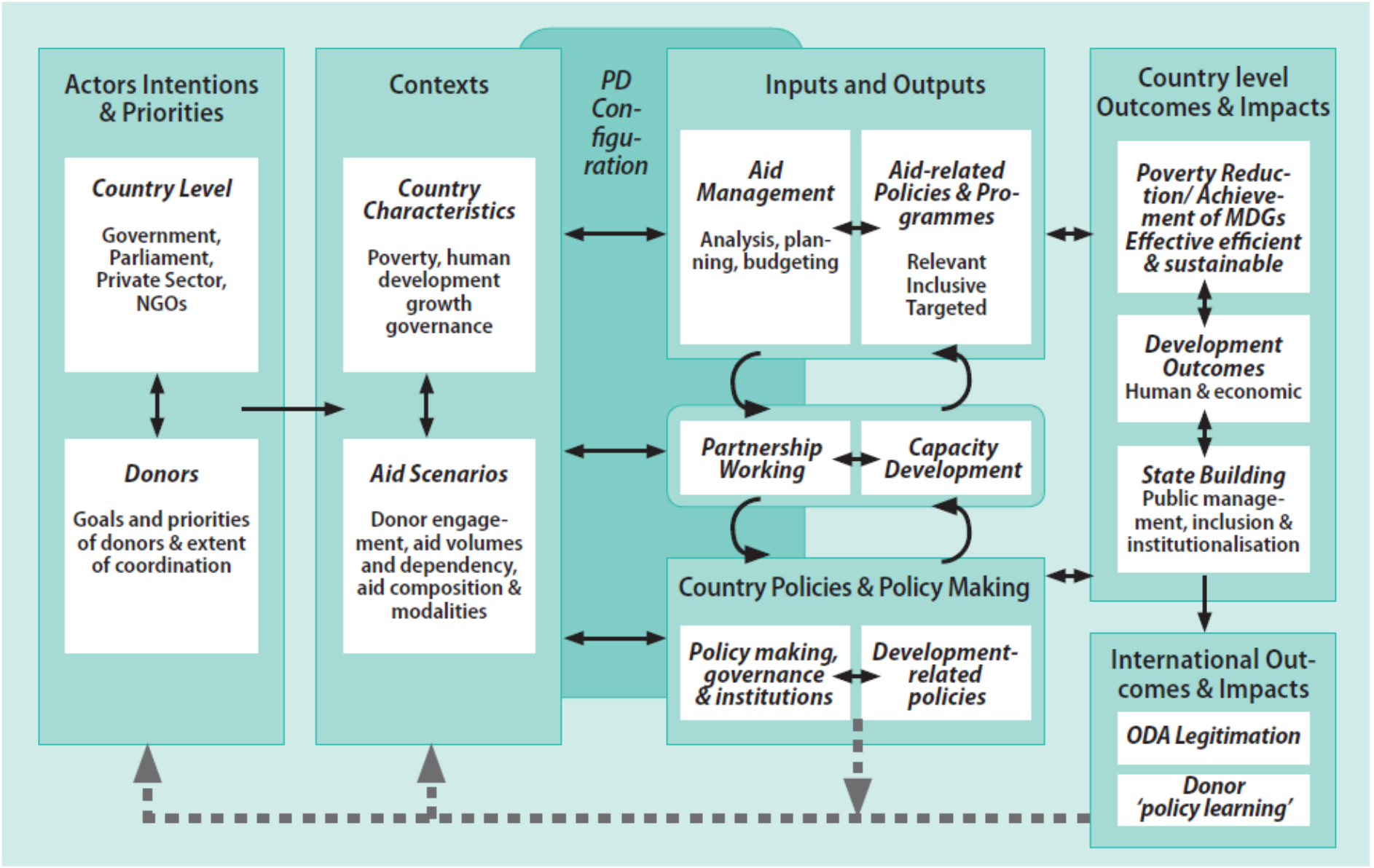
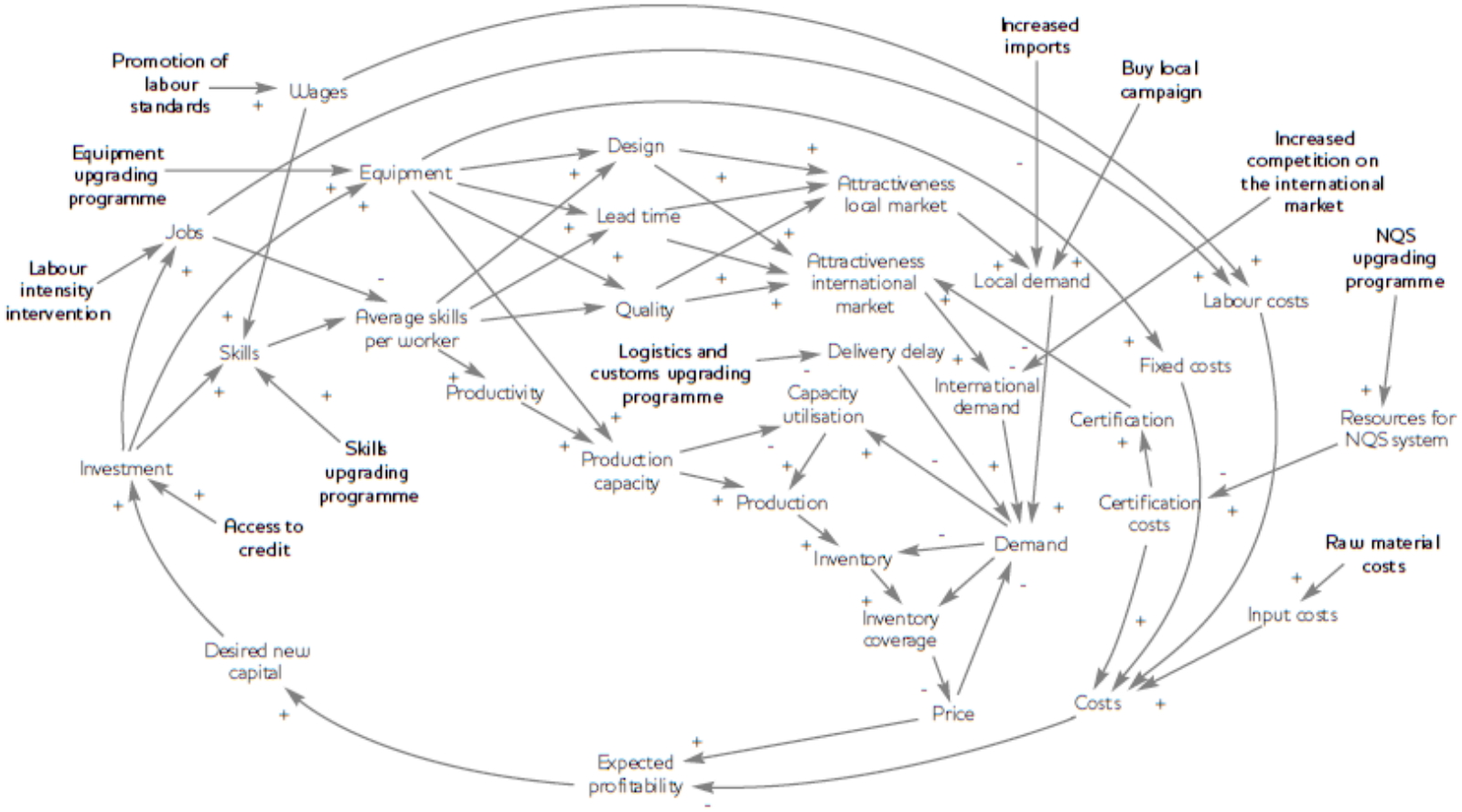


Figure 3.2 A Policy Model for PD Evaluation



A complex system… the leather shoes sector in Ethiopia



Derwisch & Loewe (2015) "Systems Dynamics Modelling in Industrial Development Evaluation" *IDS Bulletin* 46.1

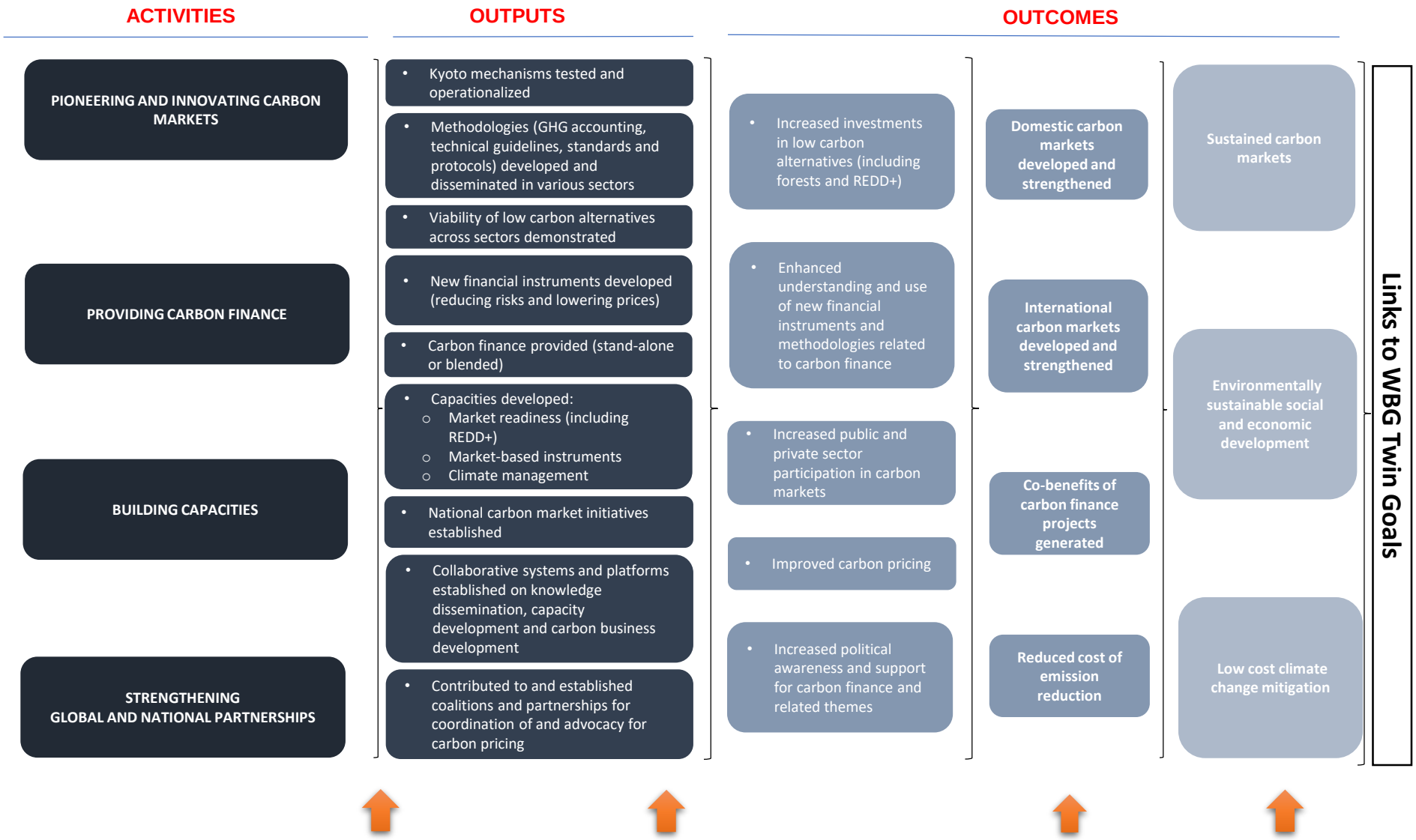
Methods for reconstructing program theories

- There are different methods available for reconstructing (see for example Chen, 1990; Leeuw, 2003; Funnell and Rogers, 2011)

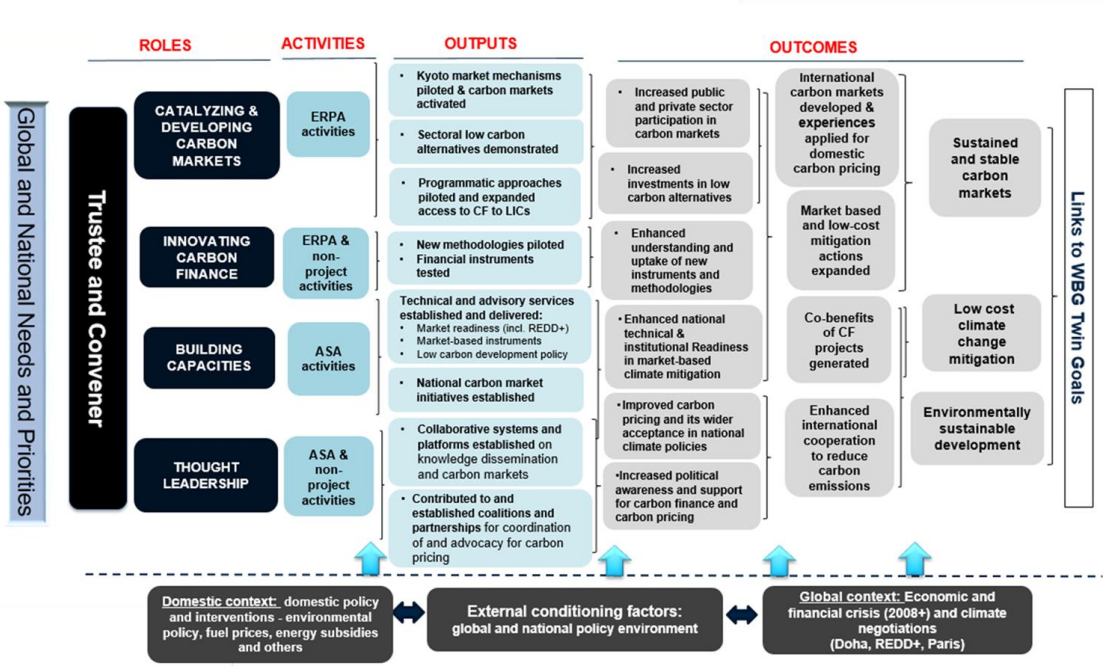
- In ex post evaluation scenarios the following principles apply:

