

Evaluating governance through theories of change

Howard White
International Initiative for Impact
Evaluation

Main points

- It would be wrong to say that ‘you can’t measure governance’ or ‘you can’t measure the impact of complex interventions’
- A good theory of change approach will greatly strengthen the evaluation

Overview

- Introduction
- Measuring governance
- Rigorous impact evaluation
- Using theories of change to explain impact
- Summary

Exercise

- Define governance

World Bank definition

Good governance is epitomized by predictable, open and enlightened policy-making, a bureaucracy imbued with a professional ethos acting in furtherance of the public good, the rule of law, transparent processes, and a strong civil society participating in public affairs. Poor governance (on the other hand) is characterized by arbitrary policy making, unaccountable bureaucracies, unenforced or unjust legal systems, the abuse of executive power, a civil society unengaged in public life, and widespread corruption

Definition: key elements

- Rule of law
- Absence of corruption
- Free participation in political life
- Active civil society
- Bureaucracy functions in rule-based, transparent manner
- Policy-making for public good in non-arbitrary manner

Can we measure these things?

Measuring governance

- National level
 - Transparency International
 - Freedom House
 - Polity IV
 - World Bank WGI
- Intervention level, project examples:
 - Nigeria elections
 - Indonesia community-managed infrastructure
 - Decentralization in Philippines and Paraguay

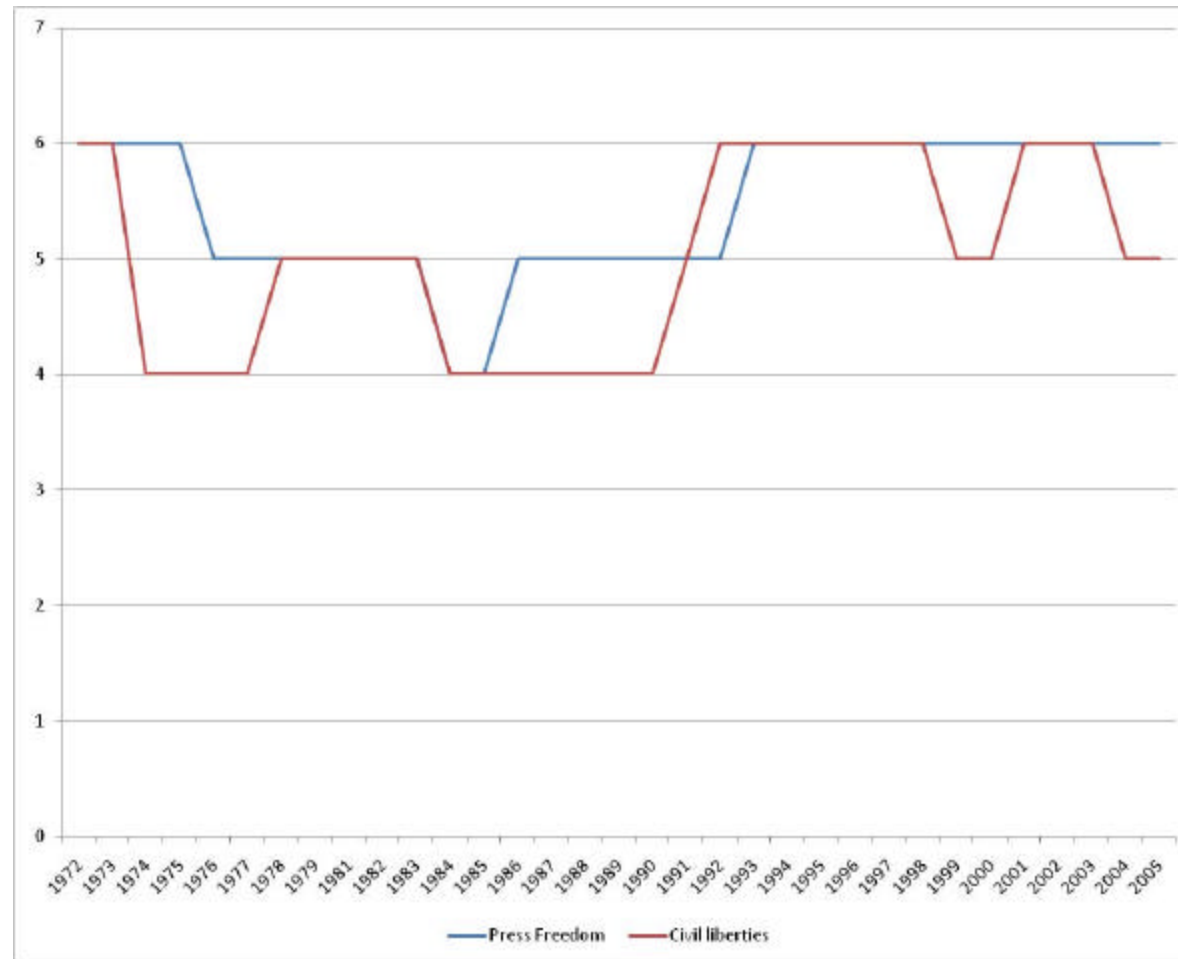
Transparency International

- Measure of corruption
- Use questionnaire, e.g. perceptions of government anti-corruption policies, if anyone in household paid bribe, cover government, private sector and land ownership

Freedom House

- Political rights: the right to vote freely, compete for public office, join political parties and organizations, and elect representatives who have a decisive impact on public policies and are accountable.
- Civil liberties: freedom of expression and belief, associational and organizational rights, rule of law, and personal autonomy without interference from the state.
- Both scored on scale 1 (Free) to 7 (Not free)