  1. Be specific
  2. Be consistent in formulations
  3. Think about the warrants (i.e. is it logical to expect that *a* contributes to *b*)
  4. Think about the underlying assumptions (i.e. under what conditions is *a* likely to contribute to *b*)

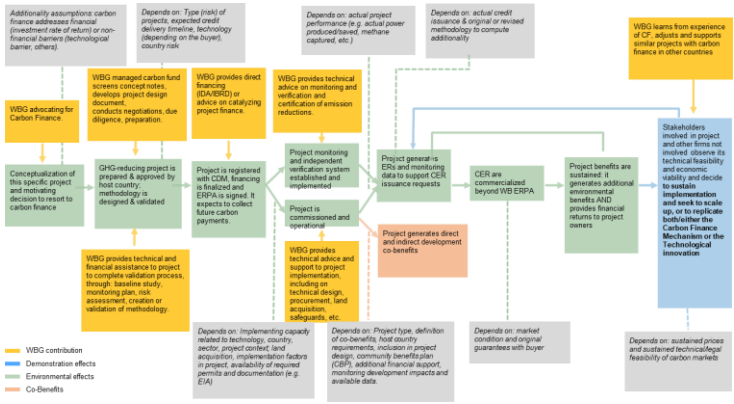
Program theory as a sense-making framework



Evaluation of WBG support to carbon finance

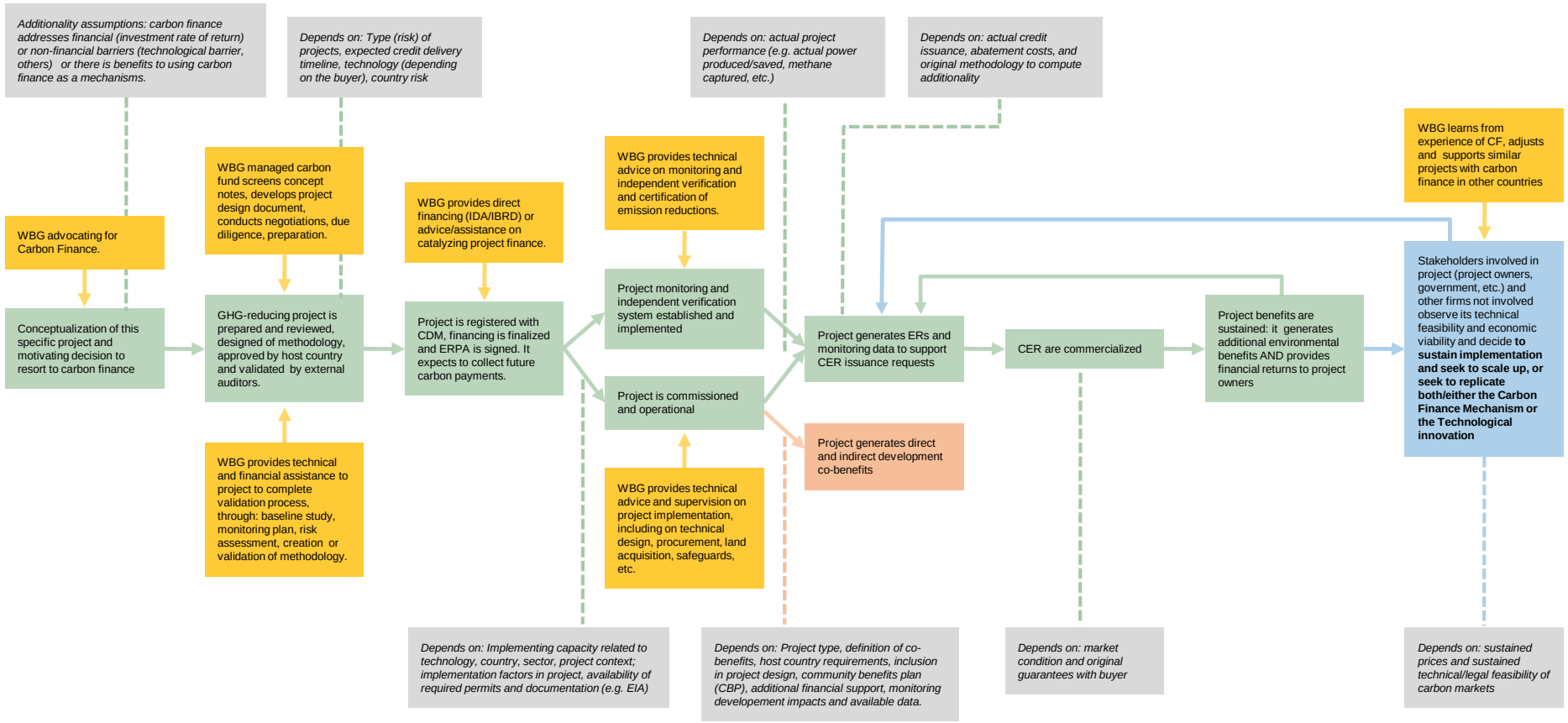


Synthetic 'high-level' ToC



Nested 'detailed' ToC

Evaluation of WBG support to carbon finance



Frameworks for developing program theories

- Policy instruments: sticks, carrots, sermons (Bemelmans-Videc et al., 2003)

- Behavioral mechanisms: social norms, profit-seeking behavior, demonstration and copying behavior, peer pressure, etc.

- Coleman's Theory of Social Action (1986)
  - Situational mechanisms
  - Action-formation mechanisms
  - Transformational mechanisms

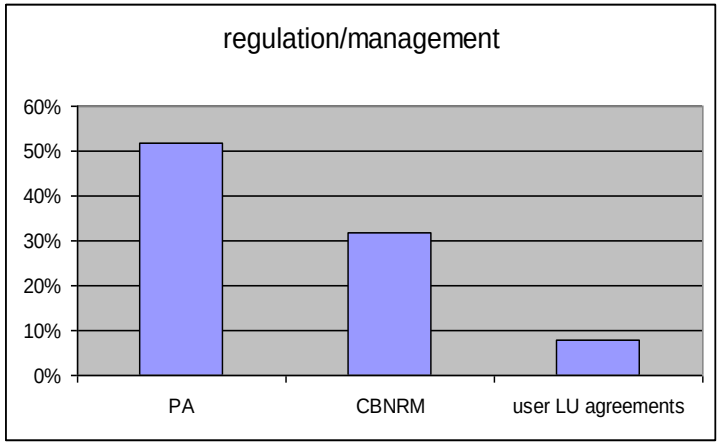
- Intervention-specific templates for program theory

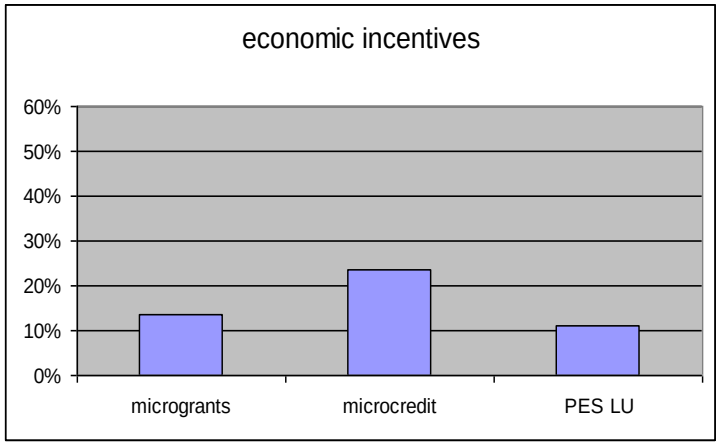
Looking across projects vs. looking across intervention activities

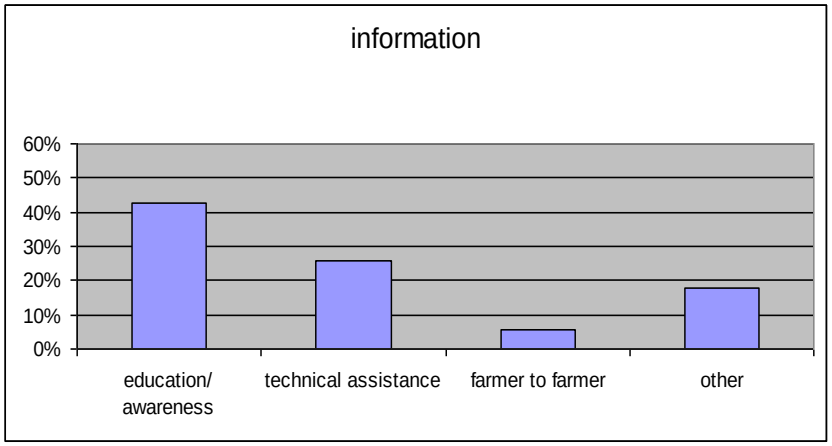
Portfolio-level: GEF-funded activities directed at rural landowners

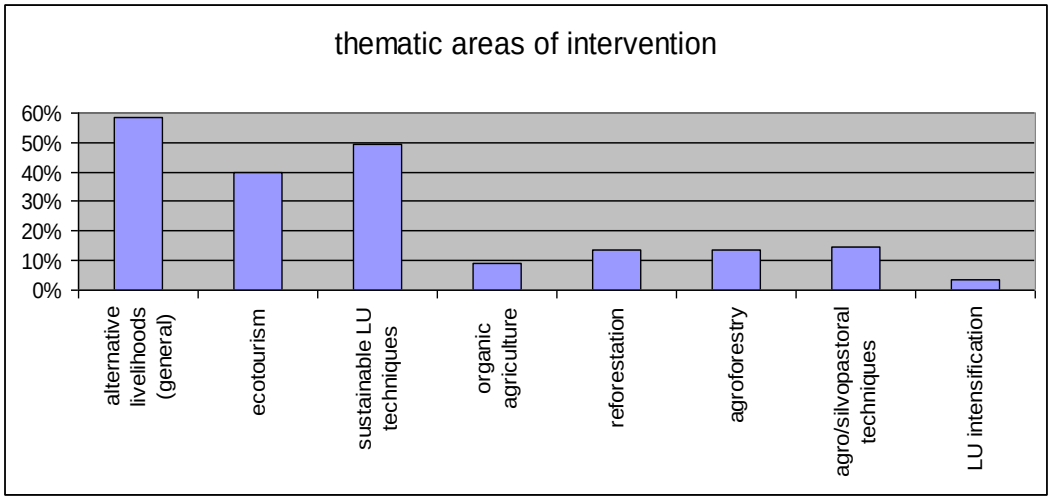
N = 332;

All FSPs and MSPs approved in period X from the Biodiversity, Land Degradation, Multi-Focal Areas focal areas.









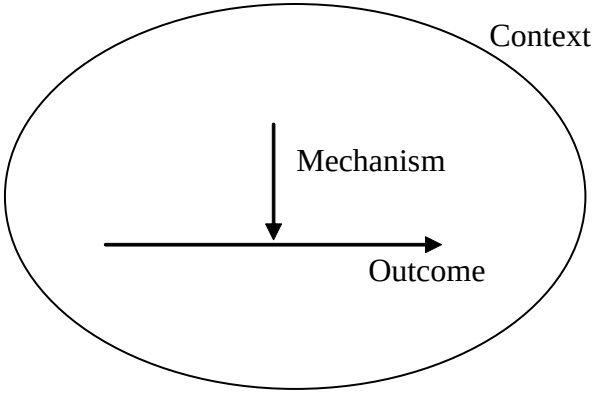
IEG Independent Evaluation Group | WORLD BANK GROUP World Bank • IFC • MIGA

Focus on behavioral mechanisms: CMO theory (Pawson and Tilley, 1997)

There is no grand theory of social change, only patterns of regularity (Merton, 1967; Elster, 1989; Pawson and Tilley, 1997; Hedström and Swedberg,1998; Astbury and Leeuw, 2010)

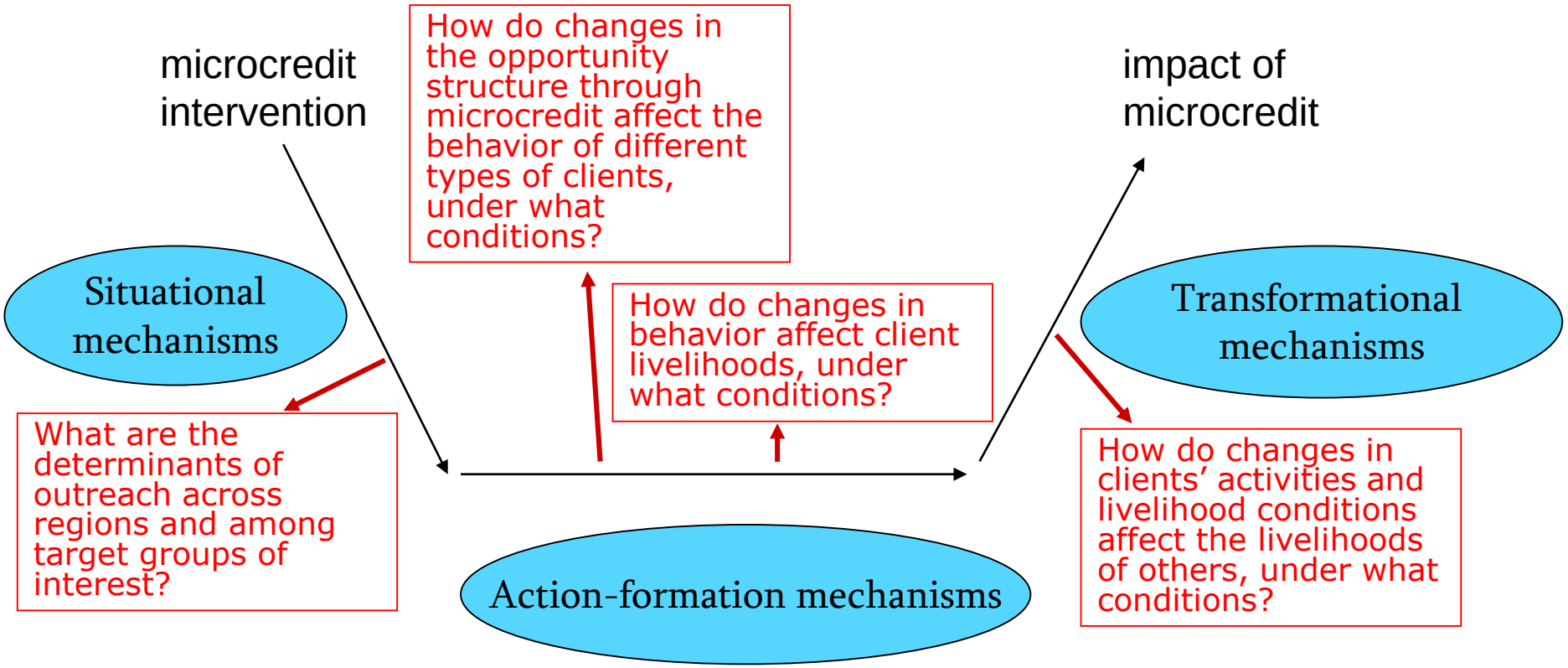
Describing patterns of change in terms of mechanisms, contextual variables and outcomes

Generative causality: under what conditions does an intervention trigger a response (mechanism) that results in particular outcome



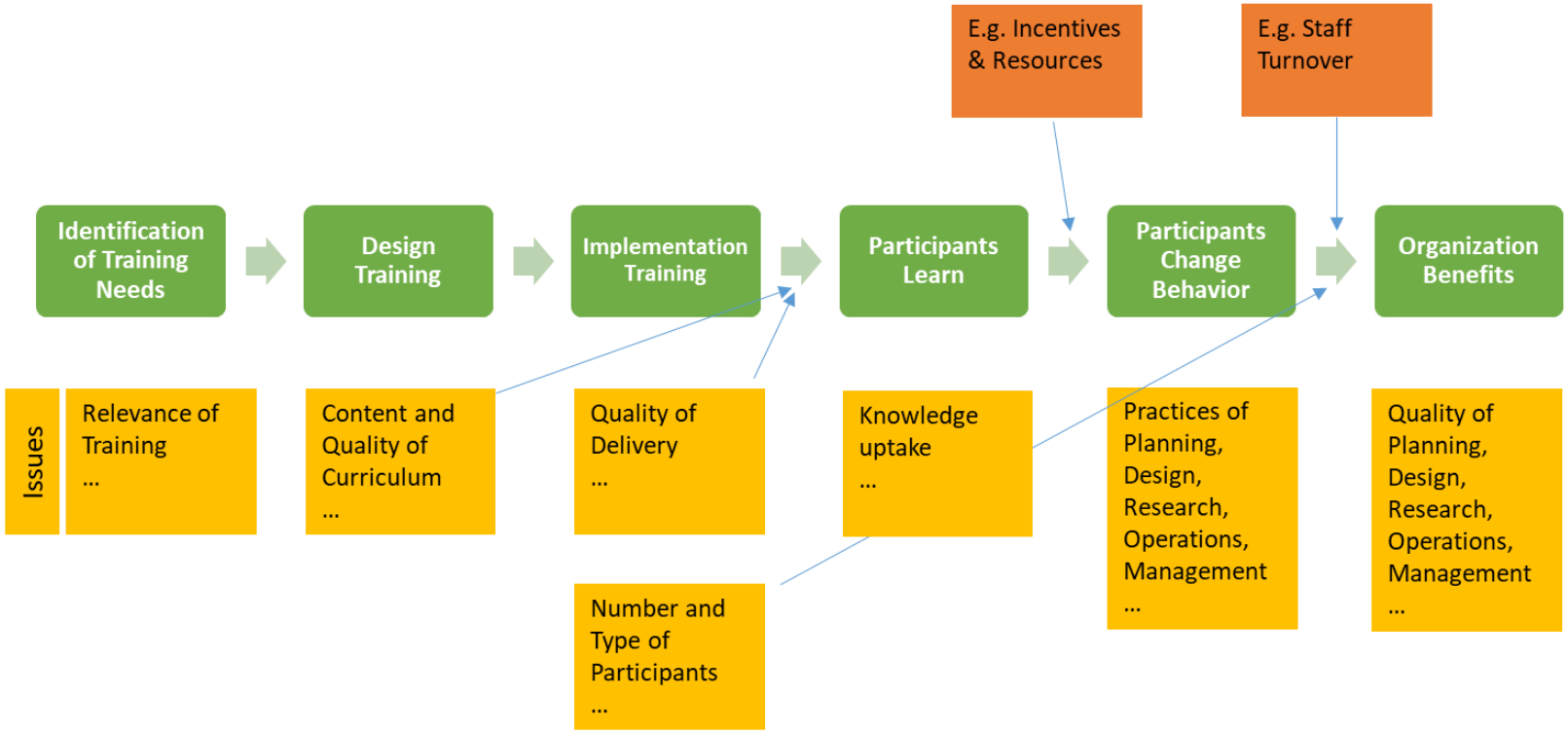
Impact theory - microcredit

Based on Coleman (1986, 1990); Hedström and Swedberg (1998), see also Leeuw (2008)



Intervention-specific templates for program theory

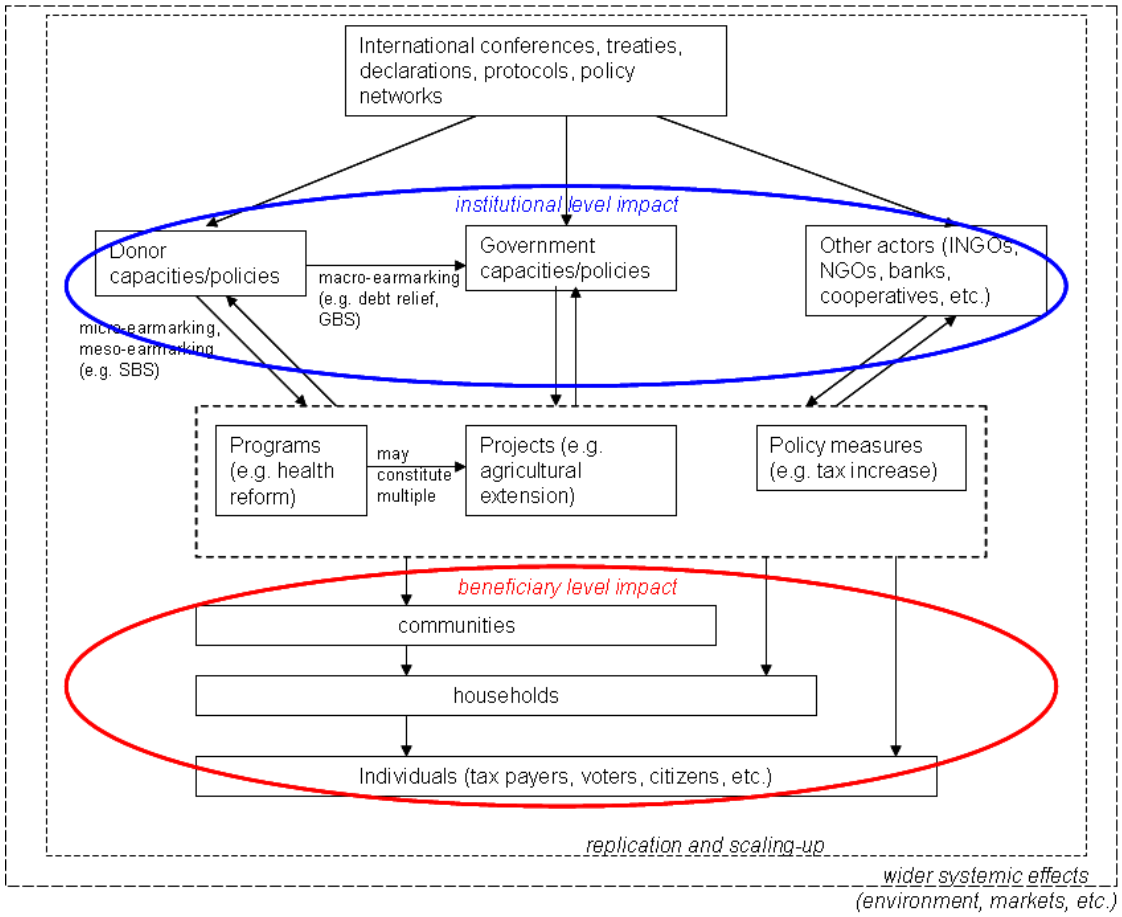
Simplified Theory of Change Training



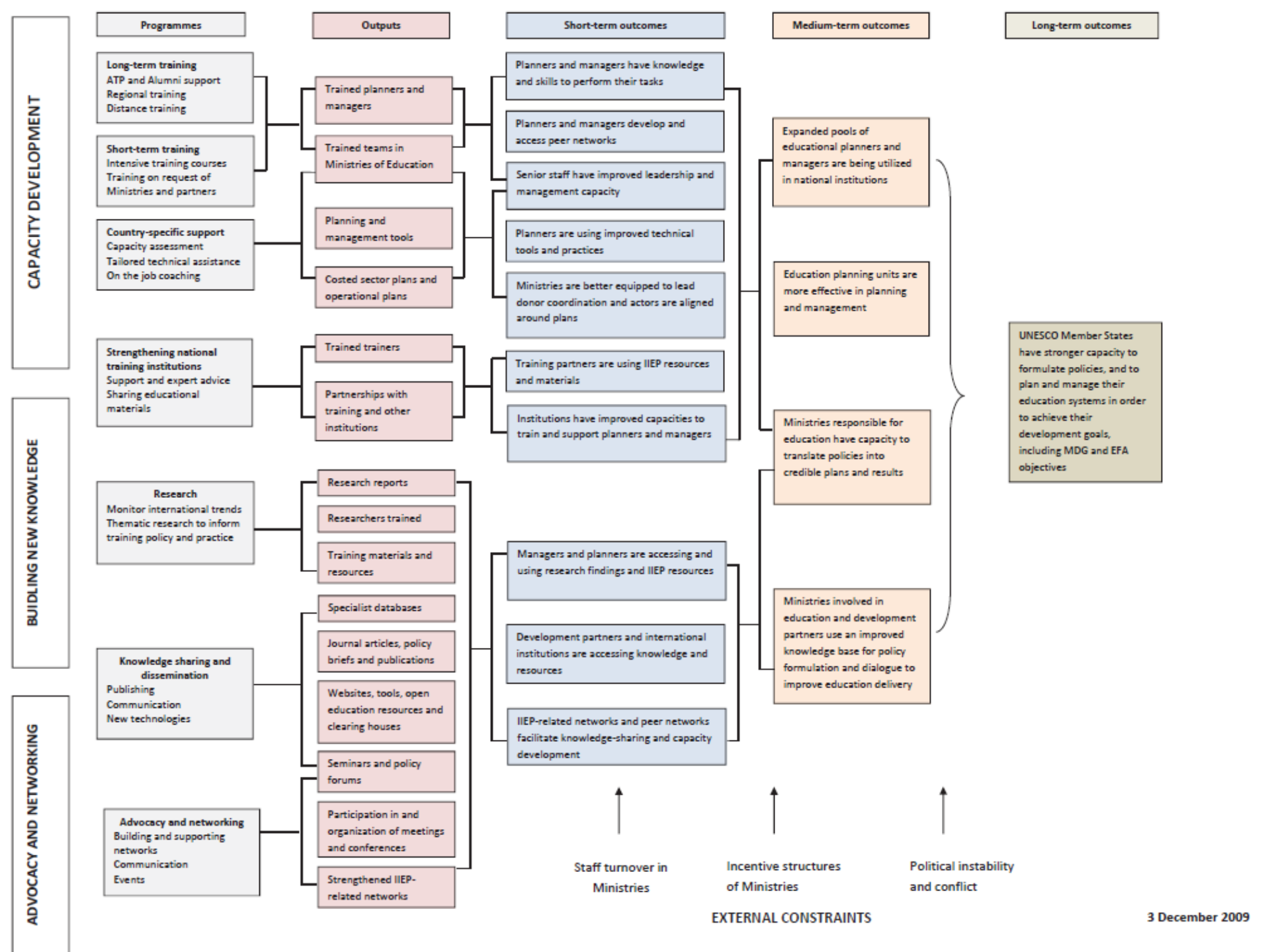
Methods

Specific methods and data sources differ according to causal step or underlying assumption