Freedom House - Egypt

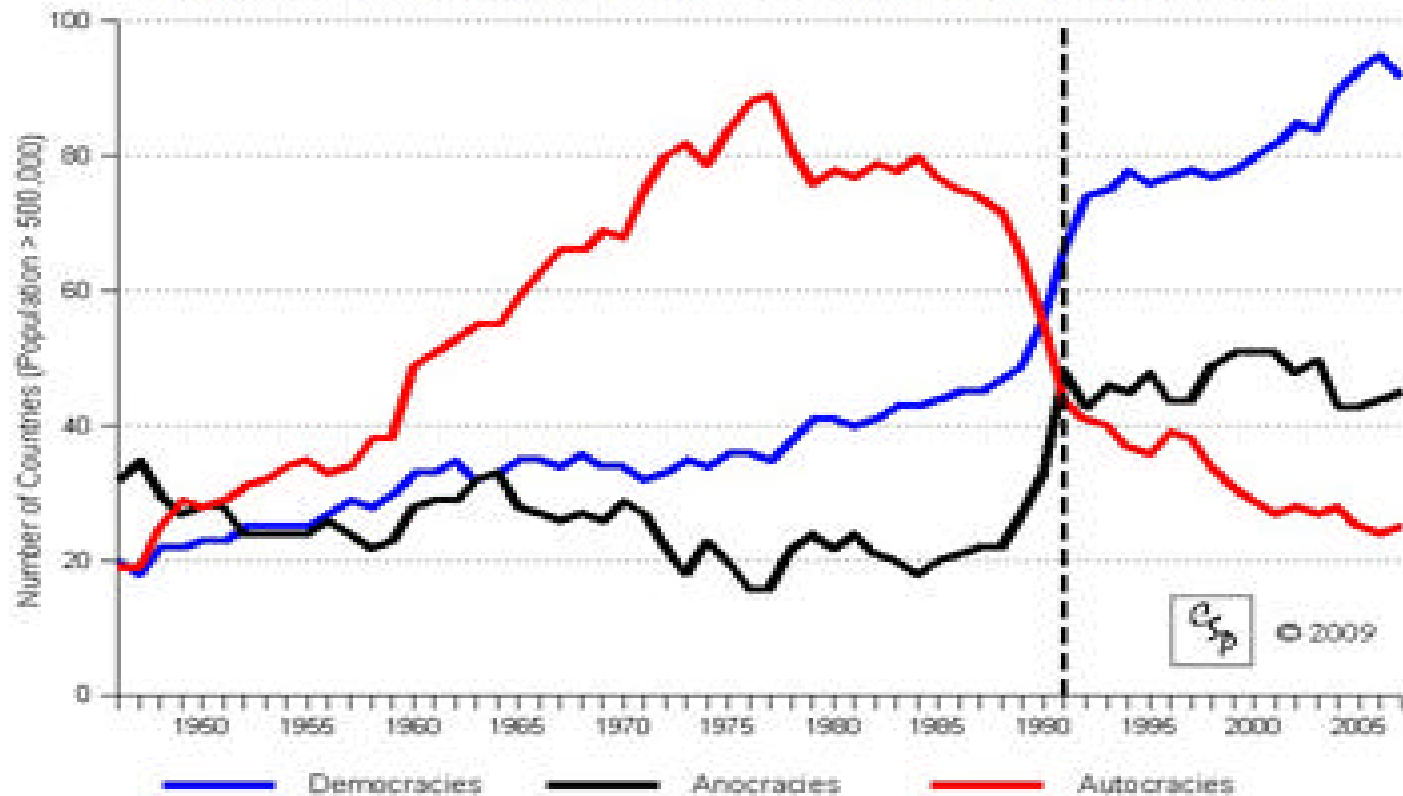


Polity IV

- Six component measures that record key qualities of executive recruitment, constraints on executive authority, and political competition
- "Polity Score" 21-point scale ranging from -10 (hereditary monarchy) to +10 (consolidated democracy)
- Regime categories: three-part categorization of "autocracies" (-10 to -6), "anocracies" (-5 to +5), and "democracies" (+6 to +10)