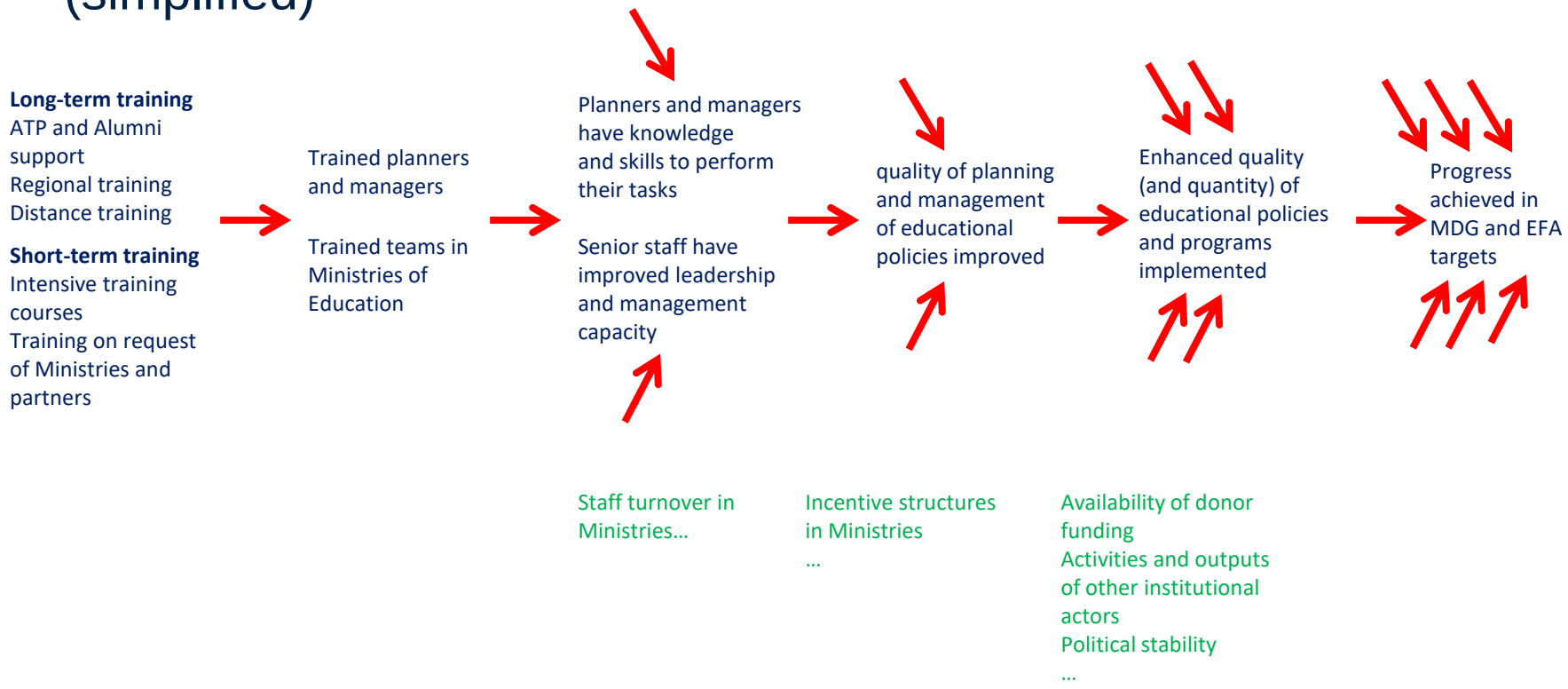
Levels of analysis



Leeuw and Vaessen (2009)


CAPACITY DEVELOPMENT
BUIDLING NEW KNOWLEDGE
ADVOCACY AND NETWORKING
Programmes
Outputs
Short-term outcomes
Medium-term outcomes
Long-term outcomes
Long-term training
ATP and Alumni support
Regional training
Distance training
Short-term training
Intensive training courses
Training on request of Ministries and partners
Country-specific support
Capacity assessment
Tailored technical assistance
On the job coaching
Strengthening national training institutions
Support and expert advice
Sharing educational materials
Research
Monitor international trends
Thematic research to inform training policy and practice
Knowledge sharing and dissemination
Publishing
Communication
New technologies
Advocacy and networking
Building and supporting networks
Communication
Events
Trained planners and managers
Trained teams in Ministries of Education
Planning and management tools
Costed sector plans and operational plans
Trained trainers
Partnerships with training and other institutions
Research reports
Researchers trained
Training materials and resources
Specialist databases
Journal articles, policy briefs and publications
Websites, tools, open education resources and clearing houses
Seminars and policy forums
Participation in and organization of meetings and conferences
Strengthened IIEP-related networks
Planners and managers have knowledge and skills to perform their tasks
Planners and managers develop and access peer networks
Senior staff have improved leadership and management capacity
Planners are using improved technical tools and practices
Ministries are better equipped to lead donor coordination and actors are aligned around plans
Training partners are using IIEP resources and materials
Institutions have improved capacities to train and support planners and managers
Managers and planners are accessing and using research findings and IIEP resources
Development partners and international institutions are accessing knowledge and resources
IIEP-related networks and peer networks facilitate knowledge-sharing and capacity development
Expanded pools of educational planners and managers are being utilized in national institutions
Education planning units are more effective in planning and management
Ministries responsible for education have capacity to translate policies into credible plans and results
Ministries involved in education and development partners use an improved knowledge base for policy formulation and dialogue to improve education delivery
UNESCO Member States have stronger capacity to formulate policies, and to plan and manage their education systems in order to achieve their development goals, including MDG and EFA objectives
Staff turnover in Ministries
Incentive structures of Ministries
Political instability and conflict
EXTERNAL CONSTRAINTS
3 December 2009

Nested impact theory of capacity development activities (simplified)



Further down the causal chain causality becomes more diffuse and attribution is not possible

Sources of theory

- Stakeholder theory:
  - explicit and tacit theory
- Research-based theory:
  - reference group behavior, naming and shaming, peer pressure, utility maximization, diffusion of innovations, social norms, anticipatory action, etc.
- Intervention vs. causal mechanism (!)

Whose theory? (1)

- Government, implementing organizations, beneficiaries (etc.) may have different expectations and assumptions regarding how an intervention is intended to work and what it may achieve

- Reconstructing different stakeholder theories is helpful in understanding the different views and beliefs of stakeholder groups

- Generating consensus on how an intervention is intended to work can be helpful in improving stakeholder relations and may benefit the intervention implementation process and subsequent benefits

Whose theory? (2)

- Stakeholder theory
- Substantive academic theory
- Empirical analysis

Questions: How would you use these sources of theory in case of the following evaluation purposes:

- Assessing whether management's strategic approach to developing public-private partnerships in health services delivery is realistic and feasible?
- Assessing the effect of performance-based financing schemes on the quality of health services delivery
- Contributing to the organization's understanding of its role and contribution in Health service delivery

IEG Independent Evaluation Group | WORLD BANK GROUP World Bank • IFC • MIGA

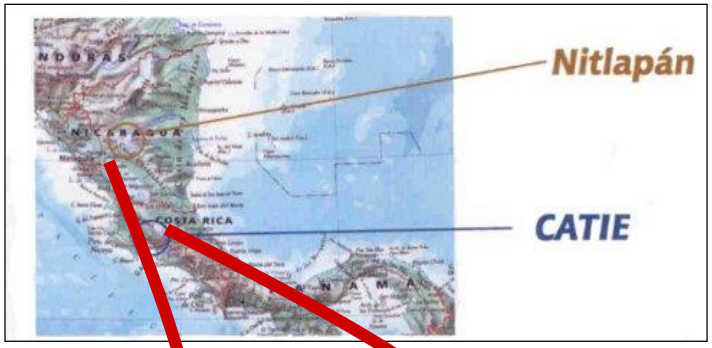
Theory-based evaluation in practice

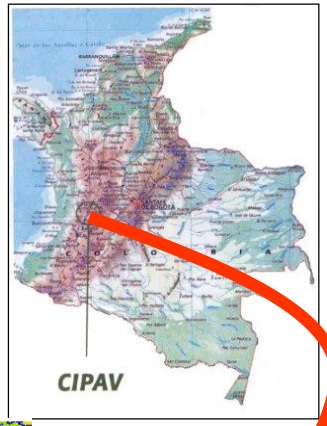
Session 3: Reconstructing a program theory (exercise)

October 2019

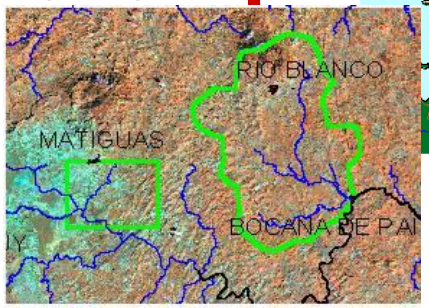
Jos Vaessen, PhD

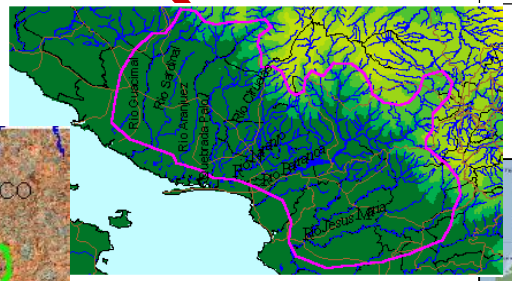
Regional project PES: Colombia, Costa Rica, Nicaragua