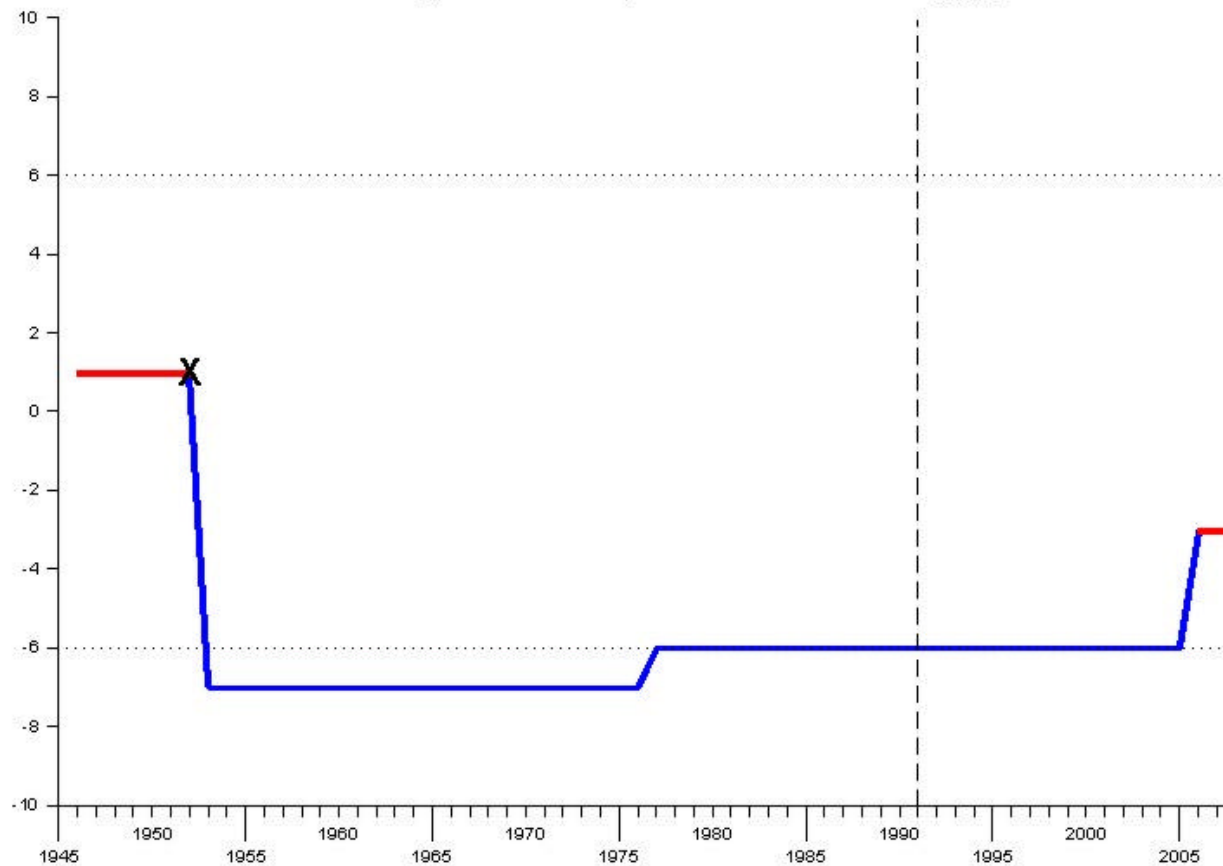
Polity 4 – global summary

Global Trends in Governance, 1946-2007



Polity IV data - Egypt

Authority Trends, 1946-2007: Egypt



Worldwide Governance Indicators

- Covering 212 countries and territories
- Six dimensions of governance: Voice and Accountability, Political Stability and Absence of Violence, Government Effectiveness, Regulatory Quality, Rule of Law, and Control of Corruption.
- Taken from 33 data sources provided by 30 different organizations.
- Data reflect the views on governance of public sector, private sector and NGO experts, as well as thousands of citizen and firm survey respondents worldwide.

Intervention level

- Defining indicators for impact evaluation
- Will come back to evaluation approaches
- And then theories of change

Impact evaluations of governance interventions

- Three RCTs
 - Elections in Nigeria
 - Community management in Indonesia
 - National Solidarity Program in Afghanistan
- Other approaches
 - Decentralization in the Philippines
 - Decentralizing health in Paraguay
 - Firm competitiveness and corruption

Elections in Nigeria

- Campaign to reduce pre-election violence and vote rigging
- Campaign by ActionAid, research by Oxford University
- 24 enumeration areas across country, intervention randomly assigned to half of them
- Results
 - Threat of violence does reduce turnout
 - Campaign reduced pre-election violence
 - Political influence and firm competitiveness

Community infrastructure in Indonesia

- Community-implemented infrastructure (roads) in 608 villages
- Corruption measure: difference project cost and engineer estimate of cost
- Three experiments: 100% audit, invitations to “accountability meeting”, and anonymous comment form
- Auditing cut “missing expenditures” from 24 to 16%; no impact from participatory experiments (though participation did increase)

National Solidarity Program in Afghanistan

- 250 project villages and 250 control: randomized
- Evaluation questions include:
 - What is the impact of NSP on structures and perceptions of local governance?
 - What is the impact of NSP on the participation of women on governance?
 - What is the impact of NSP on interpersonal trust and political participation?

Decentralization in the Philippines

- Output measure: membership local organizations and participation in local projects
- Uses propensity score matching, shows that making available information on service delivery increases likelihood of participation
- 12 districts, but matching done at household level

Decentralization health in Paraguay

- 20 municipalities, 11 decent, 9 control
- Interviews: public and private facilities, 1261 exit interviews, 1200 household surveys.
- Baseline data and after 2 years
- Outcomes: impact of decentralization on costs, efficiency, basic quality, pattern of use and equity in the use

Firm competitiveness

- More influential firms face fewer administrative and regulatory obstacles, and carry bloated payrolls, but they also invest and innovate less. These results hold up when use propensity score matching models to adjust for the fact that influence is not randomly assigned.
- Data: firm-surveys from 50,000 firms in 60 countries

Exercise

- Program
 - National Anti-corruption commission (A)
 - Program to increase women's participation in local government (B)
 - Program to economically empower women in rural areas (C)
- Exercise
 - Select 3 output indicators and 3 outcome indicators
 - How would you collect data on these six indicators; in the case of a questionnaire, describe the questions

Example

- Program to increase community role in school management
- Output indicators: knowledge of community members about school management committee, functioning school management committees (number of meetings held each year); volume of resources spent through community committees; knowledge of committee members about school matters;
- Outcome indicators: head teacher perception of community role; changes in school practices as a result of SMC decisions; efficiency of school management (need to define)

Data collection: example

- Household survey
 - Who is responsible for decisions regarding the management of the school? (Government, Village leaders, Head teacher, SMC, Other)
 - Are you aware of any committee or any group in the community that has responsibility for school management
 - Who is the chair of the school management committee?
- Note: structured to avoid 'leading question bias'

Rigorous impact evaluation

- Combines rigorous quantitative and qualitative analysis
- Quantitative methods to get a counterfactual using a credible control group
 - Experimental (RCTs)
 - Quasi-experimental (PSM, RDD, regression-based)
- Qualitative methods to help construct factual analysis
- These methods are for large n studies; if small n then need more qualitative approach (stakeholder analysis), but should still have theory of change

Example: school management committees

- Evaluation design for counterfactual analysis, either
 - Experimental – randomly assign schools to treatment and control
 - Quasi-experimental – match schools using statistical procedures
- Important elements in factual analysis
 - Knowledge of SMCs, who participates?
 - What SMCs are meant to do and how
 - Describing how SMCs interact with head and school staff
 - Resource availability

Exercise

- Program
 - National Anti-corruption commission (A)
 - Program to increase women's participation in local government (B)
 - Program to economically empower women in rural areas (C)
- Would randomization be possible in these cases? How else might you construct a control group?

Theory of change approach (theory-based impact evaluation): principles

- Map out the causal chain (programme theory)
- Understand context
- Anticipate heterogeneity
- Rigorous evaluation of impact using an appropriate counterfactual
- Rigorous factual analysis
- Use mixed methods

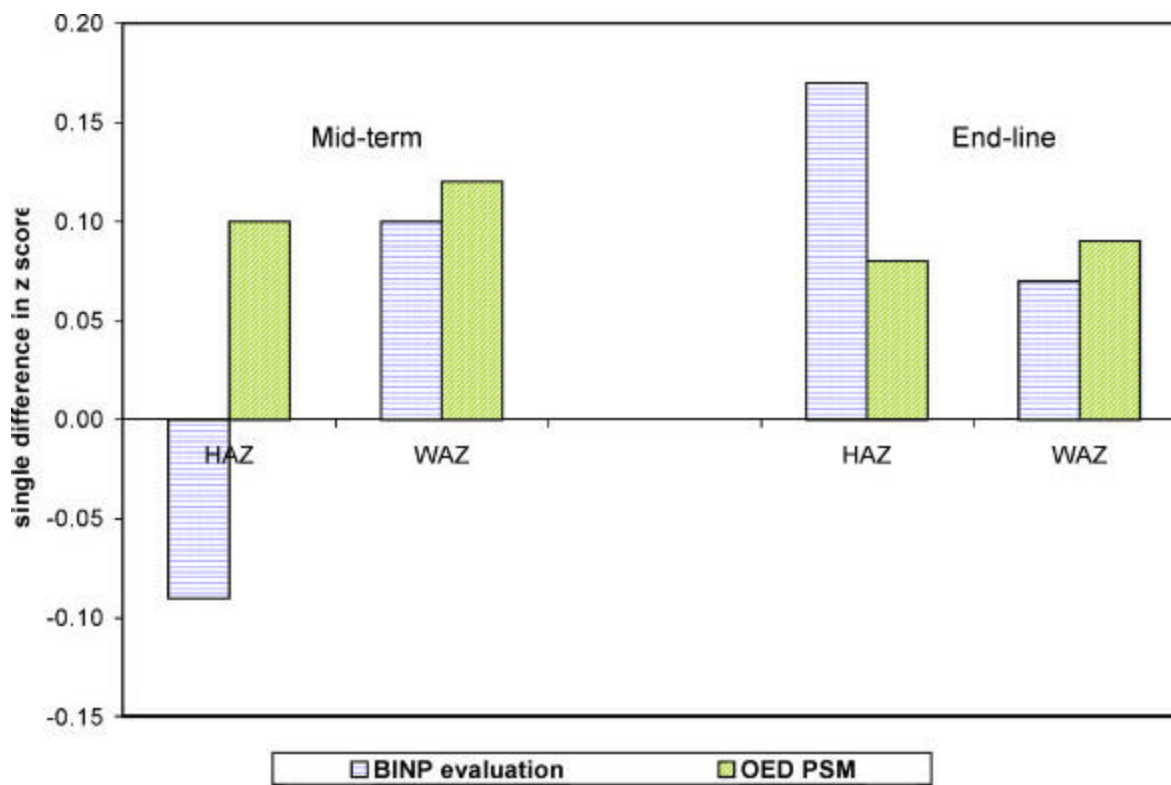
Theory-based impact evaluation: an example