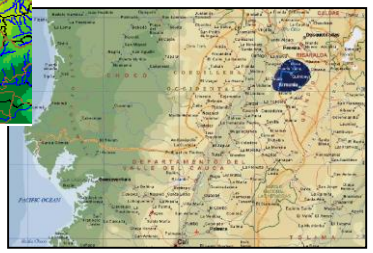
Nicaragua,
Muy-Muy





Costa Rica,
Esparza

Colombia, Quindio



Regional project PES: Colombia, Costa Rica, Nicaragua

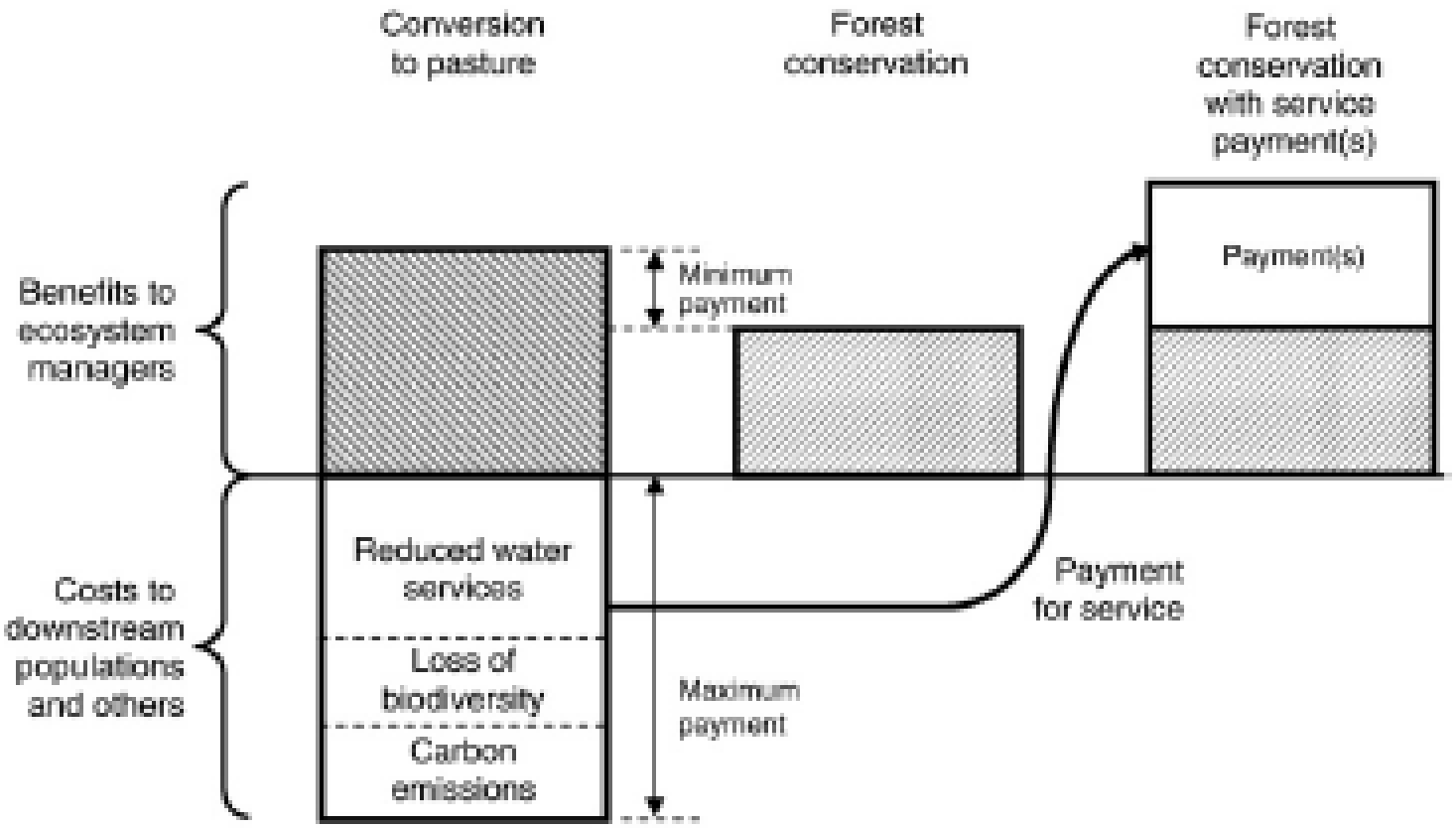
RISEMP: WB-GEF project

Pilot project on PES in agricultural landscapes

Regional project in three countries: Costa Rica, Nicaragua, Colombia

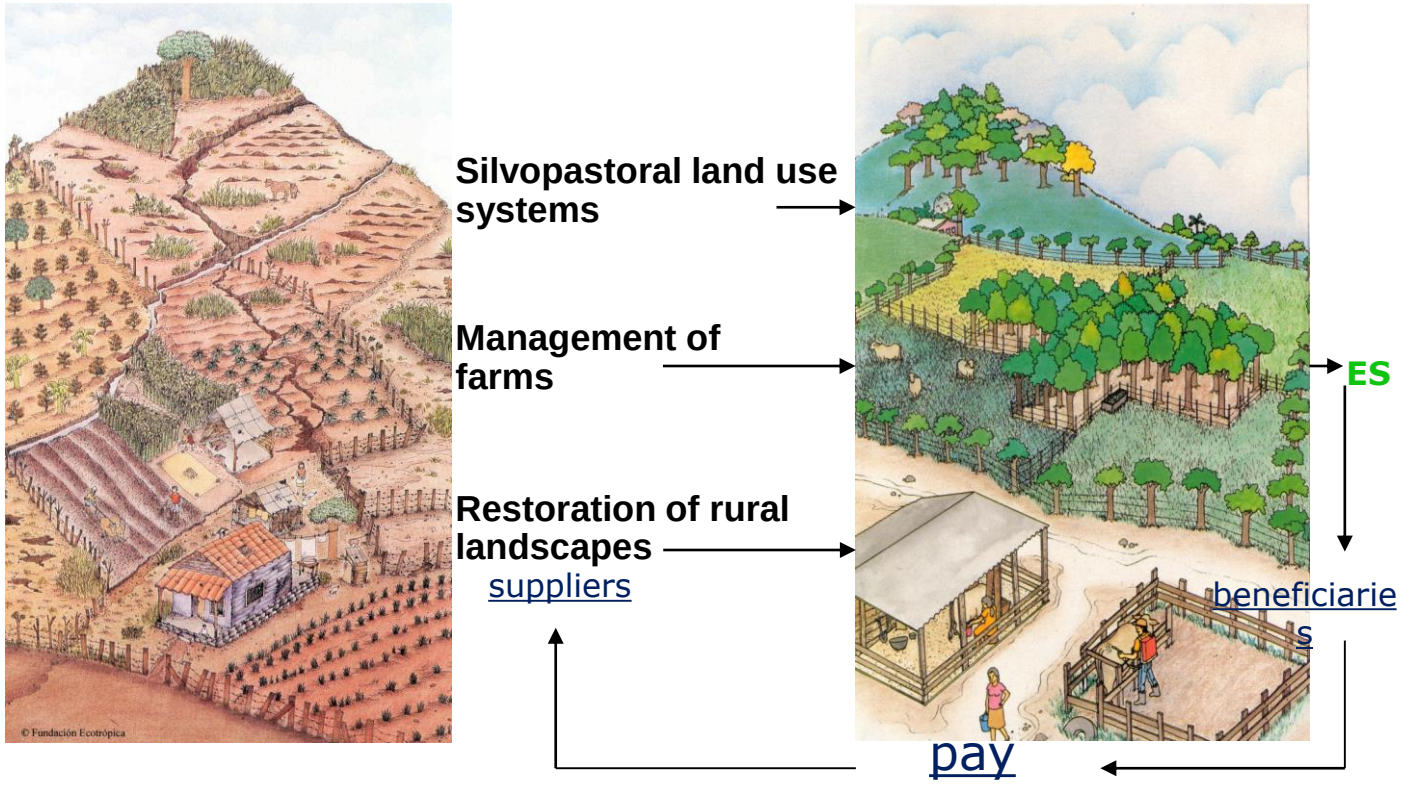
Components: payments and TA for ES; research on LU-ES relations and effectiveness of incentives; publication and dissemination of findings

Economic theory about PES



The logic of payments for environmental services. Source: Adapted from Pagiola and Platais (2007).

Regional project PES: Colombia, Costa Rica, Nicaragua



Group exercise

- Divide into groups
- Read case
- Respond to following questions:
  - what are the main patterns of intended changes that the project envisages? When trying to identify these patterns think about associations between actors, activities and processes of change/effects.
  - Please reconstruct one of these patterns into a theory of change with different causal steps and (to the extent possible) underlying causal assumptions.
- Present group findings

Principles for reconstructing program theory

1. Be specific
2. Be consistent in formulations
3. Think about the warrants (i.e. is it logical to expect that *a* contributes to *b*)
4. Think about the underlying assumptions (i.e. under what conditions is *a* likely to contribute to *b*)

IEG Independent Evaluation Group | WORLD BANK GROUP
World Bank • IFC • MIGA

Theory-based evaluation in practice

Session 4: Reconstructing a program theory (presentations and discussions)

October 2019

Jos Vaessen, PhD

IEG Independent Evaluation Group | WORLD BANK GROUP
World Bank • IFC • MIGA

Theory-based evaluation in practice

Session 5: Applications of TBE

October 2019

Jos Vaessen, PhD

TBE and methodology

- Program theory is not ‘method-specific’

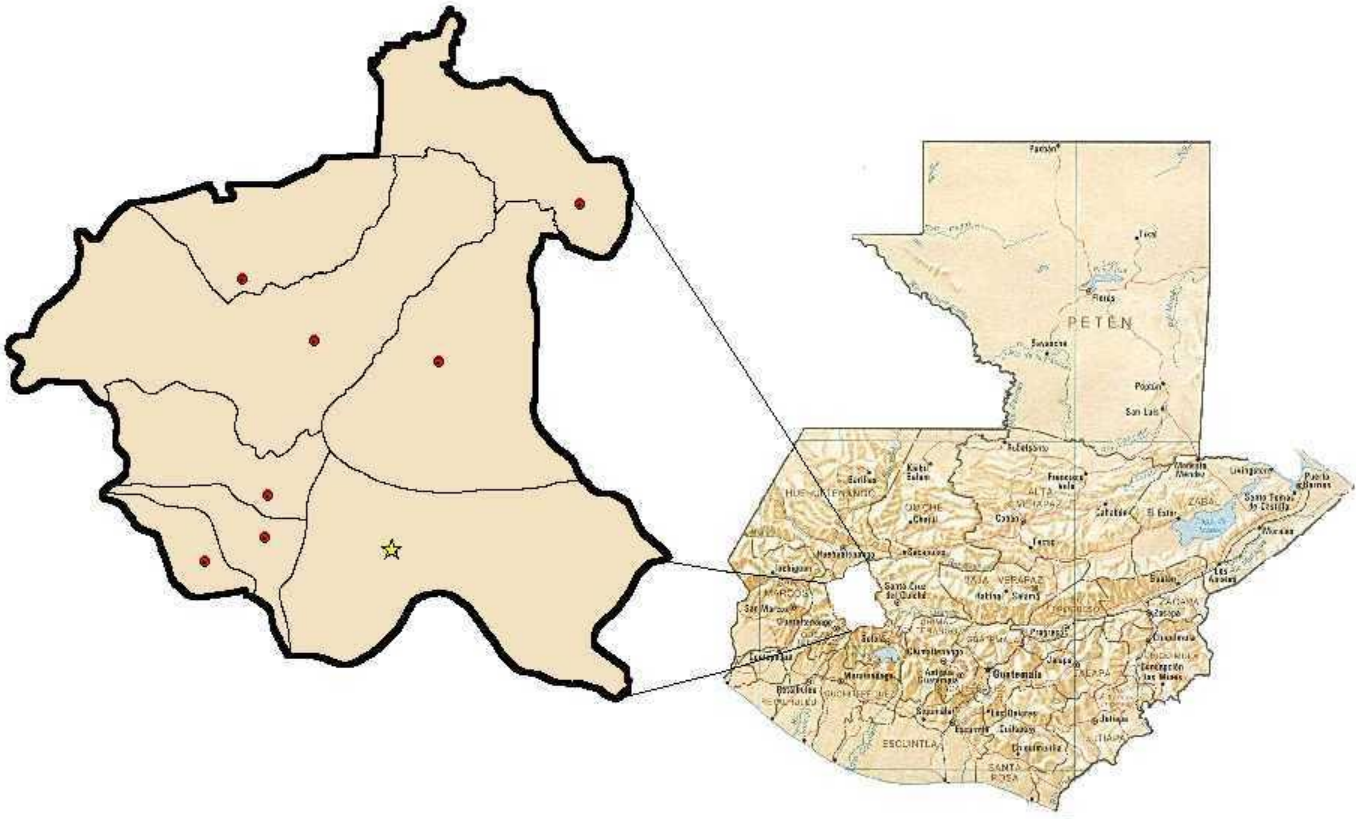
- Program theory as a framework for particular assumptions being tested / refined, using:

- (Quasi-)experimental techniques
- Regression-based techniques
- Descriptive and inferential statistical techniques
- (Advanced) modelling approaches
- Participatory techniques
- Semi-structured interviews, open interviews, focus group interviews, discourse analysis, unobtrusive measures, etc.
- Etc. etc.

Purposes of theory-based evaluation

- Understanding why interventions do or do not work (implementation versus theory failure)
- Generating a consensus on what the intervention is intended to achieve and how (formative use)
- Program theory as an overall sense-making framework
- Using program theory as a basis for data collection and analysis or M&E system
- Dealing with causality

Evaluation of training in organic agriculture



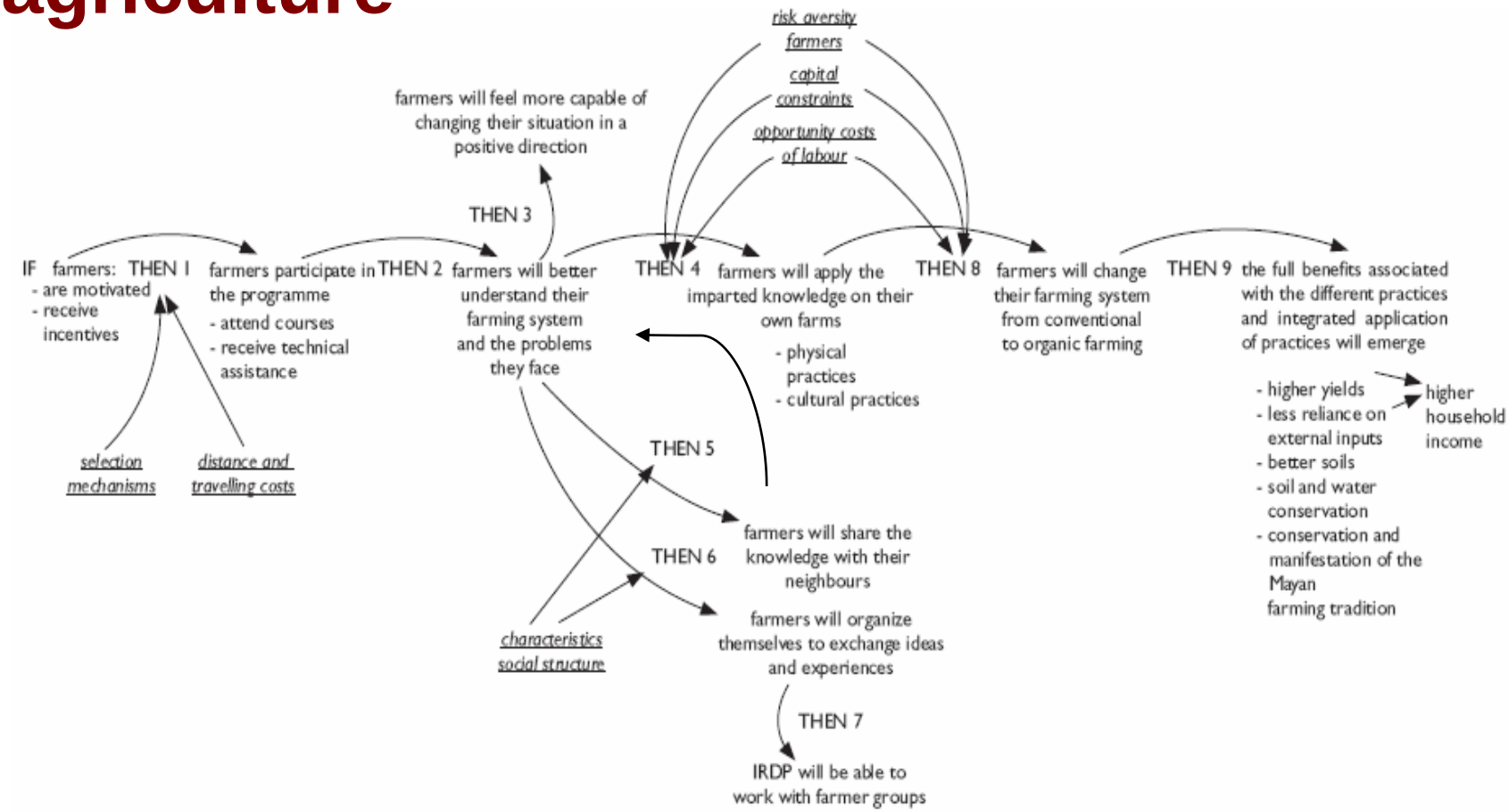
Evaluation of training in organic agriculture

- EU-supported rural development projects in 8 provinces
- In each of the provinces a national NGO provided training in organic agriculture
- In-depth evaluation (case study) of 1 out of 8 provinces
- Objective: assess implementation (participation), delivery of trainings and TA to farmers and outcomes

Multi-method approach

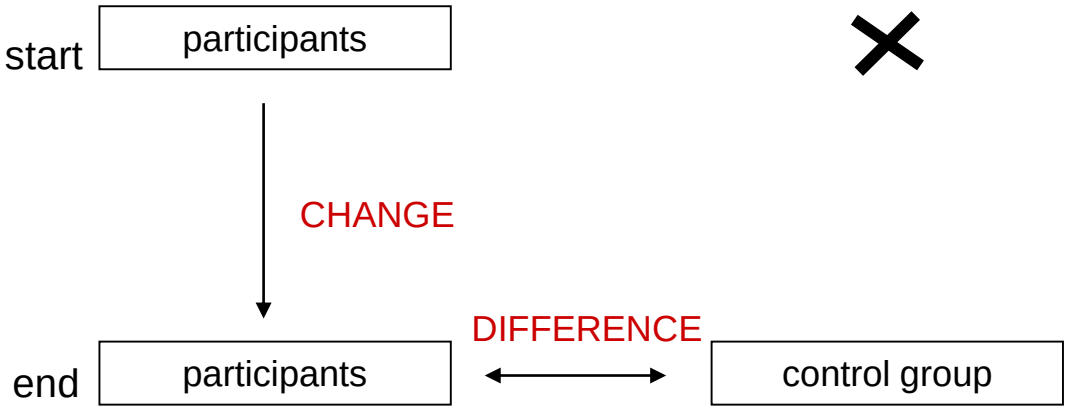
- Review of project implementation reports
- In-depth interviews with EU project staff, NGO staff, farmers
- Review of training curriculum
- Observation of training sessions
- Farms visits to inspect land use practices
- Quasi-experimental design based on baseline and ex post survey

Evaluation of training in organic agriculture



Where do the data fit into the theory?