- Bangladesh Integrated Nutrition Project (BINP)
- Growth monitoring, nutritional counselling and supplementary feeding (based on TINP)
- Implemented by NGOs at field level, using Community Nutrition Practitioners (CNPs)
- Pilot project, succeeded by National Nutrition Project

The evaluation story

- Looked like it was working – all bits in place and monitoring data showed sharp fall in severe malnutrition
- Bank agreed to scale up
- But Save the Children UK critical, though Bank's evaluation positive
- Bank's evaluation department (IEG) did evaluation – found little or no impact
- Theory-based approach explains why

Impact estimates (using propensity score matching)



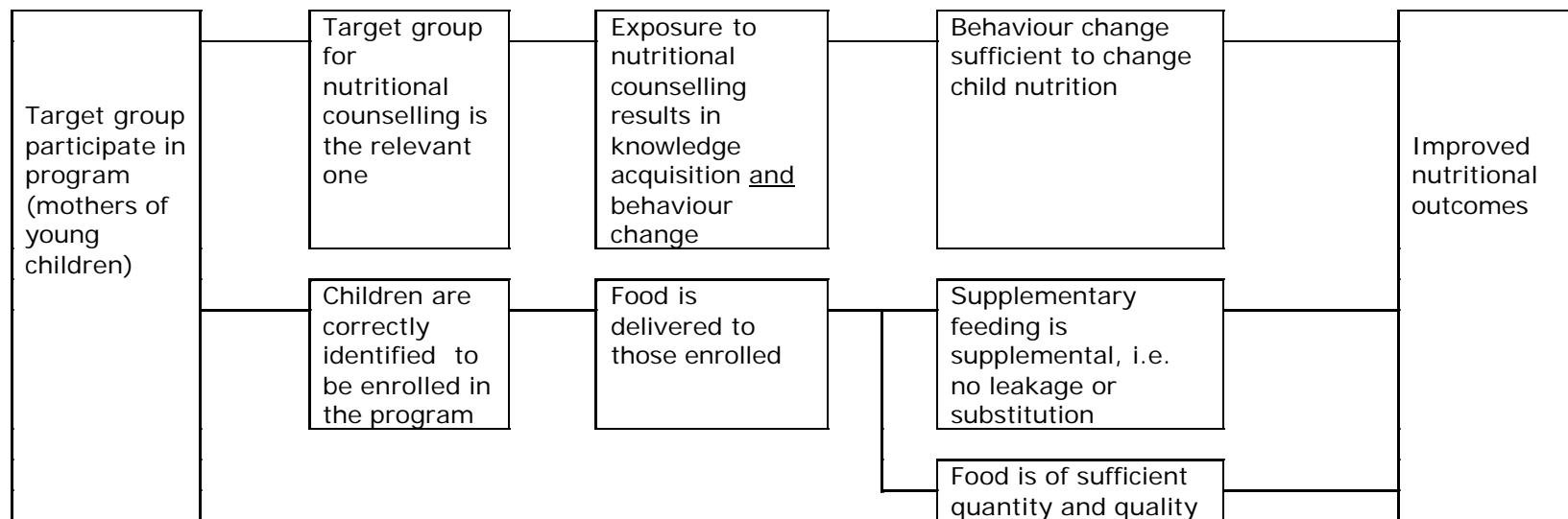
Some underlying assumptions: nutritional counselling

1. The right people (those making decisions regarding under-nourished children) are targeted with nutritional messages.
2. These people participate in project activities, and so are exposed to these messages
3. Exposure leads to acquisition of the desired knowledge
4. Acquisition of the knowledge leads to its adoption (i.e. a change in practice)
5. The new practices make a substantial impact on nutritional outcomes

Some underlying assumptions: supplementary feeding

1. The target groups have to enroll in the programme
2. The criteria are correctly applied in selecting those to receive supplementary feeding
3. Those selected for supplementary feeding attend sessions to receive the food.
4. There is no leakage (e.g. selling of food supplements or giving food to other family members), or substitution (reducing other food intake)
5. The food is of sufficient quantity and quality to have a noticeable impact on nutritional status (and, in the case of pregnant women, to be of sufficient magnitude to have an appreciable impact on birth weight).

Summary of theory

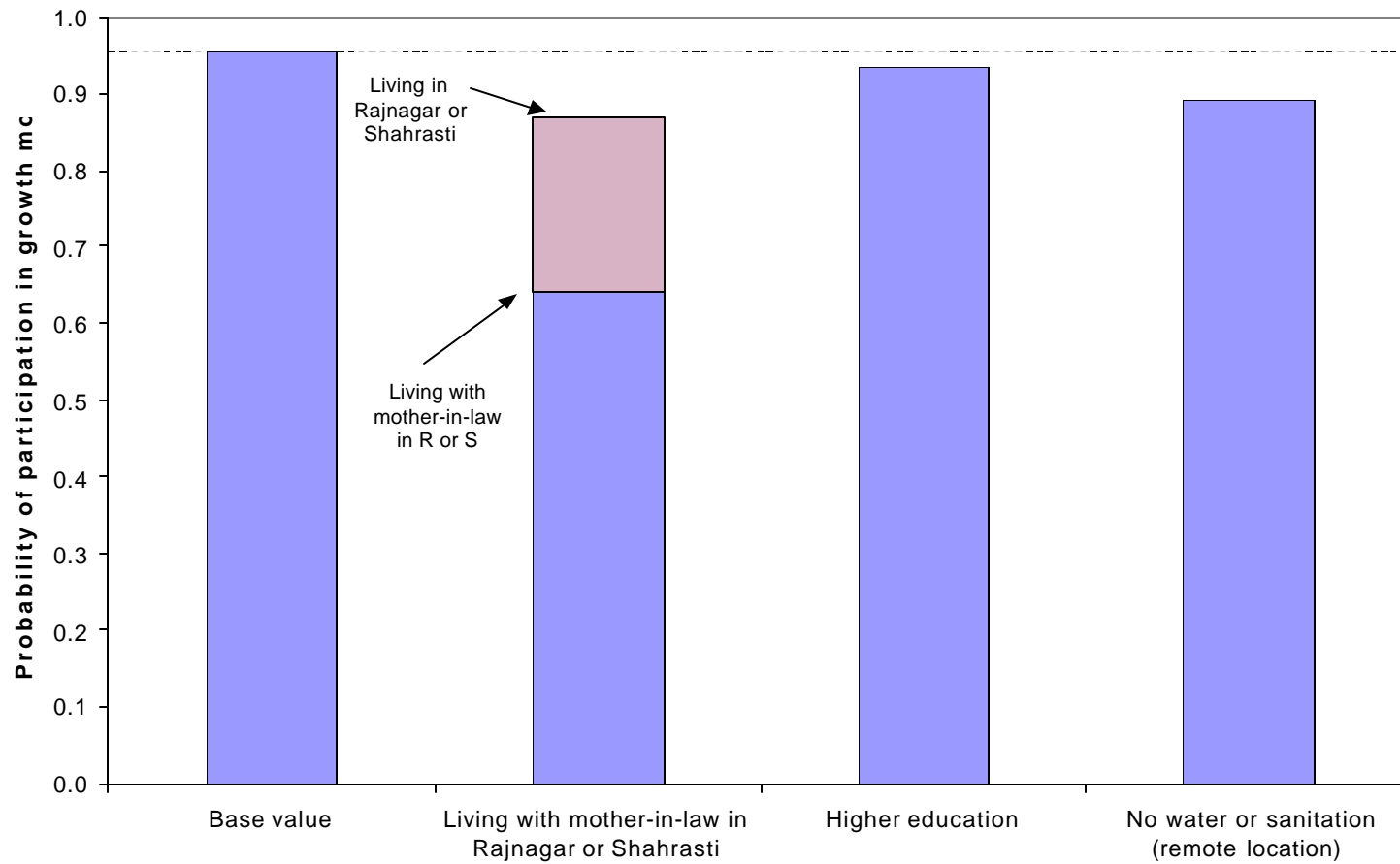


Policy relevant IEI: Theory-based IE: the Bangladesh Integrated Nutrition Project