Addressing the attribution issues: a quasi-experiment

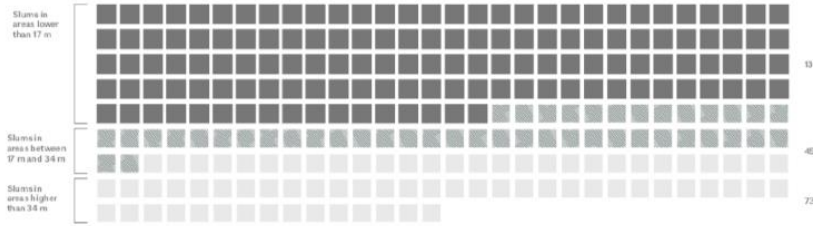


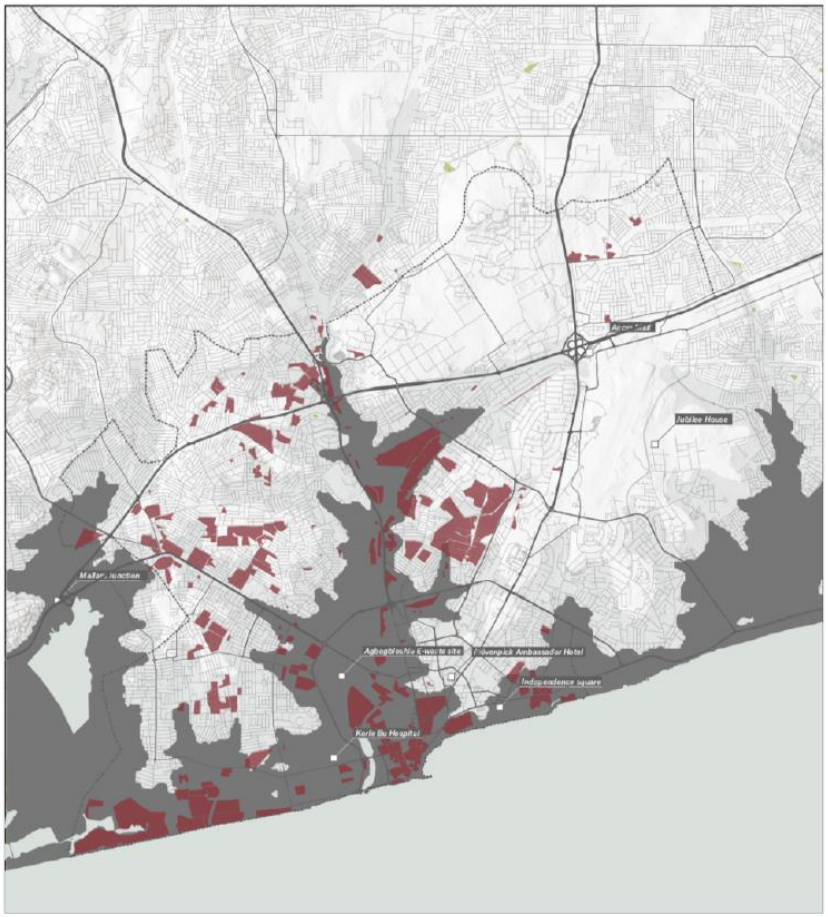
Data

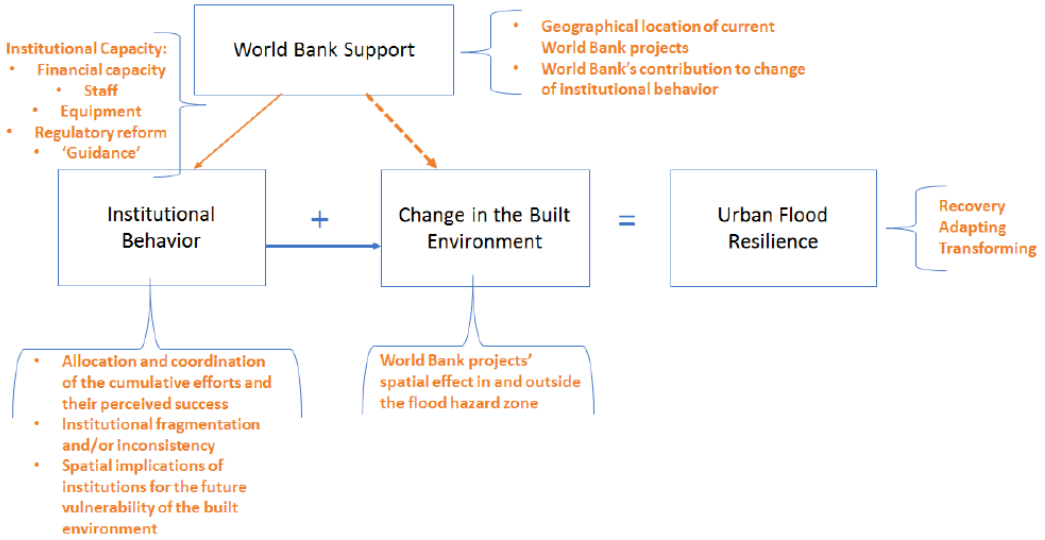
Project outcome (adoption)

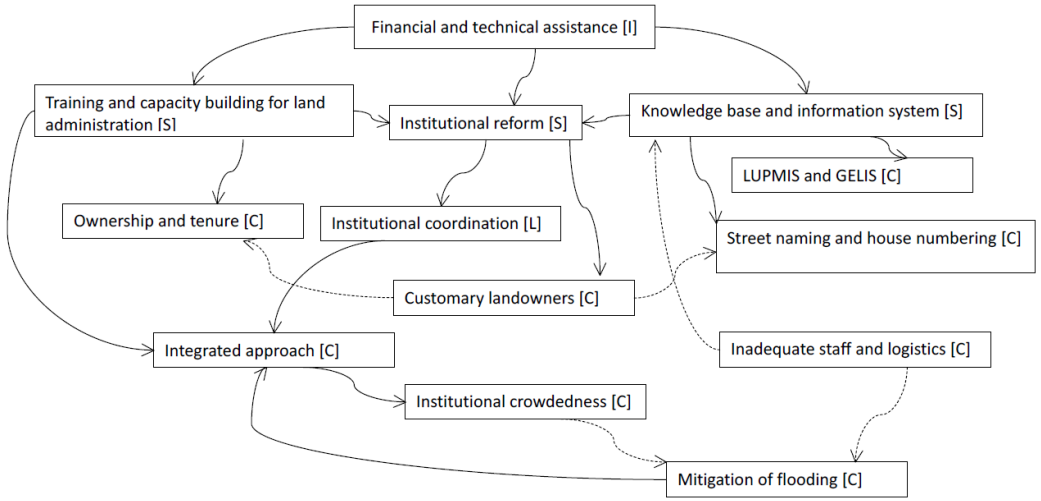
practice	participants start	participants end	control group end
burning crop residues (%)	27 % **	2 %	29 % **
applying green material (%)	25 % **	63 %	18 % **
'chemical' fertilizers (%)	96 % *	79 %	97 % *
'organic' fertilizers (%)	79 % [a]	83 %	18 % **
ditches (%)	56 % [a]	73 %	24 % **
barriers (%)	44 % [a]	58 %	21 % **
minimum tillage (%)	nihil [b]	54 %	nihil [b]
latrines (%)	15 % **	56 %	8 % **
furnaces (%)	60 %	69 %	34 % **
pig sties (%)	42 %	60 %	45 %
nurseries (%)	33 %	44 %	3 % **
medicinal plants (no. plants)	3.2 (5.3) **	8.7 (7.0)	3.2 (3.5) **
crop diversity (no. crops)	4.3 (1.7) *	4.9 (2.4)	3.2 (1.4) **
fruit tree diversity (no. trees)	4.8 (2.9) *	6.2 (3.2)	4.6 (2.3) **

Systems analysis: Flood resilience in Accra



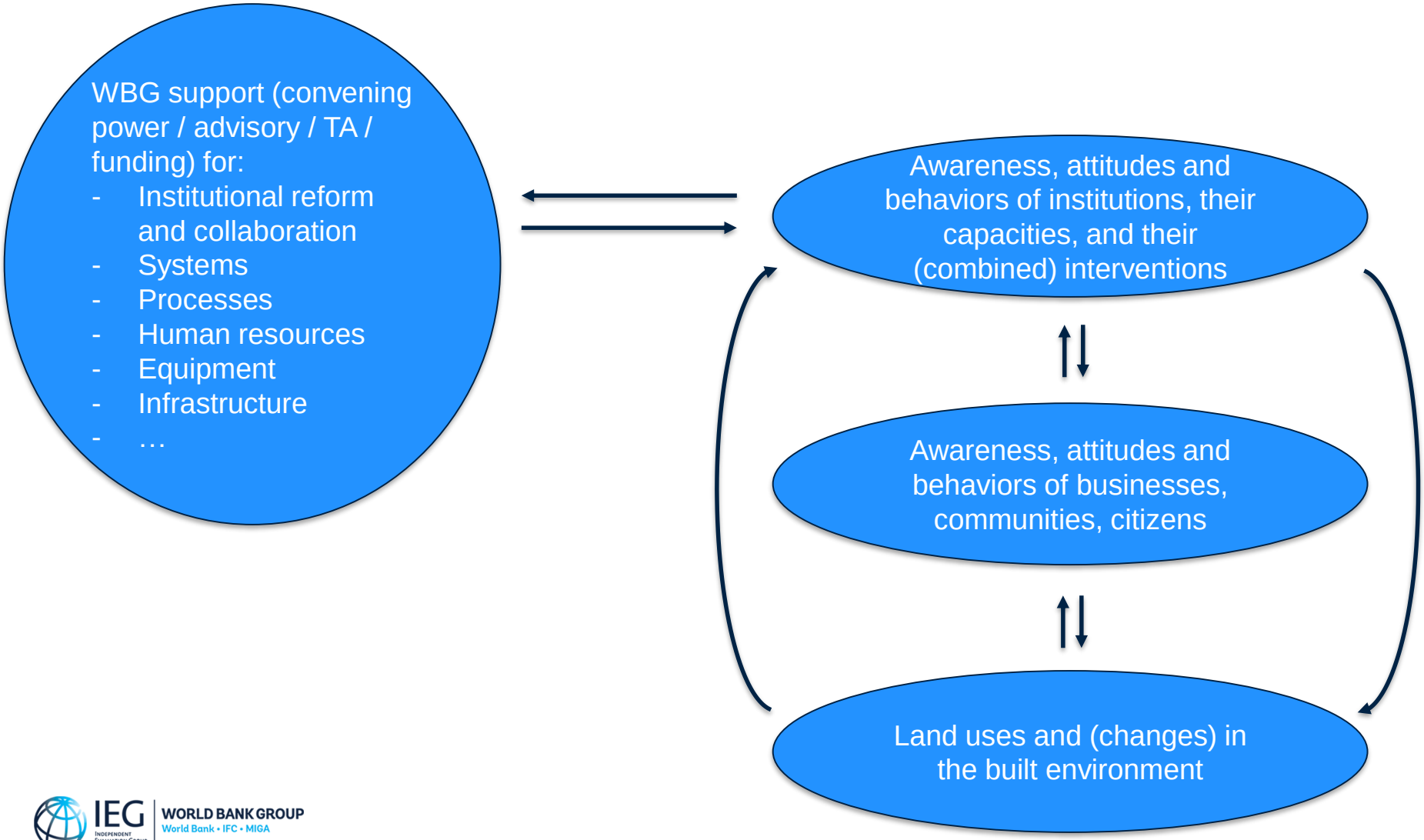




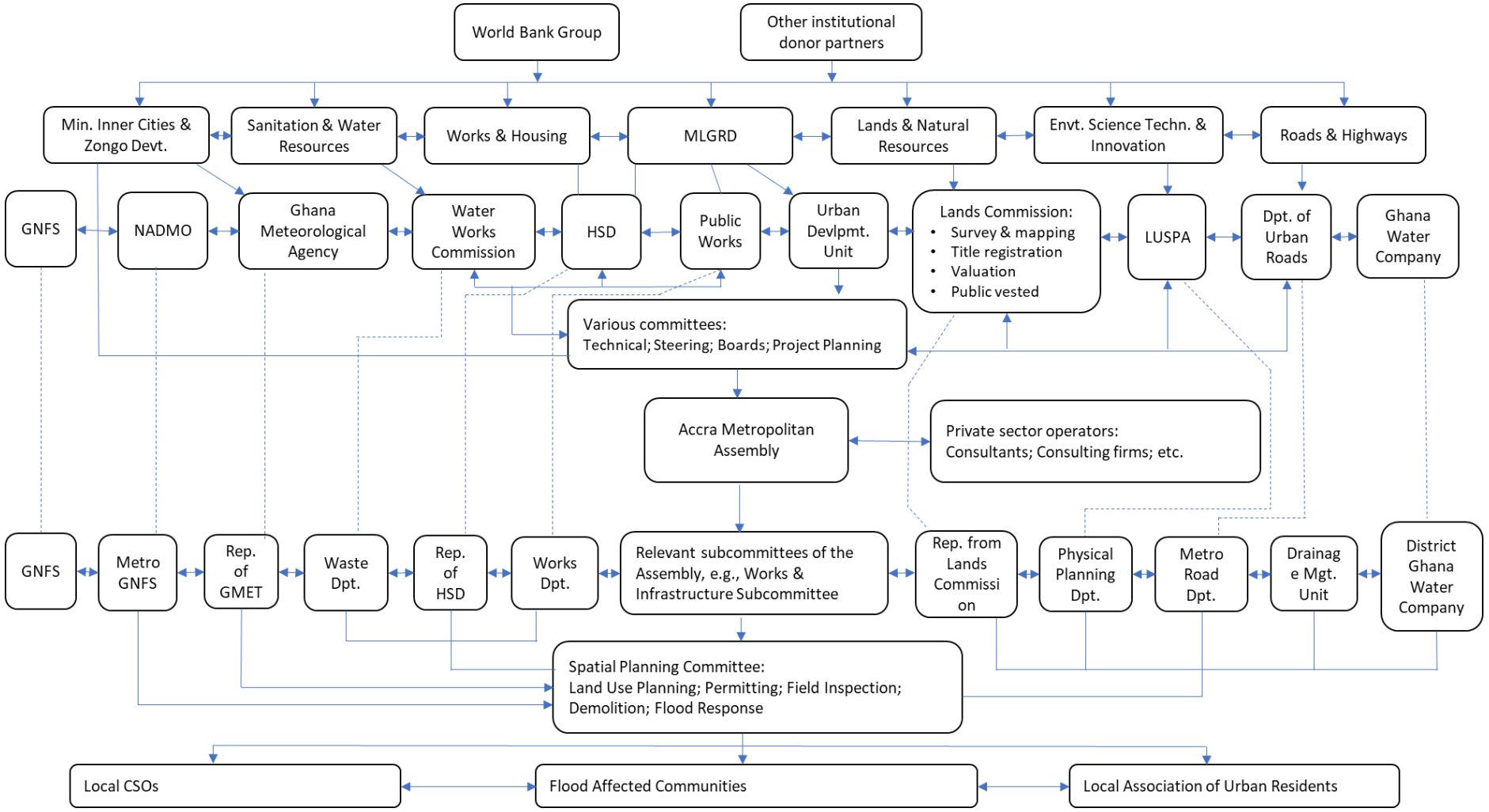


World Bank Input [I], Leverage points [L], Systems Intervention [S], Context [C]
Positive relationship → Negative relationship ⇢

Systems analysis: Flood resilience in Accra



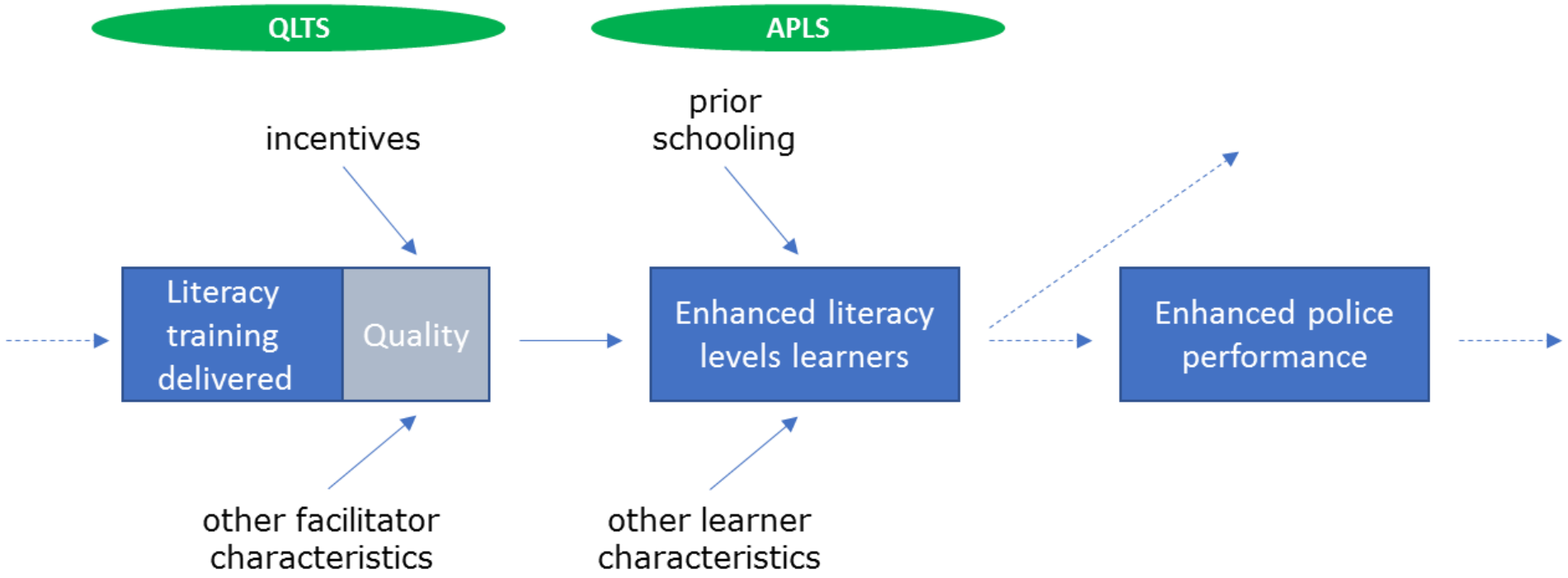
Systems analysis: Flood resilience in Accra



IEG Independent Evaluation Group | WORLD BANK GROUP World Bank • IFC • MIGA

Program theory as a framework for data collection and analysis

Example: Evaluation of police literacy training in Afghanistan



QLTS

Data collection: class observation, facilitator and commander surveys, learner focus groups

Data analysis: descriptive statistics, qualitative synthesis, OLS

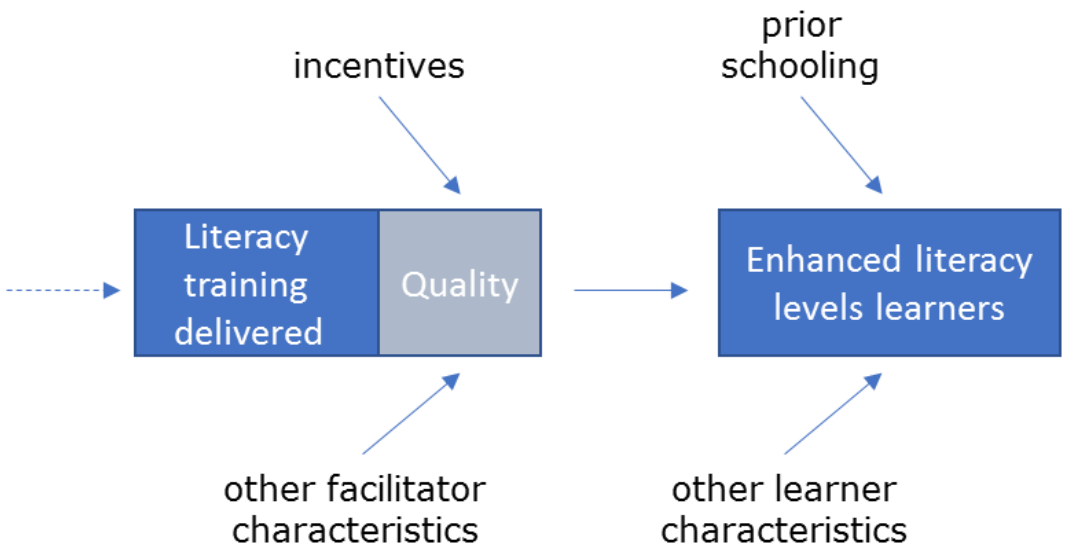
n = 274 trainings in 8 provinces

APLS

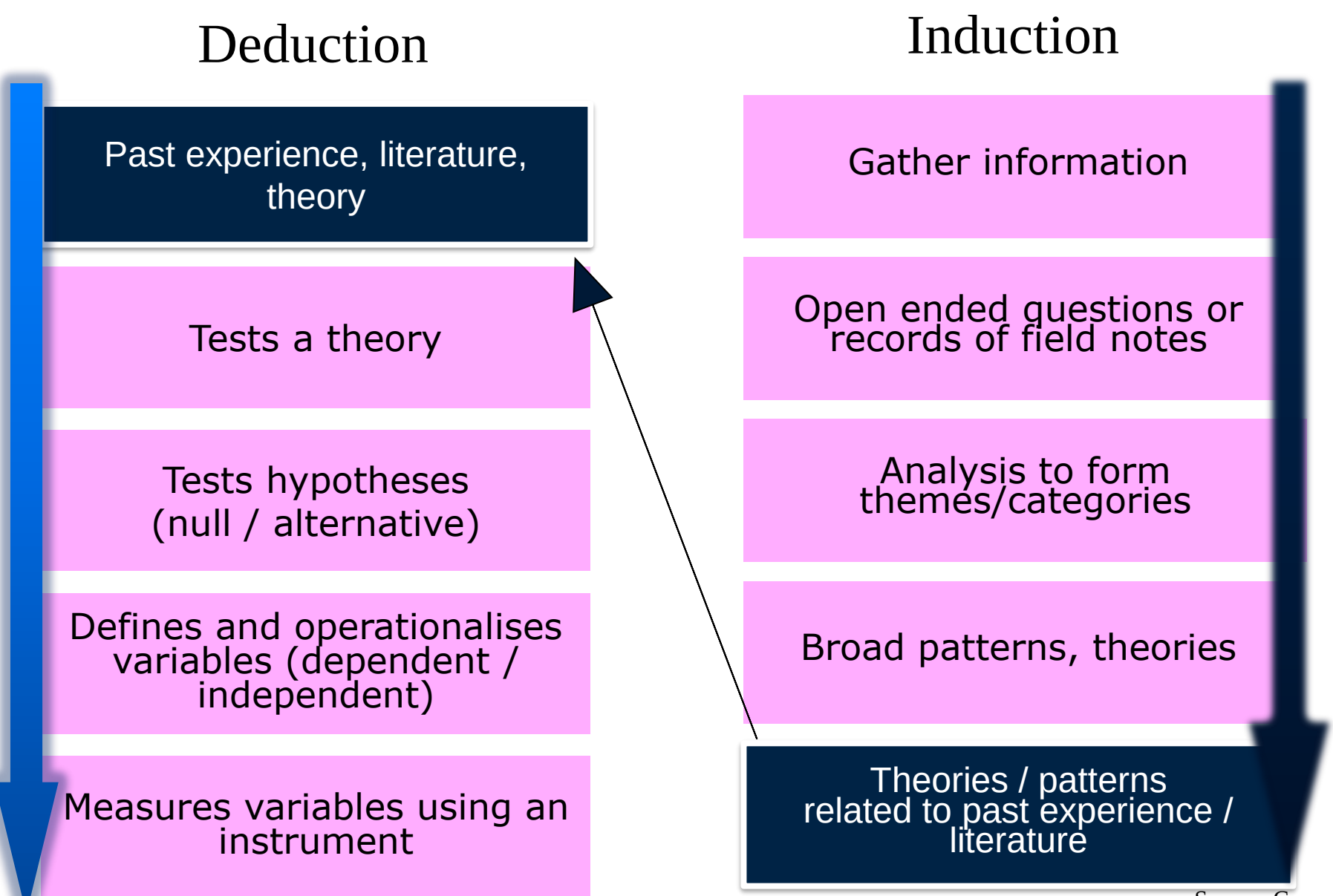
Data collection: National police literacy survey with literacy assessments