Assumption	Findings
Provide nutritional counseling to care givers	Mothers are <u>not</u> decision makers, especially if they live with their mother-in-law
Women know about sessions and attend	90% participation, lower in more conservative areas
Malnourished and growth faltering children correctly identified	No – community nutrition practitioners cannot interpret growth charts
Women acquire knowledge	Those attending training do so
And knowledge is turned into practice	No there is a substantial knowledge-practice gap
Supplementary feeding is additional food for intended beneficiary	No, considerable evidence of substitution and leakage
Adopted changes are sufficient to improve intended outcomes	Only sometimes (not for pregnant women)

Source: Howard White and Edoardo Masset (2007) 'The Bangladesh Integrated Nutrition Program: findings from an impact evaluation' *Journal of International Development* 19: 627-652

Participation rates



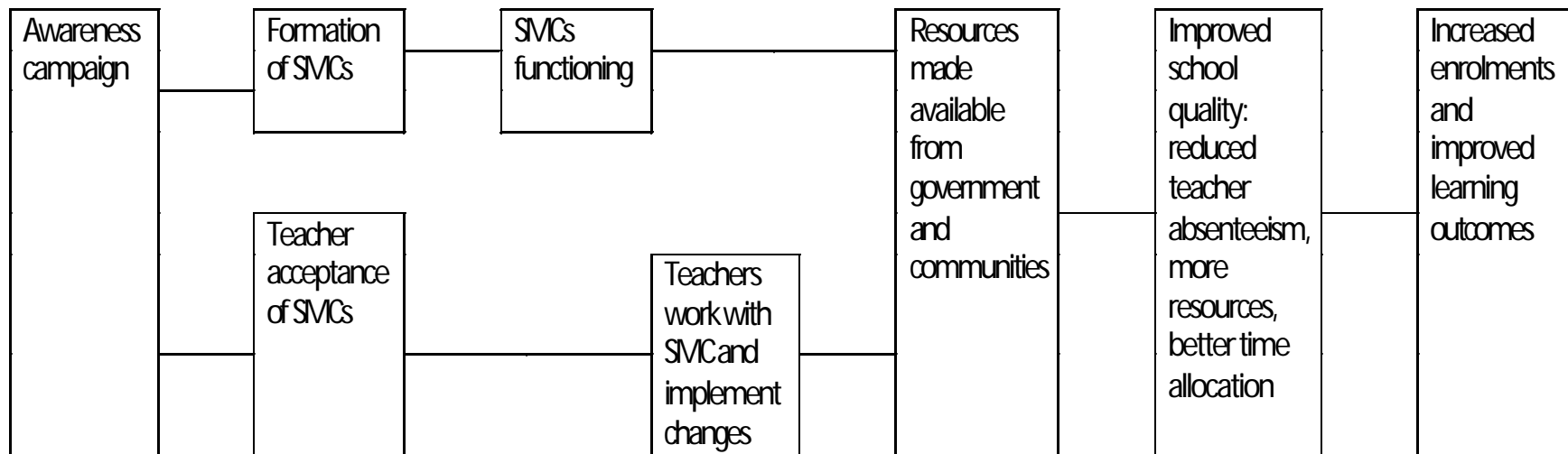
Illustrating the principles

- Map out the causal chain (programme theory): see figure
- Understand context: Bangladesh is not TN
- Anticipate heterogeneity: more malnourished children; different implementing agencies
- Rigorous evaluation of impact using an appropriate counterfactual: PSM versus simple control
- Rigorous factual analysis: targeting, KP gap, CNPs
- Use mixed methods: informed by anthropology, focus groups

Applying to governance: school management committees

- Main theory
 - Accountability
 - Efficient management (e.g. time allocation)
- Some assumptions
 - Parents know about SMCs and understand their role
 - Parents willing to participate
 - Parents have appropriate knowledge
 - Teachers willing to cooperate
 - Accountability mechanisms function

Overview of theory



Exercise

- Program
 - National Anti-corruption commission (A)
 - Program to increase women's participation in local government (B)
 - Program to economically empower women in rural areas (C)
- Sketch out the appropriate program theory. What quantitative and qualitative data would you collect to test this theory?

Thank you

Visit www.3ieimpact.org