Data analysis: descriptive statistics, multinomial logistic regression

n = 8883 patrol men and women in 27 provinces



Two modes of inference

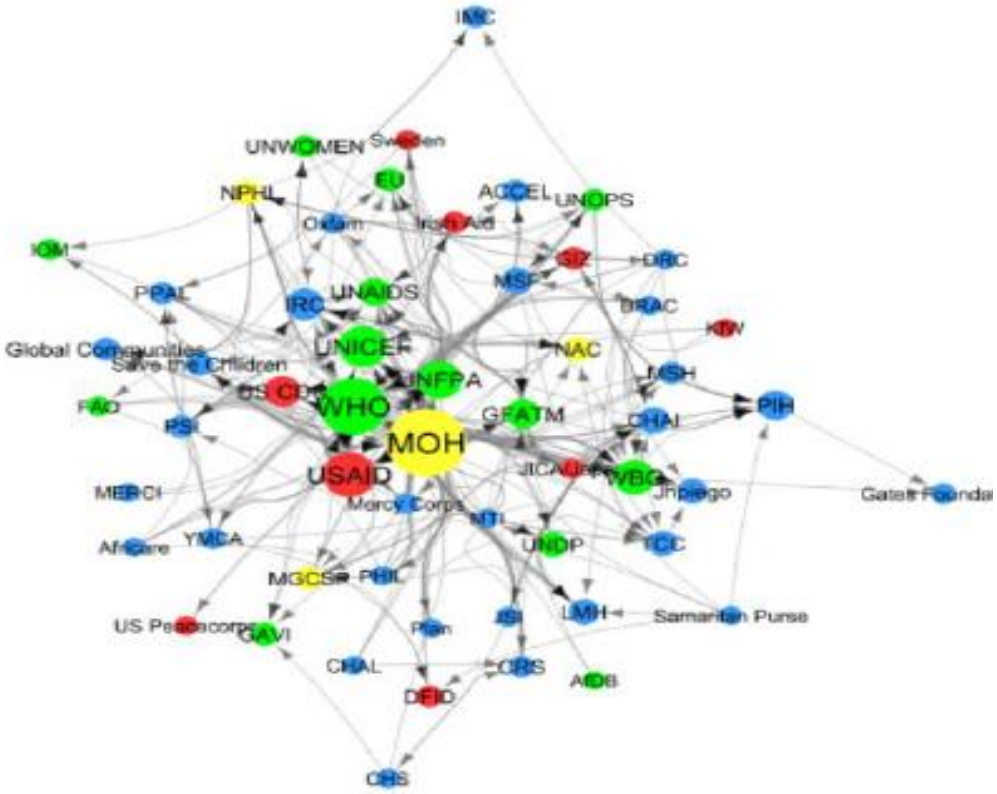


Source: Creswell, 2003

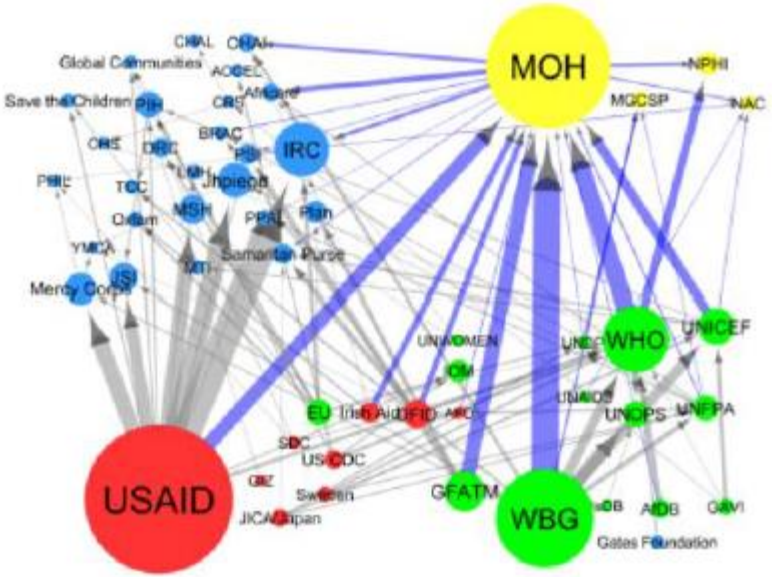
Deductive and inductive approach

1. How did outreach evolve? Was there increased outreach among the rural poor?
2. What are the factors that explain outreach/access?
3. What are the implications for poverty alleviation?

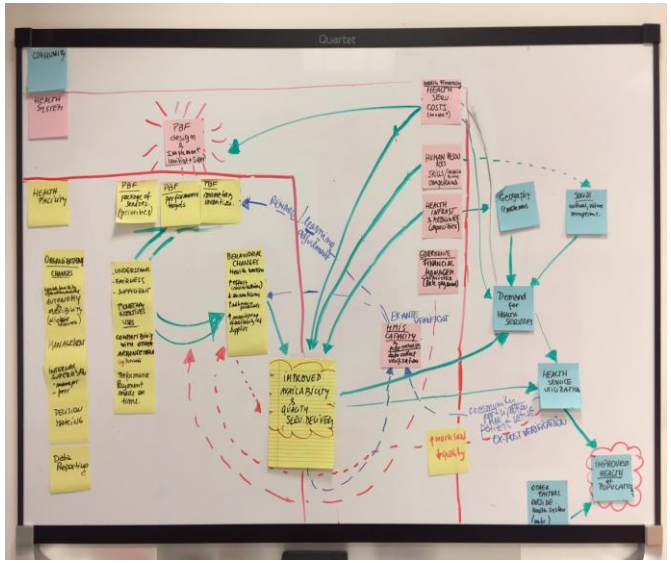
Grounded theory using network analysis

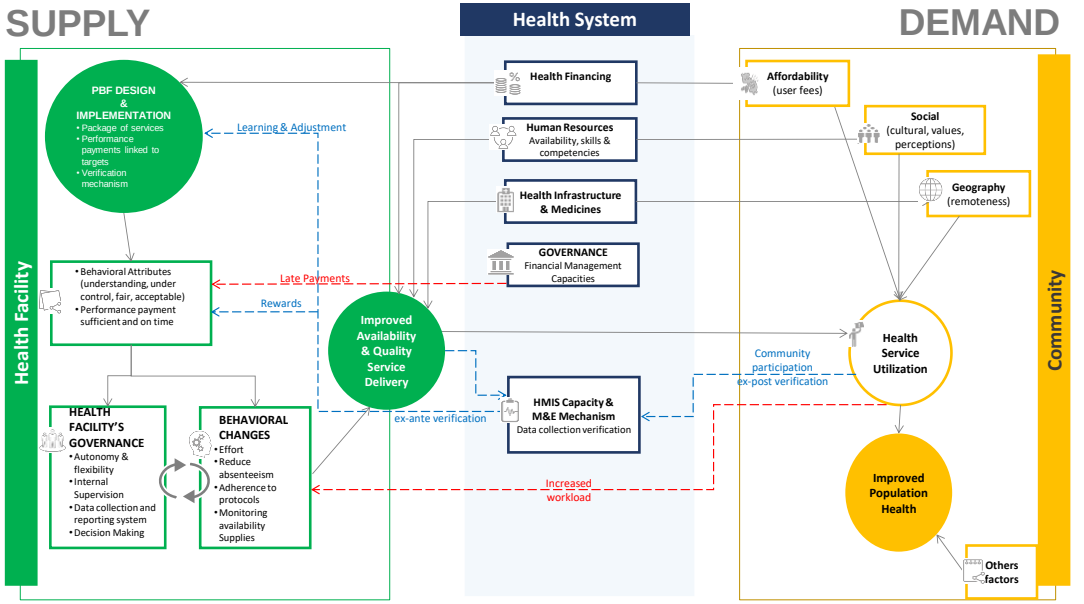


Knowledge leadership in the Health Sector in Liberia



Financial Flows in the Health Sector in Liberia





E.g, Recalibrating a theory of change based on literature

PES regional project: experimental design

Random assignment of farmers to groups receiving different incentives

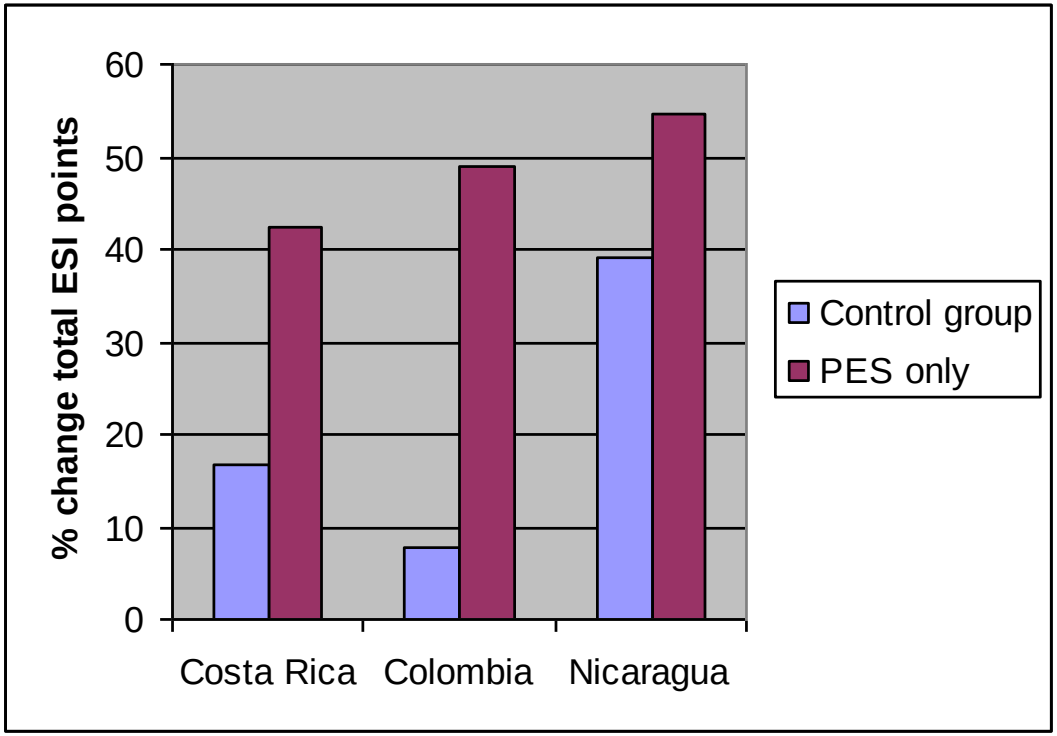
Examples of group comparisons:

PES group – control group → ***What is the effect of PES on LU changes?***

PES 4yrs group – PES 2yrs group → ***What is the effect of payment modality on LU changes?***

Some results: PES group – control group

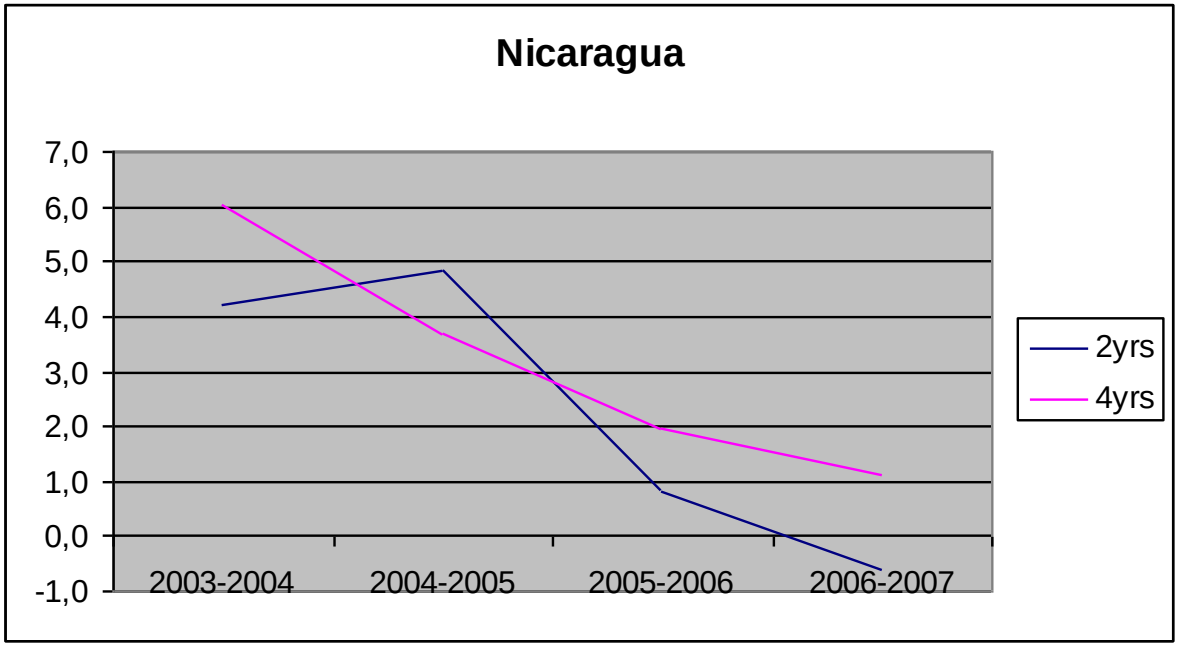
Figure 2. Incremental ESI points per hectare (2003-2007), three countries



Source: own calculations based on RISEMP project data, January 2008
Note: PSA refers to *pagos por servicios ambientales*, or PES.

Some results: PES 4yrs group – PES 2yrs group

Figure 3. Adoption behavior of the PES2yrs vs. PES4yrs group (average incremental points per farmer per year, in relation to previous year), Nicaraguan pilot site



Source: own calculations based on RISEMP project data, January 2008

Importance of a mixed methods approach: the logic of comparative advantages

The randomized experiment can test effectiveness of different incentives (PES and TA) on LU changes and subsequently the environmental and socio-economic effects of these changes *(internal validity)*

Survey data ('sub-group') analysis and case studies can tell us how incentives have *different* effects on *particular types* of farm households *(strengthens internal validity and increases external validity of findings)*

Semi-structured interviews and focus group conversations can tell us more about the nature of effects in terms of production, consumption, poverty, etc. *(construct validity)* as well as possible unintended effects (e.g. displacement effects)

Theory-based evaluation and causal analysis

Different important causal questions

Overall impact question	**Did the intervention make a difference?**		
Specific impact question	How much of a difference (on average)?	For whom? Under what circumstances?	How? Why so?
Causal question	Can we attribute the marginal (net) effect to the intervention? What is the net effect of other factors?	What role did the intervention play in producing the outcome?	What explains the outcome?
Causal theory	"Counterfactual"	"Multiple conjunctural"	"Generative" or "mechanism based"
Methods	e.g., (quasi) Experiments, stat modeling	e.g., Pattern-matching, QCA,….	e.g., Process tracing, in-depth case study

Source: Adapted from Befani, 2016 p. 20

Final note of caution: the danger of thinking inside the box and the importance of empirical evidence

- Theories are biased
- The importance of unintended effects
- Without proper empirical analysis theories may reinforce cognitive bias
- The stronger the ‘paradigm’ or ‘cognitive bias’ the stronger the need for rigorous empirical analysis

- **Conclusion:** be clear about whose theory you are reconstructing/evaluating and if possible use multiple theories in order to understand/evaluate program realities